12th Canadian Symposium on Remote Sensing
12e Symposium Canadien sur Teledetection

1989 INTERNATIONAL GEOSCIENCE AND REMOTE SENSING SYMPOSIUM (IGARSS '89)

and URSI Commission F

IGARSS'89

12TH CANADIAN SYMPOSIUM ON REMOTE SENSING
12e SYMPOSIUM CANADIEN SUR LA TELEDETECTION

Quantitative Remote Sensing:

An Economic Tool for the Nineties

AD-A223 845

N00014-89-J-1969

DTIC
ELECTE
JUL 06 1990
S D
D

(1)

DTIC FILE COPY

DISTRIBUTION STATEMENT A
Approved for public release
Distribution Unlimited

IEEE

U.R.S.

VANCOUVER, CANADA JULY 10-14, 1989

Volume 4. Thursday July 13

90 07 3 227

Printed in Canada

LIBRARY OF CONGRESS #89-84217
IEEE #89CH2768-0

VOLUME 4

IEEE No. 89CH2768-0
LIBRARY OF CONGRESS No. 89-84217

IGARSS '89

12th CANADIAN SYMPOSIUM ON REMOTE SENSING
12e SYMPOSIUM CANADIEN SUR LA TELEDETECTION

REMOTE SENSING:

AN ECONOMIC TOOL FOR THE NINETIES

JULY 10 - 14 th 1989 VANCOUVER, CANADA

12th Canadian Symposium on Remote Sensing
12e Symposium Canadien sur Teledetection
1989 INTERNATIONAL GEOSCIENCE AND
REMOTE SENSING SYMPOSIUM (IGARSS '89)
and URSI Commission F Meeting

IEEE

U.R.S.I.

CANADIAN REMOTE SENSING SOCIETY
SOCIETE CANADIENNE DE TELEDETECTION

SPONSORS:
IEEE, CRSS, URSI

AVAILABLE FOR $186 FOR SET per Institute of Electrical and Electronics Engineers, Inc. 345 East 47th St., New York, NY 10017
TELECON 7/5/90 VG

Accesion For
NTIS CRA&I ☑
DTIC TAB ☐
Unannounced ☐
Justification
By 186.00 per set /per vol
Distribution 1-5
Availability Codes
Dist | Avail and / or Special
A-1 | 21

QUALITY INSPECTED 2

IGARSS '89
REMOTE SENSING: AN ECONOMIC TOOL FOR THE NINETIES
VANCOUVER, CANADA
JULY 10-14, 1989

COPYRIGHT - 1989 by IEEE

Proceedings published by:
IGARSS '89 12th Canadian Symposium on Remote Sensing

Printed in Canada

IGARSS '89 - 12th CANADIAN SYMPOSIUM ON REMOTE SENSING
IGARSS '89 - 12th SYMPOSIUM CANADIEN SUR LA TELEDETECTION

IGARSS '89 - 12th CANADIAN SYMPOSIUM ON REMOTE SENSING SPONSORS SPONSORS COMMANDITAIRES

IGARSS ' 89 - 12th Canadian Symposium on Remote Sensing is sponsored by the Canadian Remote Sensing Society, the IEEE Geoscience and Remote Sensing Society, the International Union of Radio Science, and the Canada Centre for Remote Sensing.

It is co-sponsored by:

B.C. Ministry of Environment and Parks
B.C. Ministry of Forests, Inventory Branch
B.C. Research Corporation
Binary Image Corporation
Canadian Forest Service
Cominco Ltd.
Fisheries and Oceans Canada
MacDonald, Dettwiler and Associates Ltd.
National Research Council Canada
Pamap Graphics Ltd.
University of British Columbia

IGARSS '89 - 12th CANADIAN SYMPOSIUM ON REMOTE SENSING ORGANIZING COMMITTEE

Dr. John S. MacDonald, General Chairman

Dr. Ed Jull, Vice Chairman, URSI
Dr. Peter Murtha, Vice Chairman, CRSS
Dr. Keith Raney, Vice Chairman, IEEE/GRS-S

Dr. Jim Gower, Technical Program Chairman

Dr. Josef Cihlar, Technical Co-Chairman, CRSS
Dr. David Goodenough, Technical Co-Chairman, IEEE/GRS-S

M. Jean-Marie Dubois, French Transcription
Mr. Frank Hegyi, Social Program and Tours
Ms. Nedenia Holm, Secretary
Dr. Y. Jim Lee, Exhibits
Mr. Peter Louis, Finance
Dr. Pam Sallaway
Dr. Hans Schreier, UBC Facilities/Registration
Dr. Mark Sondheim, Publicity
Dr. Mit Tilkov
Dr. Geoff Tomlins, Publication

CRSS Technical Committee

Chair: J. Cihlar, Canada Centre for Remote Sensing, Canada
F.J. Ahern, Canada Centre for Remote Sensing, Canada
Kohei Arai, National Space Development Agency of Japan, Japan
F. Bonn, Universitie de Sherbrooke, Canada
G.A. Borstad, G.A. Borstad Associates Ltd., Canada
J. Brown, Canada Centre for Remote Sensing, Canada
R. Currie, Pacific Geoscience Centre, Canada
G. Duchossois, European Space Agency, Canada
A. Fabbri, Canada Centre for Remote Sensing, Canada
B.E. Goodison, Atmospheric Environment Service, Canada
B. Guidon, Canada Centre for Remote Sensing, Canada
F.G. Hall, NASA Goddard Space Flight Centre, USA
D. Halpern, Jet Propulsion Laboratories, USA
F. Hegyi, Ministry of Forests, Canada
J. Iisaka, Canada Centre for Remote Sensing, Canada
J. Kosalos, Honeywell Marine Systems, USA
L. McNutt, Canada Centre for Remote Sensing, Canada
J.B. Mercer, Intera Technologies Ltd., Canada
R.E. Murphy, NASA/Headquarters, USA
R.A. O'Neil, Canada Centre for Remote Sensing, Canada
P. Sallaway, PAMAP Graphics Ltd., Canada
L. Sayn-Wittgenstein, Canada Centre for Remote Sensing, Canada
R. Slaney, Radarsat Project Office, Canada
W.M. Strome, PCI Inc., Canada
T. Toutin, Canada Centre for Remote Sensing, Canada
B. Woodham, University of British Columbia, Canada
V. Zsilinszky, Ontario Centre for Remote Sensing, Canada

IEEE/GRSS Technical Committee

Chair: David G. Goodenough, Canada Centre for Remote Sensing, Canada
R.C. Beal, The John Hopkins University, USA
G. Brachet, SPOT Image, France
F. Carsey, Jet Propulsion Laboratories, USA
K.R. Carver, University of Massachusetts, USA
R. Cheney, National Ocean Service NOAA, USA
J. Cimino, Jet Propulsion Laboratories, USA
M. Ehlers, University of Maine, USA
A.F.H. Goetz, University of Colorado, USA
K.I. Itten, University of Zurich-Irchel, Switzerland
E. Ledrew, University of Waterloow, Canada
C.A. Luther, Office of Naval Research, USA
J.P. Muller, University College London, UK
J.A. MacDonald, MacDonald Dettwiler Associates, Canada
J. Megier, Ispra Establishment, Italy
R.K. Raney, Radarsat Project Office, Canada
P.N. Slater, University of Arizona, USA
C.T. Swift, University of Massachusetts, USA
Mikio Takagi, University of Tokyo, Japan
G.F. West, University of Toronto, Canada

URSI Technical Committee

Chair: J.F.R. Gower, Institute of Ocean Sciences, Canada
J.R. Apel, The John Hopkins University, USA
R.K. Crane, Darmouth College, USA
J.L. Green, NOAA Aeronomy Laboratory, USA
A. Guissard, Universitie Catholilque de Louvain, Belgium
D.C. Hogg, University of Colorado, USA
R. Humphreys, MacDonald Dettwiler Associates, Canada
E. Jull, University of British Columbia, Canada
R.K. Moore, University of Kansas Centre for Research Inc., USA
D.W. Oldenburg, University of British Columbia, Canada
R. Olsen, Communications Research Centre, Canada
B. Segal, Communications Research Centre, Canada
O.H. Shemdin, Ocean Research and Engineering, USA
J.R. Wait, University of Arizona, USA
J. Walsh, Memorial University of Newfoundland, Canada

Book 4
Table of Contents

	Page
Satellite based drought surveillance in India - Some achievements S. Thiruvengadachari, Nat. Rem. Sens. Agency, Hyderabad, India	1961
Monitoring drought affected vegetation with AVHRR E.W. Allison, Prairie Farm Rehab. Admin., Regina, Sask., Canada, R.J. Brown, CCRS, Ottawa, Ont., Canada, H. Press, J. Gairns, INTERA Tech. Ltd., Ont., Canada	1965
Monitoring rangeland change using Landsat MSS data D.W. Mullins, INTERA Tech. Ltd., Ottawa, Ont., Canada, J. Cihlar, CCRS, Ottawa, Ont., Canada	1968
Operational monitoring of forage and crop yield J. Whiting, N. Nicolichuk, Sask. Res. Council, Saskatoon, Sask., Canada, S. Benjamin, Can./Sask. Crop Ins. Corp., Melville, Canada	1972
Regional paddy monitoring system using NOAA HRPT data Y. Matsumae, T. Hosomura, H. Shimoda, T. Sakata, Tokai Univ. Research and Info. Centre, Tokyo, Japan	1975
Multi-temporal land use analysis for soil erosion in selected Prince Edward Island (P.E.I.) watersheds D.A. Wilson, Env. Canada, Dartmouth, NS, Canada, M.L. McCourt, Land Resource Info. Serv., Amherst, NS, Canada, T.M. Humes, Env. Canada, Dartmouth, NS, Canada	1979
Automatic determination of the topographic factor required in the universal soil loss equation D. Benmouffok, D.N.H. Horler, Horler Info. Inc., Ottawa, Ont., Canada, R. Leconte, CCRS, Ottawa, Ont., Canada	1986
Application of remotely sensed and Geographic Information System data, to quantitative assessment of landslide damage C.M. Trotter, P.R. Stephens, N.R. Trustrum, M.J. Page, K.S. Carr, R.C. de Rose, Division of Land and Soil Sci., DSIR, Palmerston N., New Zealand	1990
Using satellite information as a tool for range management J. Whiting, J. Polson, Sask. Res. Council, Saskatoon, Sask., Canada, D. McCartney, Can. Agric., Melfort, Sask., Canada	1995
An assessment of TM imagery for land cover change detection T. Fung, Chinese Univ. of Hong Kong, Statin, Hong Kong	1998
Utility of spatial filtering techniques for detecting soil erosion on Landsat Thematic Mapper images of Iran M.N. Disfani, Univ. of Glasgow, Glasgow, UK	2002
The statistics Canada crop condition assessment program K.D. Korporal, R. Dobbins, N.M. Hillary, Statistics Canada, Ottawa, Ont. Canada	2006

Page

Operational GIS: past, present and future 2010
F. Hegyi, Min. of Forests, Victoria, BC, Canada

Specifications for GIS and image analysis data: A government perspective 2011
P. Friesen, M. Sondheim, British Columbia Min. of Crown Lands, Victoria, BC, Canada

Integrating satellite imagery with other spatial data sets for monitoring the downstream impacts of dam construction in Nigeria 2015
P.G. Pilon, P.J. Howarth, Univ. of Waterloo, Waterloo, Ont., Canada

GIS data integration for geological investigation in the South Mountain Batholith near Halifax, Nova Scotia 2019
S.M. Shupe, M.S. Akhavi, Nova Scotia Coll. of Geog. Sci., Lawrencetown, NS, Canada

Meeting GIS database requirements for natural topography resource management through the integration of digital and satellite imagery 2023
M. Gray, E. Kenk, R. Yazdani, BC Min. of Crown Lands, Victoria, BC, Canada

The integration of remote sensing data with a Geographic Information System to develop forest cover parameters for forest wildlife impact assessment in Southern Ontario 2027
C.L. Wagner, R.N. Pierce, S.W. Hounsell, Ontario Hydro, Toronto, Ont., Canada

Integrated GIS and image analysis software: application in vegetation monitoring 2030
P. Landriau, Dipix Tech. Inc., Ottawa, Ont., Canada, M.J. Manore, CCRS, Ottawa, Ont., Canada

The use of remotely sensed data in the creation of a Geo-database for Antigonish, Nova Scotia utilizing a micro-computer based GIS 2034
T. Webster, TYDAC Tech. Inc., Ottawa, Ont., Canada, M.S. Akhavi, Nova Scotia Coll. of Geog.,Lawrencetown, Canada

Review of satellite applications to fisheries 2037
R.M. Laurs, NOAA Natl. Mar. Fish. Serv., La Jolla, CA, USA

Remote sensing and modern fishing technology - Its impact on conservation and protection of the fisheries resource 2040
J.C. Davis, Inst. of Ocean Sciences, Sidney, BC, Canada

Comparison of ship and aerial surveys of sardine school 2041
Ichiro Hara, Seikai Regional Fish. Res. Lab., Nagasaki, Japan

Remote sensing of fish schools: early results from a digital imaging spectrometer 2044
B.S. Nakashima, Dept. of Fish. & Oceans, St. John's, Nfld., Canada, G.A. Borstad, D.A. Hill, R.C. Kerr, G.A. Borstad Assoc. Ltd., Sidney, BC, Canada

The toxic Chrysochromulina Polylepis bloom in Scandinavian waters - May/June 1988 2048
O.M. Johannessen, L.H. Pettersson, J.A. Johannessen, P.M. Haugan, T. Olaussen, K. Kloster, Nansen Remote Sensing Center, Bergen, Norway

Developing operational fisheries oceanography in Newfoundland 2052
S.A. Akenhead, Dept. of Fish. and Oc., St. John's, Nfld., Canada

Assessing fishing effort by remote sensing in the Scotia Fundy region of Fisheries and Oceans Canada 2056
G. Sharp, J. Pringle, R. Duggan, Fish. and Oceans Canada, Halifax, NS, Canada

Page

La teledetection: un outil additionnel pour les etudes qui portent sur le recrutement larvaire de certaines especes marines 2061
A. Lavoie, Univ. de Montreal, Montreal, Que., Canada, C. Hudon, Peches et Oceans Canada, Que., Canada, J.-M. Dubois, Univ. de Sherbrooke, Sherbrooke, Que., Canada, G. Vigeant, Env. Canada, Saint-Laurent, Que., Canada

Automated spectral recognition of alteration mineralogy in imaging data using a knowledge-based expert system 2065
T. Rubin, Stanford Univ., Stanford, CA, USA

Inversion of imaging spectrometry data using singular value decomposition 2069
J.W. Boardman, Univ. of Colorado, Boulder, CO, USA

A hybrid system of data analysis for imaging spectrometry 2073
D. Jayasinghe, J. Miller, York Univ., North York, Ont., Canada

Combining spectral and spatial mode imagery from an imaging spectrometer 2077
J.D. Dunlop, Univ. of Waterloo, Waterloo, Ont., Canada

Use of the compact airborne spectrographic imager (CASI): Laboratory Examples 2081
Laboratory Examples
G.A. Borstad, D.A. Hill and R.C. Kerr, G.A. Borstad Assoc. Ltd., Sidney, BC, Canada

Changes in canopy leaf area index and biochemical constituents of a spruce forest as measured by the AIS-2 Airborne Imaging Spectrometer 2085
C. Banninger, Inst. for Im. Proc. and Comp. Graphics, Graz, Austria

Determining experimentally induced variation in coniferous canopy chemistry with airborne imaging spectrometer data 2090
N.A. Swanberg, P.A. Matson, NASA Ames Res. Cen., Moffett Field, CA, USA

Red edge measurements and canopy structure: a first look with AVIRIS data 2093
C.E. Leprieur, Lab. d'Etudes et Rech. en Teledet. Spat., Toulouse, France

Systeme de mesures de reflectances pour les spectrometres imageurs 2097
A. Royer, N.T. O'Neill, B. Williams, P. Cliche, Univ. de Sherbrooke, Sherbrooke, Que., Canada, R. Verreault, Univ. de Quebec, Chicoutimi, Qc, Canada

Analysis of scattering behavior from a known random surface using frequency diverse projection imaging 2101
C.E. Nance, A.J. Blanchard, Univ. of Texas at Arlington, Arlington, TX, USA

Multiple volume scattering effects in microwave polarimetric remote sensing 2104
L. Tsang, K.H. Ding, B. Wen, D. Winebrenner, Univ. of Washington, Seattle, WA, USA

Inference of geophysical parameters from multifrequency polarimetric radar observations and model imversion 2108
J.J. van Zyl, C.F. Burnette, T.G. Farr, NASA/JPL, Pasadena, CA, USA

The effect of agricultural parameters on radar cross-section: A linear regression model 2112
G.C. Hussey, G.J. Sofko, Univ. of Saskatchewan, Saskatoon, Sask., Canada, B. Brisco, CCRS, Ottawa, Ont., Canada, J. Koehler, M.J. McKibben, Univ. of Saskatchewan, Saskatoon, Sask., Canada

C-Band backscatter sensitivity to multi-scale geometry and soil moisture variability of agricultural surfaces 2116
A. Beaudoin, Q.H.J. Gwyn, Univ. de Sherbrooke, Sherbrooke, Que., Canada, T. Le Toan, CESR, Toulouse, France

Page

Experimental investigation of the dependence of radar backscattering on wind speed, wind stress and wave height 2122
S. Gogineni, Univ. of Kansas Cen. for Res., Inc., Lawrence, KS, USA, K.B. Katsaros, Univ. of Washington, Seattle, WA, USA

The dependence of the microwave radar cross section on ocean surface slope statistics: comparison of measurements and theory 2125
D.E. Weissman, Hofstra Univ., Hempstead, NY, USA

Speckle statistics in airborne SAR imagery of the ocean surface 2126
T. Eltoft, Univ. of Tromso, Tromso, Norway, H. Johnsen, FORUT, Tromso, Norway

Remotely sensed surface temperature observations in HAPEX 2127
T.J. Schmugge, USDA Hydrology Lab., Beltsville, MD, USA, J.P. Goutorbe, Cen. Nat. de Rech. Met., Toulouse, France

Remotely sensed measurements of surface temperatures over sparse canopies 2130
C.S.T. Daughtry, USDA-ARS Rem. Sens. Res. Lab., Beltsville, MD, USA, W.P. Kustas, USDA-ARS Hydro. Lab., Beltsville, MD, USA, M.S. Moran, R.D. Jackson, USDA-ARS Water Cons. Lab., Phoenix, AZ, USA

On interconnection between microwave and infrared radiation fields and natural object condition 2131
E.A. Reutov, Inst. of Radio Engineering & Electronics, Moscow, USSR

AVHRR-derived cloudiness and surface temperature patterns over the Los Angeles area and their relationships to land use 2132
B. Dousset, Univ. of California, Los Angeles, CA, USA

Estimating evapotranspiration over large areas 2138
J.C. Price, Agric. Res. Serv., USDA, Beltsville, MD, USA

A one- and two- layer model for estimating evapotranspiration with remotely sensed surface temperature and ground-based meterological data over partial canopy cover 2142
W.P. Kustas, USDA-ARS/Nat. Resource Inst.,Beltsville, MD, USA, B.J. Choudhury, NASA/GSFC, Greenbelt, MD, USA, K.E. Kunkel, Illinois State Water Surv., Champagne, USA

Control of irrigation by following the water balance from NOAA-AVHRR thermal IR data 2146
A. Vidal, Lab. de Teledetect., CEMAGREF-ENGREF, Montpellier, France, A. Perrier, INRA Bioclimat., Thiverval-Grignon, France

Remotely sensed soil moisturé input to a hydrologic model 2150
E.T. Engman, W.P. Kustas, USDA Hydrology Lab., Beltsville, MD, USA, J.R. Wang, NASA/GSFC, Greenbelt, MD, USA

Solar radiation statistics for Australian capital regions 2154
M. Nunez, Univ. of Tasmania, Australia

The technological transfer on remote sensing programs, a tool for developing countries 2156
W. Danjoy A., ONERN, Lima, Peru

Remote sensing technology applications for the Caribbean: overcoming economic, political, & institutional obstacles 2157
C. Specter, D. Gayle, Forida Intl. Univ., Miami, FL, USA

The importance of remote sensing and mapping for resource development 2161
J. Ambrose, CIDA advis., Gov't of Nepal, Kathmandu, Nepal

Sustainable development in resource management through remote sensing 2165
J.E. Otoo, Univ. of British Columbia, Vancouver, BC, Canada

Page

Use of remote sensing satellite data for investigating water resources in Africa 2169
A. Pietroniro, Univ. of Waterloo, Waterloo, Ont., Canada, W. Wishart, S.I. Solomon, MacLaren Ing. Inc., London, Ont., Canada

Remote sensing for development - in Ethiopia 2173
A. Fanta, Ethiopian Mapping Auth., Addis Abeba, Ethiopia

Evaluation of satellite data for topographic map updating in Kenya 2177
H.O. Farah, P.J. Howarth, R.A. Bullock, Univ. of Waterloo, Waterloo, Ont., Canada

Evaluating land resource information with geographic information systems in Nepal 2180
P.B. Shah, Topogr. Surv. Branch, Gov't of Nepal, Kathmandu, Nepal

Microprocessor-based water resources requirements simulation analysis 2183
M. Ortiz, R. Vasquez-Espinosa, Univ. of Puerto Rico, Mayaguez, Puerto Rico

Development of an environmental information system for Venezuela 2184
C. Gosselin, S.R. Haja, R. Simard, DIGIM Inc., Montreal, Que., Canada, S.I. Solomon, Solomon & Associates, Canada, R. Salcedo, Min. Ambiente y Rec. Nat. Ren., Caracas, Venezuela

The development of a resource information management system: A pilot project in Northern Thailand 2188
P. Vincent, M. Meaudoin, C. Prevost, G. Lemieux, R. Simard, DIGIM Inc., Montreal, Que., Canada, S. Vibulsresth, National Research Council of Thailand, Thailand

A new airborne SAR for ice reconnaissance operations 2192
J.B. Mercer, INTERA Tech. Ltd., Calgary, Alta., Canada

International Ice Patrol side-looking airborne radar (SLAR) experiment 1988. 2193
S. Osmer, Michael Alfultis, Intl. Ice Patrol, USCG, Groton, CT, USA

Lightweight airborne radar for maritime surveillance 2197
D.C. Cantrill, P.A. Gordon, Litton Sys. Canada Ltd., Etobicoke, Ont., Canada

Ice management for the offshore fishery 2201
P. Rudkin, H. Ripley, K. Ludlow, Atlantic Airways Ltd., St. John's, Nfld., Canada

Digital processing of X-Band VARAN-S airborne SAR images; 2203
R. Albrizio, P. Blonda, A. Mazzone, Inst. Elab. Segnali ed Immagini, Bari, Italy, F. Pasqualik Cen. di Geodesia Spaziale, Matera, Italy, G. Pasquariello, Inst. Elab. Segnali ed Immagini, Bari, Italy, F. Posa, Univ. de Bari, Bari, Italy, N. Veneziani, Inst. Elab. Segnali ed Immagini, Bari, Italy

Helicopter radar simulations from fine resolution airborne SAR imagery 2209
R.T. Lowry, G. Wessels, G. McAvoy, INTERA Tech. Ltd., Calgary, Alta., Canada

C-Band SAR results obtained by an experimental airborne SAR sensor 2213
R. Horn, German Aerospace Research Establishment, Oberpfaffenhofen, FR Germany

A new method of aircraft motion error extraction from radar raw data for real time SAR motion compensation 2217
J.R. Moreira, DFVLR, Wessling, FR Germany

Page

The development of a polarimetric SAR for the detection of man-made objects 2221
R. Saper, Atlantis Sci. Sys. Group Inc., Nepean, Ont., Canada, M.T. Rey, DND, Ottawa, Ont., Canada, R. Wintjes, Atlantis Sci. Sys. Group Inc., Nepean Canada, C.E. Livingstone, CCRS, Ottawa Canada, M.R. Vant, DND, Ottawa, Ont., Canada, R. Austin, Canadian Astronautics Ltd., Ottawa Canada

A realtime quick look processor for airborne SAR 2225
A. Moreira, Inst. for Radio Freq. Tech., Wessling, FR Germany

An integrated ice surveillance system for support of offshore oil and gas exploration 2229
P. Rudkin, H. Ripley, S. Gillis, K. Ludlow, Atlantic Airways Ltd., St. John's, Nfld., Canada

Results of the real-time adaptive radiometric correction implemented in the DFVLR L/C-Band SAR 2232
J.R. Moreira, W. Poetzsch, DFVLR, Wessling, FR Germany

Applications of the SIMD MPP to processing and analyzing remotely-sensed imagery 2235
J.C. Tilton, NASA Goddard Space Center, Greenbelt, MD, USA

Applications of MIMD transputer arrays to automated image understanding of SPOT 2236
J.-P. Muller, Univ. Coll. London, London, UK

Applications of MIMD Hypercube to imaging spectrometry data analysis 2237
J.E. Solomon, NASA/JPL, Pasadena, CA, USA

Frequency domain filtering of remotely sensed satellite data using parallel processing 2238
B. Oldfield, J.E. Robinson, G.A. Betzos, Syracuse Univ., Syracuse, NY, USA

Digital enhancement of STAR-1 SAR imagery for linear feature extraction 2242
E.R. Sudibjo, G.D. Lodwick, S.H. Paine, Univ. of Calgary, Calgary, Alta., Canada

A technique for road detection from high resolution satellite images 2246
L. Lalitha, NNRMS-ISRO, Bangalore, India

Computer-aided linear planimetric feature extraction 2250
J.S. Zelek, MacDonald Dettwiler, Richmond, BC, Canada

Innovative techniques for the Spaceborne Imaging Radar-C instrument 2254
B.L. Huneycutt, NASA/JPL, Pasadena, CA, USA.

Advances in spaceborne active microwave sensing 2260
C. Elachi, NASA/JPL, Pasadena, CA, USA

Bistatic radar in space: A new dimension in imaging radar 2261
H.M. Braun, Dornier Sys. GMBH, Friedrichshafen, FR Germany, P. Hartl, Univ. of Stuttgart, FR Germany

Wide swath ocean topography mapping with interferometric altimeters 2265
E. Rodgriquez, F. Li, NASA/JPL, Pasadena, CA, USA, C. Parsons, E. Walsh, NASA/GSFC, Greenbelt, MD, USA

A phased array bread board for future remote sensing applications 2269
R. Zahn, E. Schmidt, Dornier Sys. GmbH, Friedrichshafen, FR Germany

Performance of a proposed synthetic aperture radar with variable antenna height 2273
K. Tomiyasu, General Electric Co., Philadelphia, PA, USA

Design of a dual polarized microstrip patch array for an advanced SAR antenna 2275
C.A. McDonach, D. Nguyen, G.S. Gupta, A. Luscombe, SPAR Aerospace Ltd., Ste. Anne de Bellevue, Que., Canada

Page

A generalized method for evaluation of the mutual coupling effect on the radiation and scanning characteristics of microstrip array antenna 2278
A.A.M. Ali, B.R. Vishvakarma, Banaras Hindu Univ., Varanasi, India

Design of a spaceborne SAR at higher incidence angle using alternate transmitted chirp slope reversal technique 2282
A.M. Jha, A. Majmudar, N.S. Pillai, Space Appl. Cen., Ahmedabad, India

RASS applied to wind profiler radars 2285
P.T. May, Univ. of Colorado, Boulder, CO, USA, R.G. Strauch, K.P. Moran, NOAA/ERL/WPL, Boulder, CO, USA

Application of Seasat scatterometer and passive microwave radiometer data to regional short-range weather forecasting 2289
S. Peteherych, Atmos. Env. Service, Toronto, Ont., Canada

Determination of near-surface wind speeds from sunglint data in band 2 of the AVHRR and comparison with meteorological data for the western Mediterranean 2292
A.P. Cracknell, S. Khattak, R.A. Vaughan, Univ. of Dundee, Dundee, UK

Application of Special Sensor Microwave/Imager data for analysis of cyclonic storms in midlatitudes over the sea 2296
K.B. Katsaros, G.W. Petty, I. Bhatti, D. Miller, Univ. of Washington, Seattle, WA, USA

Estimation of atmospheric thermal lapse rates using conventional acoustic radar measurements 2299
J.E. Letkeman, K.R. Peterman, Radian Corp., Austin, TX, USA

A comparison of the individual and combined performance of ground-based radio-acoustic and radiometric temperature sounding systems 2300
J.A. Schroeder, NOAA/ERL/WPL, Boulder, CO, USA

Measuring winds and rainfall from space--the Radar Wind Sounder (RAWS) 2303
R.K. Moore, W. Xin, Univ. of Kansas Cen. for Res., Inc., Lawrence, KS, USA

Boundary layer structure prevailing over Calcutta 2307
J. Das, A.K. De, A. Ganguli, Indian Statistical Inst., Calcutta, India

Dissection of multimodal directional ocean wave spectra form LEWEX 2311
T.W. Gerling, R.C. Beal, Johns Hopkins Univ., Laurel, MD, USA

Comparison of simulated and measured SAR image spectra during the Labrador extreme waves experiment 2315
C. Bruning, L.F. Zambresky, Eur. Cen. for Med. Range Weather Forecasts, Reading, UK

The effect of sampling variability on the estimation of wave number and propagation direction from SAR wave spectra 2316
F. Monaldo, John Hopkins Univ., Laurel, MD, USA

Resolution of ocean wave propagation direction in single-pass airborne SAR imagery 2320
P.W. Vachon, CCRS, Ottawa, Ont., Canada, R.K. Raney, Radarsat Project Off., Ottawa, Ont., Canada

On the focussing issue of Synthetic Aperture Radar imaging of ocean surface waves 2324
W. Alpers, Univ. Bremen, Bremen, FR Germany, C. Bruning, Eur. Cen. for Med. range weather forc., UK, J. Schroter, Alfred-Wegener-Inst., Bremerhaven, FR Germany

The use of ambiguity functions for SAR measurement of ocean waves 2328
A.S. Milman, A.O. Scheffler, Env. Res. Inst. of Mich., Ann Arbor, MI, USA

Page

SAR measurements of ocean waves and other surface phenomena during the GOASEX experiment 2331
D.J. Napolitano, J.F. Vesecky, Stanford Univ., Stanford, CA, USA, F. Gonzalez, NOAA, Seatlle, WA, USA, S. Petreherych, Atmos. Env. Serv. of Canada, Downsview, Canada

Pier-based measurements of ocean wave spectra using three-frequency Ku- and L-Band radars 2335
D.L. Schuler, W.C. Keller, Naval Res. Lab., Washington, DC, USA

Mapping bottom topography with X-band SLAR 2338
J. Vogelzang, Rijkswaterstaat Tidal Waters Diectorate, Emmeloord, Netherlands, G.J. Wensink, Delft Hydraulics, Netherlands, G.P. de Loor, Physics and Electronics Lab TNO, Netherlands, H.C. Peters, Rijkswaterstaat North Sea Directorate, Netherlands, H. Pouwels, National Aerospace Lab NLR, Netherlands

Removal of 180 ambiguity in SAR images of ocean waves 2342
W. Rosenthal, F. Ziemer, GKSS Res. Cen., FR Germany, K. Raney, P. Vachon, CCRS, Ottawa, Ont., Canada

Statistical descriptions of keel-related features in the under-ice draft distribution 2346
J. Key, A.S. McLaren, Univ. of Colorado, Boulder, CO, USA

Role of ice properties in wave-ice interaction during LIMEX '1987 2350
W.D. Winsor, J.I. Clark, CCRE, St. John's, Nfld., Canada, B.M. Eid, C.M. Morton, MacLaren Plansearch Ltd., Halifax, NS, Canada

Identification and volume estimation of icebergs by remote sensing in the Barents Sea 2355
S. Vefsnmo, S.M. Lovas, S. Loset, Norwegian Hydrotech. Lab., SINTEF, Trondheim, Norway, T. Ness, Fjellanger Widerøe A/S, Norway

Observations of sea ice drift off Newfoundland using satellite imagery and ice beacons 2359
I. Peterson, S.J. Prinsenberg, Bedford Inst. of Oceanog., Dartmouth, NS, Canada

Effects of the sea ice patterns during the August reversal of the Beaufort Sea gyre on atmosphere forcing 2363
E. LeDrew, D. Johnson, Univ. of Waterloo, Waterloo, Ont., Canada, J. Maslanik, Univ. of Colorado, Boulder, CO, USA

Some observations of the snowcover on sea ice in the Gulf of Bothnia 2364
G.B. Crocker, Memorial Univ. of Nfld., St. John's, Nfld., Canada, D. Werle, AERDE Env. Res., Halifax, NS, Canada

Multisensor remote sensing of thin ice 2365
I.G. Rubinstein, R.O. Ramseier, York Univ., North York, Ont., Canada

Estimation of the thickness of undeformed first year ice using radar backscatter 2366
E.D. Soulis, W.C. Lennox, J.F. Sykes, Univ. of Waterloo, Waterloo, Ont., Canada

A 3-dimensional modelling of forest canopies for high resolution imagery 2370
R. Fournier, J.R. Miller, York Univ., North York, Ont., Canada

Preprocessing of airborne multispectral scanner data for forest damage assessment in hilly terrain 2374
F. Frey, A. Bischof, P. Meyer, K.I. Itten, K. Staenz, Univ. of Zurich-Irchel, Zurich, Switzerland

Page

Classification of forest types by microwave scatterometer 2378
M. Hallikainen, T. Tares, J. Hyyppa, E. Somersalo, Helsinki Univ. of Tech., Espoo, Finland

Integration of laser rangefinder and multispectral video data for forest measurements 2382
D.H. Currie, V.L. Shaw, F.G. Bercha, The Bercha Group, Calgary, Alta., Canada

Application of aerial multispectral videography and colour/colour IR photography in sugar maple decline assessment 2385
X. Yuan, J. Vlcek, Univ. of Toronto, Toronto, Ont., Canada, D. King, Ryerson, Toronto, Ont., Canada, D. McLaughlin, Ont. Min. of Env., Toronto, Ont., Canada

MEIS FM: a multispectral imager for forestry and mapping 2390
R.A. Neville, S.M. Till, CCRS, Ottawa, Ont., Canada

Cooperation in remote sensing research: The Remote Sensing Working group of Forestry Canada 2394
Y.J. Lee, Pacific Forestry Centre, Victoria, B.C., Canada

Count of Brazilian Parana pine crowns in 35 mm aerial color photographs 2397
A.A. Disperati, J. Skalski Jr., Univ. Federal do Parana, Parana, Brazil

Comparison of satellite forest cover maps to traditional forest inventory maps 2401
J. Polson, D.R. Ens, Sask. Res. Council, Saskatoon, Sask., Canada, J. Whiting, Sask. Parks, Prince Albert, Sask., Canada

Characteristics of first order intensity entropy as a measure of image region texture 2405
G.E. Carlson, Univ. of Missouri-Rolla, Rolla, MO, USA

Applications of geostatistics to image analysis 2409
J.A. MacDonald, J.R. Carr, Univ. of Nevada-Reno, Reno, NV, USA

Defining subtle features that are not confined to reflection domains 2412
J.E. Robinson, B. Oldfield, Syracuse Univ., Syracuse, NY, USA

Performance of two texture-based classifiers of cloud fields using spatially averaged Landsat data 2416
S.K. Sengupta, R.M. Welch, M.S. Navar, S.D. School of Mines & Tech., Rapid City, SD, USA

Image segmentation by iterative parallel region growing and splitting 2420
J.C. Tilton, NASA/GSFC, Greenbelt, MD, USA

Review on image analysis with mathematical morphology in remote sensing 2424
G. Flouzat, Univ. Paul Sabatier, Toulouse, France

Mathematical morphological operators in remote sensing 2430
R.M. Haralick, Univ. of Washington, Seattle, WA, USA

Real-time morphological transformation for linear feature extraction on aerial images 2431
D. Ionescu, R. Kadamani, Univ. of Ottawa, Ottawa, Ont., Canada

Operational satellite-derived SST charts for Canadian waters 2432
B.W. Wannamaker, SEA SCAN, Caledon East, Ont., Canada

Integrating water quality surveillance with automated biomonitoring from remote water quality platforms 2433
E.L. Morgan, W.L. Pennington, M.A. Eggleton, Tennessee Tech. Univ. Cookeville, TN, USA

Page

Airphoto interpretation and salmon river restauration 2434
J.-M.M. Dubois, A. Gosselin, Univ. de Sherbrooke, Sherbrooke, Que., Canada

The information content of AVHRR, MSS, TM and SPOT data in the Skagerrak Sea 2439
K. Sorensen, Norwegian Inst. for Water Res., Oslo, Norway, T. Lindell, J. Nissel, Univ. of Uppsala, Uppsala, Sweden

Thermal remote sensing for walrus population assessment in the Canadian Arctic 2444
D.G. Barber, Univ. of Waterloo, Waterloo, Ont., Canada, P.R. Richard, Dept. of Fish. and Oc., Winnipeg, Man., Canada, K.P. Hochheim, EMSI, Winnipeg, Man., Canada

Determination of probable fishing areas for the albacore (Thunnus Alalunga) in Chile's central zone 2447
M.A.B. Barbieri, E.R. Yanez, Univ. Catolica de Valparaiso, Valparaiso, Chile, M.S. Farias, R.H. Aguilera, Univ. de Chile, Santiago, Chile

Development of algorithms for remote sensing of marine Trichodesmium 2451
G.A. Borstad, G.A. Borstad Assoc. Ltd., Sidney, BC, Canada, J.F.R. Gower, IOS, Sidney, BC, Canada, E.J. Carpenter, State Univ. of New York, Stony Brook, NY USA, J.G. Reuter, Portland State Univ., Portland, OR, USA

Meteorological satellite imagery for marine weather forecasting 2456
N. McLennan, L. Neil, Pacific Weather Centre, Vancouver, BC Canada

Teledetection des sites a potentiel aquicole dans la Baie des Chaleurs au Quebec: Resultats preliminaires 2459
M. Grenier, Univ. de Sherbrooke, Sherbrooke, Que., Canada, J.-M. Dubois, Univ. de Sherbrooke, Sherbrooke, Qc, Canada, A.-M. Lavoie, Univ. de Montreal, Montreal, Que., Canada, E. Lambert, Univ. de Montreal, Montreal, Que., Canada

A precise APT receiver for cloud and sea surface temperature imaging 2463
J.F.R. Gower, Inst. of Ocean Sci., Sidney, BC, Canada, T.B. Lowe, Galaxy Consult., Victoria, BC, Canada, B. Wannamaker, Sea Scan, Caledon East, Ont., Canada

Topographic mapping from GPS-supplemented airborne radar 2466
J.B. Mercer, P. Button, INTERA Tech. Ltd., Calgary, Alta., Canada, M. Millot, M. Karspeck, F. Leberl, VEXCEL Corp., Boulder, CO, USA

Coherent SAR simulation 2467
V.H. Kaupp, N.D. Blackford, W.P. Waite, H.C. MacDonald, Univ. of Arkansas, Fayetteville, AK, USA

Automated stereo matching using intelligent interpolation in difficult to match regions 2468
J.P. Strong, H.K. Ramapriyan, NASA/GSFC, Greenbelt, MD, USA, W.P. Haynes, ST Systems Corp., Lanham, MD, USA

Effects of speckle and additive noise on the altimetric resolution of interferometric SAR (ISAR) surveys 2469
C. Prati, F. Rocca, A.M. Guarnieri, Politecnico di Milano, Milano, Italy

Generation of geocoded spaceborne SAR image products 2473
E. Meier, Ch. Graf, D. Nuesch, Univ. of Zurich, Zurich, Switzerland

Closed form solutions to the single-image and stereo synthetic aperture radar mapping equations 2478
M. Kobrick, NASA/JPL, Pasadena, CA, USA

Page

Accuracy in positioning with stereoscopic SLAR imagery for topographic mapping 2479
I. Wu, D.C. Lin, Cen. Space. and Rem. Sens. Res. (CSRSR), Chung-Li, Taiwan, Rep. China

Effects of radar image distorsions in a radargrammetric application of VARAN-S data 2483
L. Polidori, Univ. Paris - 7, Paris, France, P. Kamoun, AEROSPATIALE, Cannes La Bocca, France, H. Raggam, Ins. of Image Proc. & Comp. Graph, Graz, Austria

SAR image-terrain database registration 2487
J. Curlander, W. Kober, F. Leberl, VEXCEL Corp., Boulder, CO, USA

Using MIMICS to model microwave backscatter from tree canopies 2491
K.C. McDonald, M.C. Dobson, F.T. Ulaby, Univ. of Michigan, Ann Arbor, MI, USA

Global altimeter measurements of extremes of wind speeds and waveheights and their impact on forecasting 2492
E.B. Dobson, John Hopkins Univ., Laurel, Maryland, USA

Radiative transfer theory for active remote sensing of a forested canopy 2493
M.A. Karam, A.K. Fung, Univ. of Texas at Arlington, Arlington, TX, USA

Radar modeling of tropical mangal forest stands 2497
Y. Wang, Univ. of California, Santa Barbara, CA, USA, M.L. Imhoff, NASA/GSFC, Greenbelt, MD, USA, D.S. Simonett, Univ. of California, Santa Barbara, CA, USA

X-Band backscatter modeling of orchard canopy 2501
N.S. Chauhan, George Washington Univ., Washington, DC, USA, R.H. Lang, NASA/GSFC, Greenbelt, MD, USA

Electromagnetic wave extinction within a forested canopy 2502
M.A. Karam, A.K. Fung, Univ. of Texas, Arlington, TX, USA

Spectral reflectance properties of lichens: remote sensing discrimination of lichen and rock 2506
E.A. Cloutis, Univ. of Alberta, Edmonton, Alta., Canada

Contribution of 0.57um/0.88um ratio technique in geobotanical studies 2507
A.P. Cracknell, A.K. Saraf, Univ. of Dundee, Dundee, UK

A background geobotany model applied to geologic exploration in tropical forest environments: First results from Landsat-TM and DTM for Pojuca Cu-Zn deposits, Carajas Province 2511
W.R. Paradella, Inst. de Pesquisas Espaciais, Sao Jose dos Campos, Brazil, B. Bruce, CCRS, Ottawa, Ont., Canada, J.K. Hornsby, C. Kushigbor, INTERA Tech. Ltd., Ottawa, Ont., Canada

Geobotanical and biogeochemical anomalies over a fluorite mineralization in Kerio Valley, Kenya 2514
B.D. Odhiambo, Univ. of Waterloo, Waterloo, Ont., Canada, J. Gaciri, Univ. of Nairobi, Nairobi, Kenya, A. Falconer, Region. Cen. for Ser. in Surv., Nairobi, Kenya

Monitoring vegetation regrowth on placer mine tailings, Bonanza Creek, Yukon Territory 2518
J.K. Hornsby, INTERA Tech. Ltd., Ottawa, Ont., Canada, B. Bruce, CCRS, Ottawa, Ont., Canada, G. Mackensie-Greive, Env. Protect. Serv., Whitehorse, Yuk., Canada

Page

Correlation study of remote sensing and geological data for mineral exploration, Farley Lake area, Manitoba 2522
V. Singh, W.M. Moon, Univ. of Manitoba, Winnipeg, Man., Canada, M. Fedikow, Dept. of Energy and Mines, Winnipeg, Man Canada, A. Rencz, Geol. Survey of Canada, Ottawa, Ont., Canada

SAR for geology in Alberta 2528
J.K. Hornsby, R.V. Dams, INTERA Tech. Ltd., Calgary, Alta., Canada, I.A. Sutherland, A.M. Hougham, Alberta For. Lands and Wildlife, Edmon., Canada

Using passive, thermal remote sensing techniques for detecting subsurface gravel accumulations in vegetated, unconsolidated sedimentary terrains 2532
G.S. Burns, NASA/J.C. Stennis SC, Bay St. Louis, MI, USA, D.E. Scholen, USDA Forest Service, Atlanta, GA, USA

Backscatter analysis of airborne radar: implications for background geobotany 2535
J. Harris, J.K. Hornsby, INTERA Tech. Ltd./CCRS, Ottawa, Ont., Canada

Data access and economic intelligence from satellite remote sensing for worldwide resource exploration, development and environmental management 2539
F.B. Henderson III, The Geosat Committee, Inc., Norman, OK, USA

Airborne and satellite verification of arms control agreements: Past, Present, and Future 2543
T.W.P. Brogden, Ryerson Polytech. Inst., Toronto, Ont., Canada, A.W. Dorn, Univ. of Toronto, Toronto, Ont., Canada, D. Scott, The Markland Group, Hamilton, Ont., Canada

Study of an advanced civil earth remote sensing system - results of an assessment of technology and markets through 2000 2547
D. Okerson, Sci. Applic. Intl. Corp., McLean, VA, USA

Planning an advanced commercial remote sensing system: evaluation of spaceborne, environmental sensor types which address the civil and military markets 2548
D.A. Brown, C.A. Hood, D. Okerson, M. Schwaller, R. Suresh, Sci. Appl. Intl. Corp., Washington, DC, USA, S.W. McCandless, User Sys. Eng. Inc., Annandale, VA, USA

GRS-S participation in International Space Year (ISY) 2550
C. Specter, Florida Intl. Univ., Miami, FL, USA

Technology transfer - A Canadian experience 2554
H. Epp, CCRS, Yellowknife, NWT, Canada, J. Whiting, Sask. Res. Council, Saskatoon, Sask., Canada

Is remote sensing an economic tool in Third World countries? 2558
B. Turner, N.M. Walters, CSIR, Pretoria, South Africa

Infrastructure - a necessity for operational use of remote sensing 2561
H. Hamnes, E. Henriksen, N.H. Lunde, S.A. Solbakk, FORUT (Foundation of Applied Research), Tromso Norway

The remote sensing industry in Canada: a profile of capabilities 2565
R.A. Ryerson, CCRS, Ottawa, Ont., Canada, J. Kerr, Dept. of External Affairs, Ottawa, Ont., Canada

A tool to assist forest management decision making 2566
W.B. Brown, BC Forest Serv., Smithers, BC, Canada, R.V. Quenet, Nawitka Resource Cons., Victoria, BC, Canada

Doppler centroid estimation ambiguity for Synthetic Aperture Radars 2567
C.Y. Chang, J.C. Curlander, NASA/JPL, Pasadena, CA, USA

Page

PRF ambiguity resolving for SAR 2572
H. Runge, R. Bamler, DLR, Oberpfaffenhofen, FR Germany

Theory and quantitative comparison of doppler centroid estimation methods 2576
J. Siewerth, DFVLR, Oberpfaffenhofen, FR Germany

Fast digital signal processing techniques for spaceborne Synthetic Aperture Radars 2579
R. Albrizio, Inst. Elabor. Segnali Immagini, Bari, Italy, G. Aloisio, Univ. Bari, Bari, Italy, C. Marzo, Agen. Spaziale Ital., Italy, A. Mazzone, Inst. Elab. Segnali Immagini, Bari, Italy, G. Milillo, Agen. Spaziale Ital., Italy, F. Pasquali, Cen. di Geodesia Spaziale, Matera, Italy, N. Veneziani, Inst. Elab. Segnali Immagini, Bari, Italy

Error sensitivities of a secondary range compression algorithm for processing squinted satellite SAR data 2584
F. Wong, I. Cumming, MacDonald Dettwiler, Richmond, BC, Canada

A phase preserving Synthetic Aperture Radar processor 2588
R.K. Raney, Radarsat Project Off., Ottawa, Ont., Canada, P.W. Vachon, CCRS, Ottawa, Ont., Canada

SAR super-resolution u ing a perturbed point spread function 2592
D. Blackneil, GEC - Marconi Res. Cen., Chelmsford, Essex, UK, S. Quegan, Univ. of Sheffield, Sheffield, UK

Fast SAR processing with the Fermat Number Transform 2596
J. Dall, Tech. Univ. of Denmark, Lyngby, Denmark

Effect of reference function weighting and selection of optimum weighting coefficient for Synthetic Aperture Radar 2600
Archana Majmuda, A.M. Jha, N.S. Pillai, Space App. Cen., Ahmedabad, India

Doppler parameters estimation criteria for the spaceborne SAR and corresponding attitude determination accuracy limits 2604
Tapan Misra, A.M. Jha, N.S. Pillai, Space App. Cen., Ahmedabad, India

Millimeter-wave sensor image enhancement 2608
W.J. Wilson, H. Suess, NASA/JPL, Pasadena, CA, USA

Application of real time digital image analysis and enhancement to airborne infrared fire detection and mapping 2610
W.D. McColl, L.E. Nelly, Innotech Avia. Enter. Ltd., Nepean, Ont., Canada, R.A. Neville, CCRS, Ottawa, Ont., Canada

Unsupervised coastline detection and tracing in SAR images 2611
J.S. Lee, I. Jurkevich, Naval Res. Lab., Washington, DC, USA

A linear transformation technique for spatial enhancement of multispectral images using a higher resolution data 2615
K.V. Shettigara, Univ. of New S. Wales, Sydney, Australia

TM/Landsat data processing for inland water monitoring 2619
M. Godoy, L.M. Novo, Inst. de Pesquisas Espaciais, S.Paulo, Brazil

Remote sensing the suspended sediment distribution on the Mingjiang river mouth in different tide phases as shown by MSS and TM imagery 2623
Delu Pan, S. Lin, Second Inst. of Ocean., Hangzhou, PR China

Geometric and radiometric correction of airborne MSS imageries over rugged terrain area 2627
N.Y. Chen, H.T. Wang, Nat. Cen. Univ., Chung-Li, Rep. China

Page

Evaluation du satellite SPOT pour l'extraction des elements planimetriques des cartes topographiques au 1:50 000: Resultats preliminaires 2628
L. Charbonneau, A. Coulombe, S. Perras, A. Royer, F. Bonn, Univ. de Sherbrooke, Sherbrooke, Que., Canada

Remote sensing detection of hydrocarbons and hydrocarbon seeps 2632
E.A. Cloutis, Univ. of Alberta, Edmonton, Alta., Canada

Parameters affecting reflectance of coniferous forests in the region of chlorophyll pigment absorption 2633
B. Curtiss, Univ. of Colorado, Boulder, CO, USA, S.L. Ustin, Univ. of California, Davis, CA, USA

Transmission, bidirectional reflectance and depolarization of an optical wave for a single plant leaf 2637
Qinglin Ma, A. Ishimaru, P. Phu, Univ. of Washington, Seattle, WA, USA, Y. Kuga, Univ. of Michigan, Ann Arbor, MI, USA

Detection of stress of coniferous forest trees with the VIRAF spectrometer 2641
C. Buschmann, U. Rinderle, H.K. Lichtenthaler, Univ. of Karlsruhe, Karlsruhe, FR Germany

Comparison of spectral data gathered from a labratory spectrometer and TM images with and without shadow correction 2645
R.L. Thiessen, J.R. Eliason, Washington State Univ., Pullman, WA, USA

Spectral emissivity observations in HAPEX 2649
T.J. Schmugge, USDA Hydrology Lab., Beltsville, MD, USA, F. Becker, Univ. Louis Pasteur, Strasbourg, France

Spectral reflectance characteristics of cotton & rice in Egypt 2652
H.M. El-Khattib, F.M. Hawela, Soil and Water Res. Inst., Giza, Egypt

Operational TOVS soundings using a physical retrieval approach 2653
A. Reale, NOAA, Washington, DC, USA, M. Goldberg, J. Daniels, ST Systems Corp., Lanham, MD, USA

Spectroscopic measurement of atmospheric water vapor and schemes for determination of evaporation from land and water surfaces using the AVIRIS 2658
J.E. Conel, R.O. Green, V. Carrere, J.S. Margolis, G. Vane, C. Bruegge, R. Alley, JPL/CIT, Pasadena, CA, USA

Column atmospheric water vapor retrievals from airborne imaging spectrometer data 2664
B.-C. Gao, A.F.H. Goetz, CIRES, Univ. of Colorado, Boulder, CO, USA

Remote sensing of the thermal forcing on the tropical Pacific 2669
W.T. Liu, NASA/JPL, Pasadena, CA, USA, C. Gautier, California Space Inst., La Jolla, CA, USA

Physical retrieval of precipitation cell parameters using passive 118-GHz observations 2672
A.J. Gasiewski, Georgia Inst. of Tech., Atlanta, GA, USA, D.H. Staelin, Massachusetts Inst. of Tech., Cambridge, USA

Detection and analysis of polar lows using satellite imagery and sounder data 2676
J. Turner, Nat. Env. Res. Council, Cambridge, UK

Statistical retrieval of precipitation cell-top altitude using passive 118-GHz observations 2680
A.J. Gasiewski, Georgia Inst. of Tech., Atlanta, GA, USA, D.H. Staelin, Massachusetts Inst. of Tech., Cambridge, USA

Page

Sensitivity of satellite signals to aerosol pollutants 2684
N.T. O'Neill, A. Royer, L. Hubert, Univ. de Sherbrooke, Sherbrooke, Que., Canada, J. Freemantle, York Univ., Downsview, Ont., Canada

Implications from LIMEX/LEWEX for spaceborne SAR sensing of ocean wave spectra 2688
R.K. Raney, Radarsat Proj. Off., Ottawa, Ont., Canada, P.W. Vachon, CCRS, Ottawa, Ont., Canada

Detectability of ocean waves by the Synthetic Aperture Radar on ERS-1 2692
A.O. Scheffler, C.L. Rufnach, R.A. Shuchman, Env. Res. Inst. of Mich., Ann Arbor, MI, USA

Strategies to assess the capacity of an on-board real-time SIR-C processor to produce ocean wave spectra 2693
F. Monaldo, Johns Hopkins Univ., Laurel, MD, USA

The influence of platform attitude dynamics on the doppler frequency synthesis of SIR-B ocean data near hurricane Josephine 2696
D.G. Tilley, J.L. Kerr, C.L. Vogt, Johns Hopkins Univ., Laurel, MD, USA

Radar surface signatures based on the two-dimensional tidal ciculation of Phelps Bank, Nantucket Shoals, MA 2700
S.R. Chubb, Sachs/Freeman Assoc., Inc., Landover, MD, USA, G.R. Valenzuela, Naval Res. Lab., Wahington, DC, USA, D.A. Greenberg, Bedford Inst. of Oceanog., Dartmouth, NS Canada

Comparisons between measured and theoretical results for C- and Ku-Band backscatter from naturally-occuring internal wave current patterns (FDC) 2704
T.W. Dawson, B.A. Hughes, Defence Res. Estab. Pacific, Victoria, BC, Canada

Surface current and wave modulation measurements over bottom topography - a comparison with theory 2710
P. Smith, D. Johnson, Stennis Sp. Cen., MS, USA

SAR ocean surface feature analysis in Conception Bay, Newfoundland 2714
K.A. Thomson, Seaconsult Ltd., St. John's, Nfld., Canada, S.A. Akenhead, Dept. of Fish. and Oc., St. John's, Nfld Canada, B.G. Snaderson, Memorial Univ. of Nfld., St. John's, Canada

Airborne SAR characteristics of arctic ice shelves and multiyear landfast sea ice, and the detection of massive ice calvings and ice islands 2715
M. Jeffries, W.M. Sackinger, Univ. of Alaska, Fairbanks, Alaska USA

Studies of ice sheet hydrology using SAR 2719
R. Bindschadler, NASA/GSFC, Greenbelt, MD, USA, P. Vornberger, ST Systems Corp., Lanham, MD, USA

A matched-filter technique for removing hyperbolic effects due to point scatterers: simulation and application on Antarctic radar data 2722
G. Raju, R.K. Moore, Univ. of Kansas Cen. for Res., Inc., Lawrence, KS, USA

Radar reflections from water injected into an Antarctic glacier 2726
R.H. Dean, C.H. Davis, W. Xin, Univ. of Kansas Cen. for Res., Inc., Lawrence, KS, USA

Bottom characteristics of the Antarctic ice sheet from radar 2727
R.K. Moore, W. Xin, G. Raju, Univ. of Kansas Cen. for Res., Inc., Lawrence, KS, USA

Slope correction by relocation for satellite altimetry 2730
A.P.R. Cooper, Scott Polar Res. Inst., Cambridge, UK

Page

Coherent radar contour mapping of ice stream thickness 2734
C.H. Davis, R.K. Moore, G. Raju, W. Xin, Univ. of Kansas Cen. for Res., Inc., Lawrence, KS, USA

N.B. * Denotes duplicate page numbers.

SATELLITE BASED DROUGHT SURVEILLANCE IN INDIA SOME ACHIEVEMENTS

S. Thiruvengadachari

National Remote Sensing Agency, Hyderabad, India 500 037

ABSTRACT

Though the resiliency of Indian agriculture helped to avoid the disastrous consequences of severe drought conditions in 1987 the event emphasized the need for continous and effective drought management measures. The first need is the development of a nationwide drought assessment and monitoring system to provide early warning and surveillance . The configuration of such a system, based on current remote sensing capabilities, is described. NOAA AVHRR data at LAC level for the years since 1986, through the major agricultural season (June-Dec.), and covering the whole country have been processed into two weekly NDVI imageries. Extensive ground data on rainfall, aridity anomaly, crop sowing, development and condition and various other anciliiary data were usd to compare with seasonal districtwise VI profiles and the drought assessment methodology has been developed. Operational drought monitoring in six states of India comprising of 113 districts will be taken up from June 1989, to be later extended to the entire country.

Key words : Agricultural Drought, NOAA AVHRR, Vegetation index, Indian National System, Surveillance

1. INTRODUCTION

Drought continues to be a major factor of uncertainty in India agriculture inspite of the significant technical advances since independence (Krishnan 1979). Severe drought conditions were experienced during 1987. Though Indian agriculture proved resilient enough to avaoid the disastrous consequences the event emphasised the need for continued efforts at effective drought managements. Thus the immediate requirement is the devellopment of a national/regional system for assessment and monitoring of agricultural drought conditions through the cropping season, to provide periodic information on the prevalence, severity level and persistence of drought conditions. It is expected that such a system would help resource managers to optimally allocate scare financial resources to where and when they are most needed. This will also help in reconciling the differing drought perceptions of central and state Governments. It is in this context that a project for development of a National Agricultural Drought Assessment and Monitoring System (N-ADAMS) has been under progress since 1986 at the National Remote Sensing Agency (Dept. of Space, Govenment of India) with the collaboration of India Meteorological Department, Central Water Commission and the various State Governments. The project has been sponsored by the Department of Agriculture and Cooperation, Government of India.

The N-ADAMS envisages primary dependance on NOAA AVHRR data, supplemented and supported by ground observations of rainfall and significant sowings of major crops. Since current drought assessment by State Governments involves the use of a large number of drought indicators, many of which are not readily (and currently) amenable to satellite sensing, the project will focus on agricultural drought conditions (Jayaseelan and Thiruvengadachari, 1986). Though satellite data can provide informaton on various drought parameters many are still in experimental stage and not proven for large area applications. The N-ADAMS has been consciously built around current operational capabilities (Thiruvengadachari, 1988).

2. CURRENT AGRICULTURAL DROUGHT MONITORING METHODOLOGIES

Currently agricultural drought conditons are characterised by ground observations of rainfall, aridity anomaly and agricultural conditions in regard to sown area and crop condition. Sparse realtime rainfall information, generally measured at tahsil headquarters, as well as nonunique rainfall-use-efficiency over space and time (Le Hourou 1984) makes rainfall a poor indicator of drought conditions. The aridity anomaly charts of the entire country, brought out by India Meteorological Department, are based on 194 observatories, and show the departures of actual aridity from normal aridity, giviing an indication of the severity of water deficit-to-water demand relationship on a weekly basis, representative of large areas such as meteorological subdivisions. The aridity anomaly suffers from the same limitations as rainfall. Though the ground observatins of agricultural drought conditions by the State Departments of Agriculture and Revenue are exhaustive such a system involves a significant amount of subjective judgement at various states (Jayaseelan and Thiruvengadachari; 1986). The periodicity and extent of ground observations vary significantly between different States, and frequently the submission of such reports from the districts to State headquarters is delayed or missed. The nature of sparse ground observations also make it difficult to assess in near realtime average drought conditions over a district.

In place of the discrete point ground observations,

N-ADAMS will provide direct spatial information on vegetation stresses caused by drought conditions.

3. NATIONAL AGRICULTURAL DROUGHT ASSESSMENT AND MONITORING SYSTEM (N-ADAMS)

N-ADAMS has the objective of providing realtime information on drought prevalence, relative severity level and persistence through the kharif (rainy) season all over the country with district as the reference unit. The N-ADAMS will use Local Area Coverage (LAC) data from NOAA AVHRR received at the Indian ground station operated by NRSA. A two weekly period has been chosen as optimum for time compositing of NOAA NDVI (Normalised Difference Vegetation Index) data, based on the experience of 1986, 1987 and 1988 cloud cover condtions. A longer period would be unacceptable in view of the short duration of many crops grown during the kharif season. The first standard meteorological week of 7 days starts on the first of January in any year.

The two weekly composited NDVI imagery of India has a pixel size of 1 km squeare and is geometrically corrected to a positional accuracyof 5 to 6 kms. with reference to digitised India map on 1:2.5 million scale. The time compositing algorithm uses maximum value compositing over land with cloud masking based on visible band and TIR bands of AVHRR is under validation.

NOAA AVHRR LAC level data for June to December in 1986, 1987 and 1988 has been processed into two weekly NDVI data sets, consisting of colour coded imagery showing the spatial vegetation condition as well as districtwise VI histogram, for the different states of the country. Districtwise seasonal VI profiles of various VI parameters such as the GNDVI (VI values averaged over all the pixels of the district), MNDVI (averaged over the green pixels, based on a threshold VI of 0.05) and VA (vegetation area based on the VI threshold), have been generated for the past three years. The use of these VI averages have also been reported in other drought monitoring systems elsewhere in the world (Philipson 1988).

The development of N-ADAMS methodology was based on the analysis of the VI profiles of past three years, supported by extensive dynamic ground data on rainfall, aridity anomaly, sowing progress of major crops and crop development and condition as well as static data on land utilisation details, general cropping pattern and crop calendar, irrigation support and past drought history. Most of the dynamic ground data was collected in realtime through the cooperative efforts of concerned State Agricultural Directorates. Standard forms and reporting procedures were followed to ensure objectivity, reliability and uniformity.

The experience of these three years' data indicates that operational use of N-ADAMS will continue to require ground information on rainfall as well as significant sowings of the principal crops of the district (Fig.1). This information however, is already being reported either completely or partially by ground agencies and may only require conversion to a realtime format. The configuration of N-ADAMS has been dictated by realistic appraisal of ground data availability and capability.

4. ANALYSIS OF PAST VI DATA

The analysis of 1986 to 1988 VI data sets was preceded by basic studies on agricultural drought definition and characterisation as well as current ground monitoring mechanisms (Harikishan and Thiruvengadachari 1987, Prasad and Thiruvengadachari 1987; Thiruvengadachari and Harikishan 1987, and Thiruvengadachari and Prasad 1987). Pilot studies using Landsat and NOAA data for selected districts of Andhra Pradesh and Tamil Nadu States helped in the first-cut specification of drought monitoring methodology (kishan, 1987, Chandrasekhar, 1988).

The comparison of seasonal VI profiles and ground measured rainfall, computed aridity anomaly and observed crop development and condition was attempted in the six States of Tamil Nadu,Andhra Pradesh, Karnataka, Maharashtra, Gujarat and Orissa, comprising of 113 districts.

Only broad lag correlations could be observed between VI profile and the rainfall distribution. The correlation was better in the case of severe and persistent rainfall deficiency rather than slight-to-moderate deficits.

Paradoxically the more-than-normal rainfall during 1988 Kharif season resulted in significantly lower VI values in many districts. While the exact reason for this inversion is not clear, it is suspected that the heavy rainfall might have need to saturated soils and water logging and hence the vegetation with water background yielding lower VI values. Thus conjunctive use of rainfall and NDVI is called for.

The aridity anomaly is representative of a meteorological sub-division and hence provides poor correlation with district VI profile; except during severe and widespread drought conditions.

The correlation betweeen moisture stress observed in a specified seasonal profile (as deviations from the normal year profile) and stressed crop conditions observed on ground was broadly valid, but not too much so due to sparse and subjective ground observations.

Trend analysis helped in the detection of anomalies in the growith cycle, such as delayed start, delayed/lower peak VI value and shorterning/extension of vegetative phase, as well as moisture stress observed at different growth stages. Since the district-wise VI profile represent average vegetation growth/condition, which is a mix of natural vegetation and various agricultural crops, extraction of crop specific inforamtion requires that the crop cycle be fixed on the average profile. This calls for ground reporting of significant sowing of the principal crops since they can not be spectrally identified.

The moisture stress periods identified from VI profiles have shown an excellent correlation in regard to the impact on crop yeild. It is interesting to note that other factors ofyield reductionsuch as pest/disease attack or unseasonal rainfall during flowering/grain development are not reflected in the average VI profile and hence are not confused with moisture stress.

Drought conditions result in reduction in cropped area and/or reduction in crop yield. While the latter could be qualitatively evaluated as a function of when and how severe moisture stress condition occur, the former is not reflected will in the VI profiles. The comparison of interannual seasonal VI profiles not always reflected the trends in cropped area. It is generally concluded that N-ADAMS may provide reliable indications of reduction in a vegetated area, and cropped area, only

under drought conditions as experienced in Gujarat State during 1987.

5. OPERATIONAL OUTLOOK

The State level ADAMS (S-ADAMS) will need to provide more detailed information than N-ADAMS and hence has been configured (Fig.2) to use high resolution data from earth resources satellites such as Landsat, SPOT and IRS (Indian Remote Sensing Satellite).

The changeover from NOAA-9 to NOAA-11 from Nov. 1988 is expected to introduce some problems in inter-annual comparison of seasonal VI profiles for drought assessment. The earlier acquisition time and different calibration coefficients of NOAA-11, will make the 1989 VI values generally higher than those of NOAA-9 during 1986 to 1988. An analysis indicated that the different calibration coefficients alone will increase the VI value by as much as 0.02 while the total impact including that of smaller solar zenith angle could be an increase of 0.05 on the district average VI value. The lack of concurrent NOAA-9 and NOAA-11 AVHRR data will make inter sensor calibration difficult and unreliable. Efforts are under way to solve this probelm prior to the beginning of cropping season in June 1989,

ACKNOWLEDGEMENTS

The author acknowledges the permission granted by Director, NRSA for publication of this paper. Grateful thanks are due to Prof. U.R. Rao, Secretary of Department of Space, Government of India, for the inspiring guidance received during the project. Thanks are also due to Secretary, Department of Agriculture and Cooperation, Governemnt of India for sponsoring this project. The help and cooperation received from the different collaborating agencies is also gratefully acknowledged. Last but not the least, the author acknowledges the cooperation received from the team members ofthe drought project. The typing support was provided by Mr. Babulal.

REFERENCES

1. Chadrakeshar,N. 1988, A Comparative Study of Landsat (MSS) and NOAA (AVHRR) Data for Agricultural Drought Indicators, M.Tech., Thesis, Anna University, Madras, India.

2. Jeyaseelan, A.T., & Thiruvengadachari, S., 1986, Current Drought Monitoring System in Andhra Pradesh State, Report No. IRS-UP-NRSA-DRM-TR-03-86, NRSA, Hyderabad, India.

3. Jeyaseelan, A.T., & Thiruvengadachari,S., 1986, Recent Drought History of Andhra Pradesh State, Report No. IRS-UP-NRSA-DRM-TR-04-86, NRSA, Hyderabad,India

4. Kishan, K., 1987, A Landsat Comparitive Study of Drought Severity for 1985 and 1986 in Anantapur District, Andhra Pradesh, M. Tech., Thesis, Regional Engineering College, Warangal, A.P., India.

5. Krishnan, A., 1979, Key paper on Definition of Droughts and Factors relevant to specifications of Agricultural and Hydrological Aspects of Droughts, New Delhi, India, 1, 67-102.

6. Le Hourou, H.N., 1984. Rain Use Efficiency - A Unifying Concept on Arid Land Ecology, J. Arid Environ. 7.3.

7. Philipson, W.R., 1988, Operational Interpretation of AVHRR vegetation indices for World Crop Information, Photogrammetric Engineering & Rem. Sen. Vol.54, No.1, pp 53-59.

8. Prasad, T.S., and Thiruvengadachari, S., 1987. Drought Prediction. A Review" Report No. RSAM-NRSA-DRM-TR-03/87, NRSA, Hyderabad, India

9. Thiruvengadachari, S., and Harikishan J., 1987, Satellite Monitoring of Agricultural Drought in Hindupur Division in Anantapur District of A.P. State, Report No. RSAM-NRSA-DRM-TR-08/87,NRSA, Hyderabad, India.

10. Thiruvengadachari,S., and Harikishan, J., 1987, Drought Early Warning Based on Monthly Rainfall Anomalies, Report No. RSAM-NRSA-DRM-TR-11/87, NRSA, Hyderabad, India.

11. Thiruvengadachari, S., and Prasad T.S., 1987, Aridity Anomaly as Drought Indicator. An Evaluation, Report NO. RSAM-NRSA-DRM-TR-09/87, NRSA, Hyderabad, India.

12. Thiruvengadachari, S., 1988 Remote Sensing Capabilities and Contribution to Better Drought Monitoring, Workshop on Agricultural Droughts, NRSA,Hyderabad, India.

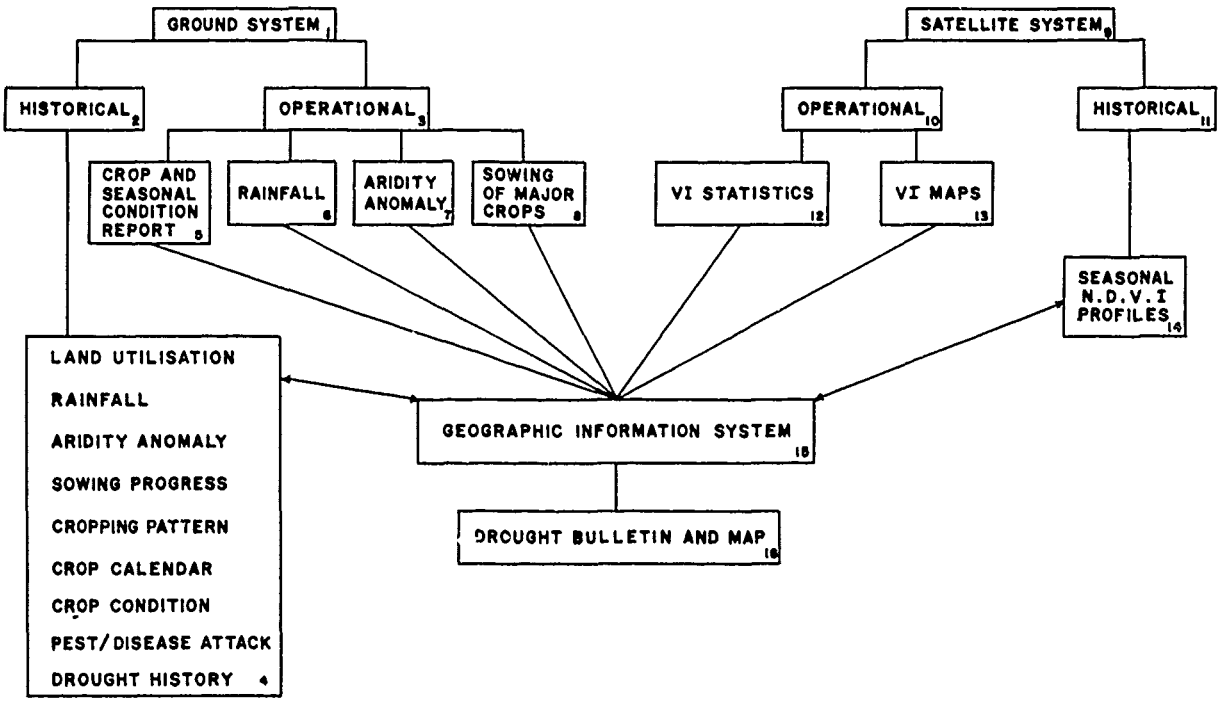


Fig. 1: NATIONAL AGRICULTURAL DROUGHT ASSESSMENT AND MONITORING SYSTEM (N-ADAMS)

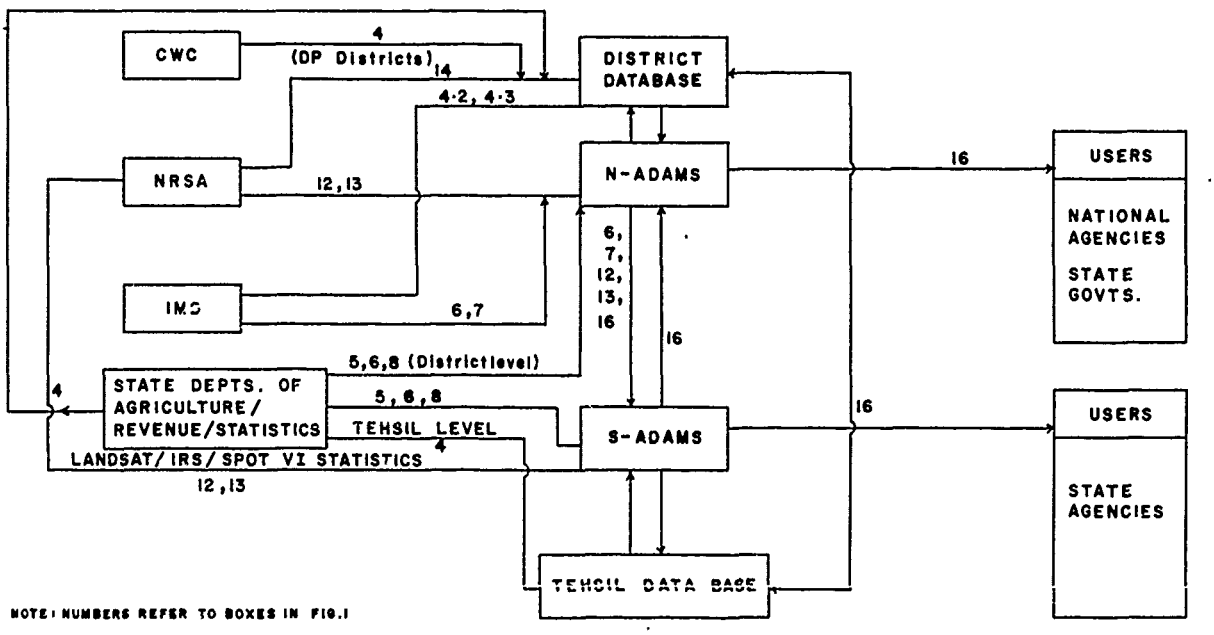


Fig. 2: INTEGRATED AGRICULTURAL DROUGHT ASSESSMENT AND MONITORING SYSTEM

MONITORING DROUGHT AFFECTED VEGETATION WITH AVHRR

E. W. Allison
Prairie Farm Rehabilitation Administration
Regina, Saskatchewan

R. J. Brown
Canada Centre for Remote Sensing
Ottawa, Ontario

H. E. Press, J. G. Gairns
Intera Technologies Ltd.
Ottawa, Ontario

ABSTRACT:

Data from the Advanced Very High Resolution Radiometer (AVHRR) on the NOAA series of meteorological satellites can be used to generate a cost effective near real-time vegetation-based drought indicator. During the summer of 1988, the Canada Centre for Remote Sensing and the Prairie Farm Rehabilitation Administration used these data to monitor vegetation conditions and help determine drought program assistance payments. This paper discusses factors affecting this satellite technique and outlines procedures used to monitor drought affected vegetation in over 400 different administrative units in Western Canada.

Keywords: AVHRR, Drought, Crop Condition

INTRODUCTION:

In 1988, drought severely impacted farming operations in southern Alberta, Saskatchewan and Manitoba by reducing on-farm water supplies, forage supplies and crop yields. The drought caused by the cumulative effects of below normal fall soil moisture reserves, abnormally low winter and spring precipitation, and abnormally hot dry growing conditions in May and June, severely reduced forage, grain and oilseed crops, in varying degrees over large areas of the prairies.

The drought stimulated requests for area specific assistance programs. As these programs have been administered on a Rural Municipality (County of Municipal District in Alberta) basis, and there are over 400 such administrative units, this created a need to monitor drought impacts in over 400 areas of varying size. As part of its drought monitoring activities, and in support of the 1988 Federal-Provincial Drought Committee, the Prairie Farm Rehabilitation Administration (PFRA), a branch of Agriculture Canada, requested the Canada Centre for Remote Sensing (CCRS) to assist in evaluating the use of NOAA satellite data for drought monitoring.

Prior to 1988 drought monitoring during the growing season has been primarily based on field observations, limited point source meteorological data, and the results of soil moisture and crop yield models derived from these data. As summer rainfall is highly variable and there are only 48 daily and 146 weekly, routinely reporting meteorological stations in the prairie provinces, the availability of accurate and timely precipitation data and its spatial limitations have been a significant limiting factor to drought monitoring. In 1988, PFRA and CCRS supplemented the existing drought monitoring procedures by using remote sensing to measure chlorophyll reflectance, a measure of vegetation vigor, on a Rural Municipality (RM) basis. The objectives of this work were to develop a near real-time drought indicator with improved spatial coverage and to enhance the interpretation of point source meteorological data, limited field observations, and other traditional drought indicators.

This paper will describe how data from the Advanced Very High Resolution Radiometer (AVHRR) on the NOAA series of meteorological satellites were used to monitor the 1988 drought. Emphasis will be placed on the operational advantages and disadvantages of the data and studies conducted to relate these data to other drought indicators. In particular, the issue of comparison of data from year to year and from region to region will be addressed.

DATA SOURCE AND PREPROCESSING:

The AVHRR data used in this study were received from the Manitoba Remote Sensing Centre (MRSC). These data are a weekly composite of daily AVHRR images generated in such a manner that each pixel in the composite is the clearest value for the composite period. This is accomplished, at the MRSC, by registering each daily image to a base map and forming a normalized difference image from this mapped data. This normalized difference image is the difference of channels 2 and 1 divided by the sum to form a normalized difference vegetation index (NDVI). The weekly composite is formed from these daily images by

selecting, for each pixel, the radiance values from the date having the largest NDVI value (Holben, 1986). Normally the NDVI value is reduced in value if the area is covered by clouds, partially covered by clouds, or there is radiometric degradation of the imagery due to haze or large scan angles. The resultant weekly composite is thus virtually cloud free and has been corrected, to some extent, for radiometric degradation due to atmospheric effects.

The MRSC receive the AVHRR data from the Canada Centre for Remote Sensing Receiving station at Prince Albert, Saskatchewan by courier and has the processed data in the hands of the user agencies within four days of the last date which goes into the composite image.

Since it was required to have vegetation condition information on a RM basis, a coverage of these boundaries was read into an ESRI ARC\INFO Geographic Information System (GIS). Then software, recently developed by Dipix Technologies Inc., was used to transfer this coverage to the Dipix ARIES image analysis system. Within the image analysis system, the average NDVI values were calculated for each RM and stored as attributes of the RM polygons. These attributes were then transferred back to the GIS for cartographic display on a polygon (RM) basis. In this case, the change of NDVI between 1988, specific years, and an average of 1985 to 1987, were displayed by RM in graphical format. In addition, the image analysis system was used to produce photographic products such as 35 mm slides, split screen comparisons and density sliced images.

Because the AVHRR sensor. with an approximate one kilometre instantaneous field of view at nadir, measures the reflected radiation from all the ground cover and not individual fields this average NDVI value will, to a first approximation, be dependent upon the general landcover and agricultural practices (Brown and Prevost, 1985). However, in addition, there is a significant variation in NDVI caused by changes in vegetation vigour. Hence it is important that the average NDVI values for a particular area, such as a RM, be compared to values for the same area, for other years. If within year comparisons are made of the NDVI from different locations, care must be taken to insure that the ground characteristics of the areas under comparison are the same. This includes crop mix, soil type, percentage of native pasture, etc. For this reason we restricted our analysis to a comparison of average NDVI values from the same RM for the various years.

It is also important to look at the multitemporal curves of NDVI throughout a year in assessing the validity of a particular data point. It is possible that for a given area, there may still be some residue cloud present which has not been removed by the compositing process. This is quite apparent from a examination of the multitemporal NDVI data. Knowledge of precipitation data, landuse, crop mixes, cultivation practices and field conditions also assisted in the interpretation of these data.

RESULTS AND DISCUSSION:

Since the assistance programs were administered on an RM or county basis (of which there are over 400 across the Western Canadian provinces) and there are only 146 weekly reporting meteorological stations throughout the area it is apparent that no precipitation data was available for many RMs. The remotely sensed data can be used to interpolate between the point measurements. Another consideration, which is often not addressed, is the validity, or representativeness, of a point precipitation measurement in a year of significantly reduced rainfall. During the summer, much of the prairie precipitation tends to result from local convective storms rather than from large frontal systems. When there are a large number of such smaller storms, as in a normal year, individual meteorological station measurements will be more representative of the precipitation throughout an area than in drought years when fewer convective and frontal rain events occur. Hence, the remotely sensed data has an increased role to play in dry years such as 1988. In addition, it has been shown that the remotely sensed data is far more sensitive to changes in the amount of photosynthetically active material when the amount of material is small. This, of course, is the situation in a drought year. Furthermore, regression models relating soil moisture or expected yield to precipitation do not follow the peaks and valleys in the data distribution well because the combined effects of extreme temperatures, low precipitation and high evapotranspiration, which occur during droughts, are difficult to model accurately. As a result, it has been observed that the models have a tendency to overestimate crop yields in extremely dry years. In contrast, the NDVI data are a more direct measurement of the combined impacts of drought on crops, their vegetative vigor, and hence crop yields.

Figure 1 shows the multitemporal NDVI values for an RM of Southern Manitoba. This area was severely affected by the drought in 1988. Normally, the grain and canola plants are over 60 cm high at the time of heading for grains or flowering in the case of canola. However, in 1988 the plants were significantly stunted, attaining a height of only 20 to 30 cm at the same development stage. In addition, the plant density was very low. When viewed from above, substantial areas of soil could be seen between the plants compared to a normal year when there is virtually 100% ground cover. For this area the precipitation in 1988 was much below normal over the months of May to July which was compounded by a dry winter. In 1988, the average yield for the Crop Reporting District which included this RM was 14.7 bu\ac compared to a value of 33.0 in 1987. The peak of the NDVI curve for 1988 was reduced by 36% compared to 1987 and the average of 1985-1987. This qualitative approach truly illustrated the information content of the AVHRR data for vegetation monitoring and work is now progressing on making the analysis more quantitative.

CONCLUSIONS:

The AVHRR data has been demonstrated to be useful for monitoring drought conditions in the prairie provinces. These data have the advantage that they represent a direct measurement of the vegetation conditions, which can be altered by drought, and that they are continuous in spatial extent. Consequently, they are an excellent complement to point meteorological measurements. Another important consideration for an operational agency is data timeliness. It has been demonstrated that the remotely sensed data can be collected and preprocessed in a timely fashion such that it can be used in real-time crop monitoring. It is anticipated that, in the

future, these two data sources will be closely linked in new crop yield and drought detection area specific models, which will greatly enhance drought monitoring. One of the disadvantages of the remotely sensed data is the lack of a long term database which is relatable to historical crop yield data and the need for further research to make the data analysis more quantitative. Vigorous efforts are currently being made by various groups to correlate these data with crop yields and refine the monitoring techniques. In summary, even as a qualitative tool, the AVHRR has been very useful in monitoring and identifying areas impacted by drought. NOAA satellite data can be used to generate a cost effective near real-time vegetation-based drought indicator. Further studies to refine this monitoring technique are warranted because of it has great potential to improve our crop condition and drought monitoring capabilities.

ACKNOWLEDGEMENTS:

The authors would like to acknowledge inputs from Mrs. G. Ferguson who put the manuscript together in its final form and the personnel at the Manitoba Remote Sensing Centre who did an excellent job in processing and delivering the AVHRR data to us.

REFERENCES:

Brown, R.J., and C. Prevost, "Integration of High and Low Resolution Satellite Data for Crop Condition Assessment", Proc. of Eleventh Int. Sym. on Machine Processing of Rem. Sens. Data, Purdue University, Indiana, pp 189-196, 1985.

Holben, B.N., "Characteristics of Maximum-Value Composite Images from Temporal AVHRR Data", Int. J. Rem. Sens., Vol. 7, pp 1417-1434, 1986.

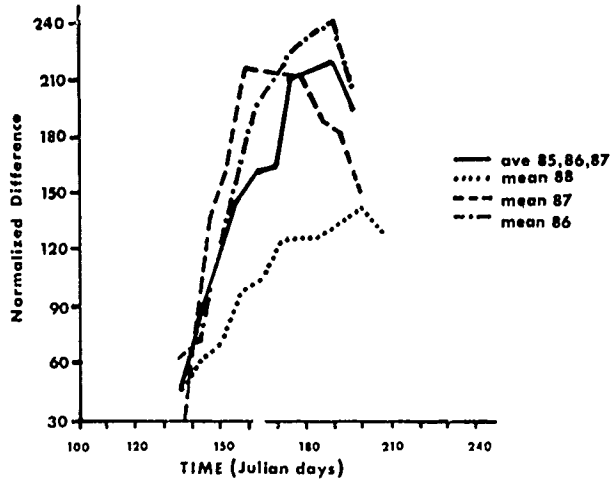


Figure 1. A typical vegetation index vs time for a Rural Municipality in Southern Manitoba.

MONITORING RANGELAND CHANGE USING LANDSAT MSS DATA

D.W. Mullins

Intera Technologies Ltd., Ottawa, Ontario

J. Cihlar

Canada Centre for Remote Sensing, Ottawa, Ontario

ABSTRACT

Since 1980, the Canada Centre for Remote Sensing has undertaken work to determine the extent to which LANDSAT Multispectral Scanner (MSS) data can be used to assist managers of the rangelands of Western Canada. Previous studies have focussed on the qualitative classification of range condition based on the tones of the enhanced satellite imagery. The aim of the present study is to evaluate the feasibility of monitoring the change in rangeland condition using quantitative digital analysis techniques.

12 cloud-free fall images covering a 20,000 km^2 area of southeastern Alberta were acquired between 1973 and 1988. A procedure previously developed at CCRS to convert digital counts to reflectance values was applied to the individual scenes. This procedure corrects for sun angle and atmospheric effects, thus making possible between-date comparisons. The images were registered to the Universal Transverse Mercator (UTM) projection and resampled to produce 50 m X 50 m pixels.

MSS band 5 difference images were calculated for selected periods spanning the data coverage. The images showed year-to-year changes in range condition and land use. Land use changes, particularly the conversion of rangeland into cropland or seeded pasture, were the dominant changes in the area. Significant differences in the level of range use between years were also observed in most difference images. A standard deviation image product was then prepared in which the pixel value represents the standard deviation of band 5 for all the years. This product highlighted specific areas where land cover condition varied substantially over the 15 year period.

Keywords: Rangeland, LANDSAT-MSS, Alberta

1. INTRODUCTION

The Public Lands Division of Alberta Forestry, Lands and Wildlife is responsible for administering approximately 1.25 million hectares of publicly owned mixed grass prairie. The grazing lands are leased to individual ranchers and grazing associations or are maintained as Provincial Grazing Reserves. One of the responsibilities of Public Lands is to monitor the condition of the rangeland in order to develop and implement appropriate management practices. This monitoring effort is hampered by the size of the area that must be covered by range inspectors. Another concern of range managers is the diminishing size of the prairie as pressures from oil and gas exploration, conversion to conventional farming, and recreational interests increase.

Starting in 1980, the Canada Centre for Remote Sensing has been working with the Public Lands Division to develop methods for using remotely sensed data in long-term monitoring of range condition and use. As part of the investigations, the relationship between biomass and reflectance values was explored (Brown et al., 1983). This and subsequent studies (e.g., Thomson et al., 1985; Adams et al., 1986; Adams et al., 1988) identified the difficulties associated with a quantitative interpretation of range condition from satellite data. Besides inconsistencies in the reflectance vs. biomass relationship and its variations with season and vegetation composition, substantial numbers of ground samples are required to provide an estimate of biomass production (Ahern et al., 1981). In addition, biomass is not the only (or necessarily the best) measure of range condition in mixed grass prairie. The studies established, however, that rangeland condition can be qualitatively assessed from enhanced LANDSAT images (Pearce, 1980; Brown et al.,

1983; Adams et al., 1988).

Using visual analysis, the interpreter can effectively examine only a few images at a time. This creates serious difficulties when examining long-term trends in vegetation condition over a large area (Pearce, 1985). Another complicating factor is the misregistration of images. As a result, it is not possible to follow the trend of one or a few pixels from year to year.

In the present study, the emphasis is placed on simplifying and standardizing the interpretation process. LANDSAT MSS band 5, which spans the red spectral region from 0.6 to 0.7 micrometres, was used as an indicator of range condition. This choice was based on the fact that, for vegetation studies, LANDSAT MSS has two uncorrelated bands (5 and 7) but since there is very little green biomass in the fall the bands are highly correlated and band 5 provides most of the useful information. Pearce (1980), Brown et al. (1983) and others found that low band 5 values correspond to a better range condition; specific reasons could be higher standing brown biomass, higher carryover, higher club moss (*Selaginella densa*), or a conversion to seeded grass cover such as crested wheatgrass (*Agropyron cristatum*). In contrast, range in poor condition due to excessive grazing or environmental factors appears as a lighter tone on the imagery. In addition to exposing the underlying soil, heavy grazing tends to eliminate the more desirable native grasses which are then replaced by more weedy species. Since there is very little green biomass in the fall, band 5 is highly correlated to band 7 and thus provides most of the useful information.

2. METHODOLOGY

The study area is a 20,000 km^2 region of mixed grass prairie in southeast Alberta. This region is characterized by native grasslands with a cover of short grasses such as needle-and-thread (*Stipa comata*) and blue grama (*Bouteloua gracilis*), medium grasses such as northern wheatgrass (*Agropyron dasystachyum*) and crested wheatgrass (*Agropyron cristatum*), club moss (*Selaginella densa*) and small forbs and shrubs. These grasses are associated with Brown and Dark Brown Chernozem soils as well as some Solonetzic soils. The soils are generally shallow, have a high sodium content, and are susceptible to wind erosion when dry. Erosion of the upper soil layer exposes the light-coloured highly reflective B horizon.

A time series of 12 fall images was selected because little or no green vegetation is present at this time of year, leaving only brown carryover biomass. In order to facilitate multi-date comparisons, the effects due to variations in solar elevation angle, atmospheric transmission, and path radiance had to be eliminated. A procedure had been developed which measures path radiance and infers atmospheric transmittance for LANDSAT MSS bands using clear water bodies. The procedure transforms the raw MSS digital counts to reflectance values which are corrected for the aforementioned factors. Clear water bodies were identified in each input scene and used to perform the conversion to reflectances. In order to enhance the subtle tonal differences in the imagery, a linear contrast stretch developed specifically for rangeland was applied to each image (Brown et al., 1981). This ensures that the minimum and maximum reflectance values correspond to digital counts of 0 and 255 respectively. The stretch is given by:

$$D = 255\ (R - R_{min})/(R_{max} - R_{min}) \qquad 1)$$

where D is the resultant digital value of a pixel, R is the input reflectance, and R_{min} and R_{max} are fixed reflectance limits. The images were then registered to the Universal Transverse Mercator (UTM) projection and resampled using a nearest neighbour algorithm to produce a pixel size of 50 m X 50 m.

Between-year difference images were calculated as:

$$(\text{Band } 5)_{t1} - (\text{Band } 5)_{t2} + 128 \qquad 2)$$

where the subscripts t1 and t2 correspond to a more recent and an older image respectively. The standard deviation image was calculated using:

$$[((x_1 - a)^2 + \ldots + (x_n - a)^2)/(n - 1)]^{1/2} \qquad 3)$$

where x is the individual pixel value from each input image (Table 1) and a is the mean value for that pixel from all 12 images.

Year	Date	Year	Date
1973	September 27	1981	September 12
1974	September 22	1982	September 14
1975	September 8	1983	no image
1976	September 20	1984	October 13
1977	no image	1985	September 30
1978	no image	1986	September 1
1979	September 5	1987	September 20
1980	no image	1988	September 22

Table 1: Dates of LANDSAT MSS images

3. RESULTS AND DISCUSSION

Due to space constraints, only a subset of the data from one area is presented. Three images spanning the study period were examined to ascertain rangeland changes that are evident in MSS band 5 data. The sub-scene chosen for discussion is located just southeast of the village of Hays, with the Bow River crossing the area from top to bottom (Figure 1).

The overall tone of the 1973 scene imaged on September 27 (Figure 1a) was lighter than in the other two years, probably reflecting low carryover in this very dry year for the region. Note also that an island appears in the centre of the reservoir at A indicating a lower water level than in the other two years.

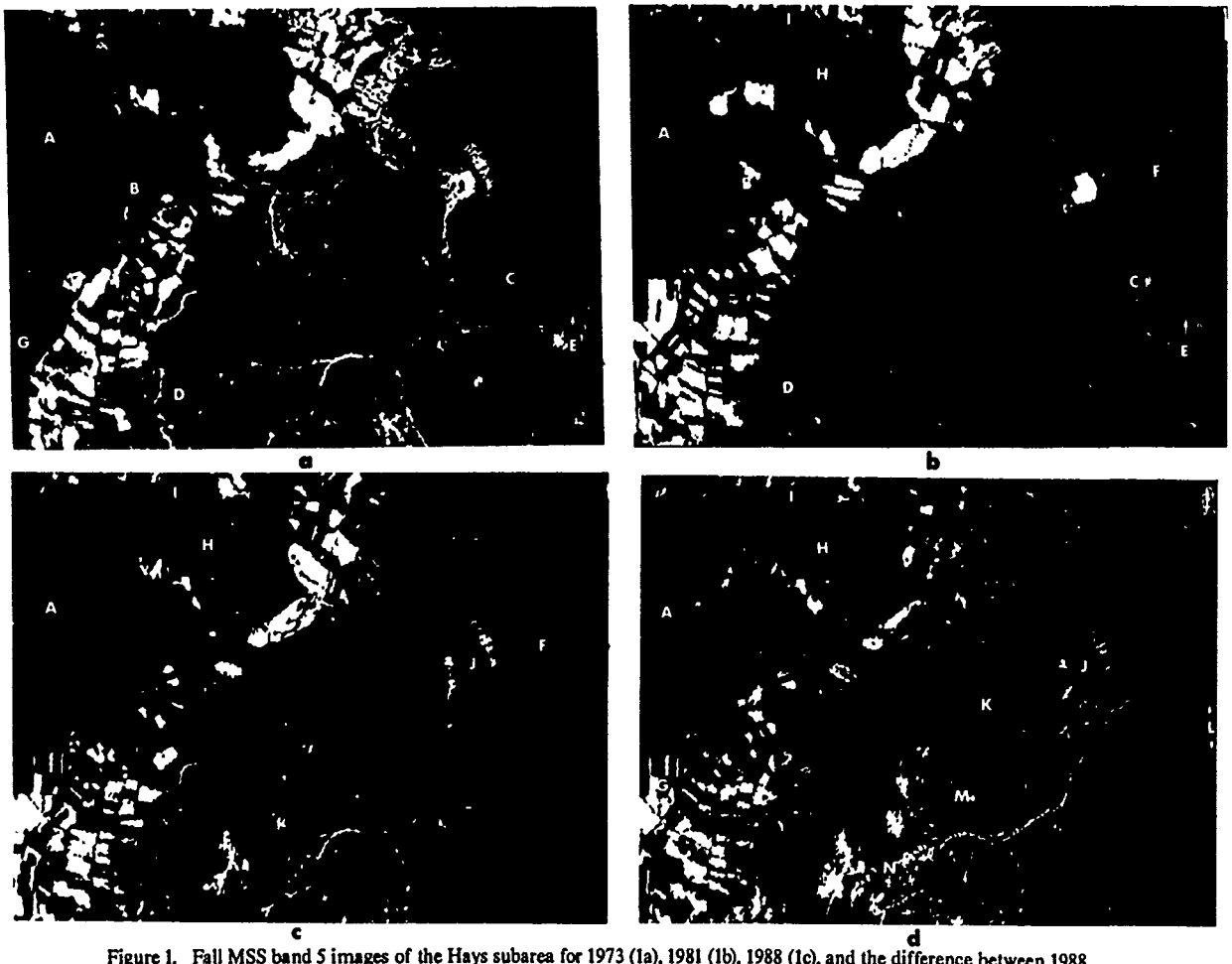


Figure 1. Fall MSS band 5 images of the Hays subarea for 1973 (1a), 1981 (1b), 1988 (1c), and the difference between 1988 and 1973 (1d). See text for discussion.

On September 12, 1981 (Figure 1b) some notable changes in land use and range condition were evident. The overall range condition appeared better than in 1973, although some specific sites were in worse condition, for example B, C, and D. On the other hand, area E did not experience a decline in range condition by 1981 although the surrounding range did. Range E was only grazed after the grasses had senesced, while the surrounding range was grazed all summer (Pearce, 1985). The dark feature at F (Figure 1b) is a fire scar from late summer of that year. Fortunately, the fire was quite light and by September 22, 1988 there was no evidence of its occurrence in the imagery (Figure 1c). Analysis of the time series showed that much of the range from area G (Figure 1a) was broken and replaced by irrigated farming between 1976 and 1979.

Also of note is the amount of prairie converted from native grass production to irrigated lands. Installations of pivot irrigation systems supporting crested wheatgrass have replaced part of the native range in area H by 1981, and almost all of the area by 1988. In the 1981 (Figure 1b) and 1988 (Figure 1c) images it appears from the light tones that the pasture between these pivot systems has been heavily grazed. Between 1981 and 1988, area I just across the road from H was broken and replaced with wheel irrigation systems.

The change in land use and range condition between 1973 and 1988 is evident in the difference image produced using equation 2 (Figure 1d). A neutral tone of grey, such as that exhibited by the reservoir at A, represents little or no change while white and black tones correspond to areas of large change. Some of the previously mentioned locations stand out very well (G, H, and I) as does area J which was native range in 1973 but was converted to irrigated alfalfa by 1988. In terms of range condition, the darker tones of regions K and L indicate range that was in better condition in 1988; the lighter tones at M and N identify range that has diminished in quality.

The standard deviation image (Figure 2) serves to highlight regions of highly variable land cover. It should be noted that this image is

shifted slightly to the north of the other images. The white tones show areas of high variability while the dark tones indicate constant conditions. Most of the white regions are farmland (including native range converted to farmland) although some spurious effects caused by the presence of clouds in one of the images are seen in the lower portion of Figure 2. The fluctuating water level of the reservoir at A is seen as a white strip around the water body. The intermediate tones of grey represent rangeland condition variations which are relatively small (compared to the changes in land use). The previously discussed areas F and L are highlighted in this fashion, as are the irrigated lands at H.

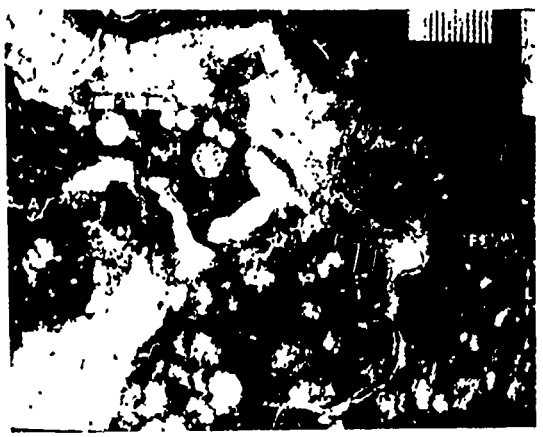


Figure 2: Image in which each pixel value represents the standard deviation of that pixel for the entire dataset.

4. SUMMARY

Year-to-year difference images based on LANDSAT MSS band 5 fall data serve to identify areas of changing range condition in the mixed grass prairie. This technique can assist range managers in highlighting areas where corrective management practices may be required. It can also be used to determine area estimates of rangeland and other land cover types in order to map the changing land use. The standard deviation image highlights particularly variable regions, although cloud cover in any one image limits the usefulness in the affected areas.

Since MSS band 5 reflectance in mixed grass areas is monotonically related to range condition, the rectified and calibrated LANDSAT MSS band 5 image could, in principle, be transformed to show a "range condition" index. For this procedure to be sound, the MSS data must be absolutely calibrated and a time-invariant MSS vs. range condition relationship established. Encouraging results have earlier been reported in both areas for mixed grass prairie vegetation (Brown et al., 1983). First attempts at using this approach for a long time series showed that the absolute calibration process is rather sensitive to operator input, thus introducing a subjective element. Further work in this direction continues. Future work will also involve the incorporation of ancillary data through the use of geographic information systems.

ACKNOWLEDGEMENTS

The authors would like to acknowledge the contribution of Mr. Barry Adams and Mr. Bob Schuler of the Public Lands Division in Lethbridge for providing background field information, and the assistance of Mr. Trevor Taylor and Ms. Anne Botman during the data processing stage.

REFERENCES

Adams, B., Bernier, M., McLeod, J.C. and O. Dupont, "Effectiveness of the Thematic Mapper for the range condition assessment in fescue grasslands of Southwestern Alberta", Proceedings of the 10th Canadian Symposium on Remote Sensing, Edmonton, Alberta, 1986.

Adams, B., Bernier, M., deValois, J., Dupont, O. and I. Sutherland, "Evaluation of LANDSAT TM enhanced images for range management applications in Alberta", Canadian Journal of Remote Sensing, Vol. 14, No. 1, pp4-16, 1988.

Ahern, F.J., Brown, R.J., Thomson, K.B.P., Pearce, C.M., Fedosejevs, G. and K. Staenz, "Optical characteristics of Alberta rangeland as related to remote sensing", Proceedings of the 7th Canadian Symposium on Remote Sensing, Winnipeg, Manitoba, 1981.

Brown, R.J., Thomson, K.P.B., Ahern, F.J., Staenz, K., J. Cihlar, Klumph, S.G. and C.M. Pearce, "LANDSAT MSS applied to rangeland management in Western Canada", 15th Symposium on Remote Sensing of the Environment, Ann Arbor, Michigan, 1981.

Brown, R.J., Ahern, F.J., Thomson, K.P.B., Staenz, K, Cihlar, J., Pearce, C.M. and S.G. Klumph, "Alberta rangeland assessment using remotely sensed data", Research Report 83-1, Canada Centre for Remote Sensing, Energy, Mines and Resources Canada, 128 pp., 1983.

Pearce, C.M., " An evaluation of satellite-generated information for inclusion in rangeland monitoring programs in Alberta", M.Sc. thesis, Dept. of Geography, University of Calgary, 206 pp, 1980.

Pearce, C.M., "An evaluation of enhanced archival LANDSAT imagery for monitoring mixed grass prairie rangelands in Alberta", Research Report for Alberta Remote Sensing Centre, Contract #85-0191, 46 pp, 1985.

Thomson, K.P.B. Gosselin, C., Adams, B.W. and I. Sutherland, "Thematic Mapper data used for rangeland management in rough fescue grasslands of Western Canada", Canadian Journal of Remote Sensing, Vol. 2, pp162-176, 1985.

OPERATIONAL MONITORING OF FORAGE AND CROP YIELD

Jeff Whiting and Nick Nicolichuk
Saskatchewan Research Council
Saskatoon, Saskatchewan, Canada S7N 2X8
and
Stan Benjamin
Canada/Saskatchewan Crop Insurance Corporation
Melville, Saskatchewan, Canada S0A 2P0

ABSTRACT

This paper details the first attempt to provide operational crop and forage yield in Canada. Yield estimates were provided within ten days after the end of every week, starting June 15 and ending August 7, 1988. The major breakthrough was the production of cloud-free imagery by the Canada Centre for Remote Sensing and Manitoba Remote Sensing Centre. This breakthrough enabled the Saskatchewan Research Council to produce the crop and forage estimates for the Canada/Saskatchewan Crop Insurance Corporation.

The methodology and results of operational monitoring of the 1988 drought are discussed in this paper. Calibration was performed using 1985, 1986 and 1987 data. NOAA AVHRR were used to estimate yield for the major grain and oil crops and forage crops. Two hundred and ninety-nine Rural Municipalities (average size 7000 km^2) were used as test sites in southern Saskatchewan. Results indicated that drought boundaries obtained from NOAA data compare favorably with those obtained from yield data provided many months later.

INTRODUCTION

One of the major economic factors in Saskatchewan is the production of cereal and oil grains. During the 1984, 1985 and 1988 crop years the production of grains and forage crops was severely reduced by a drought. A major portion of agricultural Saskatchewan was affected. Drought aid to farmers therefore became one of the most important expenditures for the provincial government. Timely data on areas affected and the severity of the drought are important to the governments in their financial compensation to the farmers.

Drought has played an important role in the success or failure of agriculture in Saskatchewan since the early days of settlement. The "Dirty Thirties" were probably the most memorable drought years with 1961, 1977, 1980 and 1984-1985 being the most recent ones (Fung 1969, Faultey et al 1986). The abnormally prolonged dry weather during the most recent drought (1984-1985 and 1988) have serious effects on crop growth. The most vulnerable sectors of the economy to drought are, of course, the farmers. Decision makers require timely information, on the drought areas, in order to help farmers over this period of crises through the government financial assistance programs. However, for these programs to be fair and equitably applied to those actually affected by the drought, accurate delineation of the drought boundary must be available in a timely manner.

Multi-spectral remotely sensed data from satellites has proven useful in assessing crop vigour. Most authors have used some form of vegetation index or normalized difference vegetation index (NDVI) using Landsat multi-spectral scanner or NOAA Advanced Very High Resolution Radiometer (AVHRR) data (Tucker 1979, Henricksen and Durkin 1986, Graetz and Gentle 1984, Henricksen 1986, Thompson and Wehmanen 1978). In most cases, where drought assessment or monitoring was required, large land areas were involved and high temporal resolution preferred. This increases the likelihood of obtaining cloud free imagery over large areas more frequently. AVHRR data provides two passes per day over the critical areas, low spatial resolution for large area coverage and two channels appropriate for vegetation analysis.

The grazing and hay lands of Saskatchewan extend from the wide open short grass prairie of the south to the northern roughland bush pastures of the Aspen Parkland (Populus tremuloides). These areas total 7,853,000 ha.

The agricultural areas in Saskatchewan are divided into relatively large federal and risk crop districts which are administrative units established by the federal and provincial governments to gather different types of census data. Within each are smaller Rural Municipalities (RM). Insurance programs designed to protect farmers against crop loss due to drought, hail and other natural disasters are available through government sponsored programs. Crop yield data is provided by individual farmers, however, crop statistics used in insurance programs are collected and tabulated by Saskatchewan Crop Insurance Corporation (SCIC).

OBJECTIVE

SRC was requested by Saskatchewan Crop Insurance Corporation to provide:

1. Snow-cover as of March 31, 1988.
2. Soil moisture conditions - April 15, 1988.
3. NOAA AVHRR data at 1 km resolution in the form of:
   a. Computer tape and film prints of crop stress (weekly from the Manitoba Remote Sensing Centre [on Wednesdays]) for 16 weeks at 40cm x 50cm and 20cm x 25cm.
   b. Film prints of potential yield for forage crop on two dates (week ending June 20 [shipping date June 29] and week ending July 4 [shipping date August 2]). Verification of the satellite data using historic data for each RM. Derivation of the statistical data for 1988.
   c. Film prints of Crop Yield Estimates for two dates: week ending June 20 and week ending July 24. Verification of the satellite data using historic RM supplied by Saskatchewan Crop Insurance Corporation [SCIC] in digital form for 3 years and provide mean and standard deviation for long-term data. Derivation of statistical data for 1988.

METHODOLOGY

Computer compatible tapes of NOAA AVHRR data were obtained for June, July and August of 1983 to 1988. The data was normalized by applying a quadratic correction curve to the data using the minimum radiance (clear water at nadir) and the brightest radiance values such as clouds. Normalized data from the visual and infrared channels were ratioed to form a Normalized Difference Vegetation Indices (i.e. 350 x [Channel 2-1]/[Channel 2+1]) for specific regions of the province. An agricultural district boundary map was overlain on the NOAA data.

Digital image analysis was performed on a DIPIX computer system. The daily Normalized Difference Vegetation Index (NDVI) was summarized by a digital image (a pixel representing 1 sq km) into a weekly composite in order to eliminate cloud cover. The maximum daily NDVI for each pixel was used in order to have the most recent data.

The Normalized Difference Vegetation Index for each week was then contrast stretched using Dipix Aries II computer software. The contrast stretch for crop yield was performed on the bases of a ratio between the weekly NDVI and the maximum NDVI for 1987.

For forage the equation calculated first an estimate of monthly rainfall and from that the amount of forage in kg ha $^{-1}$ (Smoliak et al., 1976). The weekly 1988 data was divided by the maximum 1987 data at each point. Historical yield differences for specific areas between the normal and for each forage were used to derive ratios.

The yield values each year from 1965 to 1987 were compared to 1987. The agricultural district boundaries were digitized and used to extract an average NDVI value for each in a geographic region.

RESULTS

Cibachrome colour prints of enhanced AVHRR imagery were produced from the computer enhancements to visually define the boundaries of the areas affected. Colour prints of enhanced imagery were provided with three classes, namely, dry, transitional and equal to or greater than 1987 for the chosen dates in 1988.

The 1985 imagery showed an extensive zone of drought extending from the southwest towards the northeast. During 1986 and 1987, these same drought areas supported mainly healthy crops and other vegetation. For forage the spread of the three years to the 15 year mean was larger with a greater spread at the higher end.

Analysis of the NOAA data compared to the ground data was done using ratios of 1985 to 1987 and 1986 to 1987.

For 1988, the NOAA analysis for forge yield estimate showed that Saskatchewan averaged 57% less than normal and 45% for crop yield (Table 1).

Table 1. Area of Province at or Below Last Year for Forage and Cereal Crop Using Satellite Data

	Forage: June 24-30, 88/87			Cereal Crop: July 14-21, 88/87		
% of 1987	Area (%)	Cum (%)	Area (km)	Area (%)	Cum (%)	Area km
0-39%	--	--	--	.1	.1	451.2
40-49%	18.9	18.9	74728.2	1.9	2.0	7719.6
50-59%	6.5	25.5	25906.1	7.8	9.9	30975.8
60-69%	7.5	33.0	29616.1	11.4	21.4	45349.6
70-79%	8.1	41.1	32259.7	10.1	31.5	39869.6
80-89%	8.4	49.6	33237.8	7.6	39.1	30038.7
90-99%	7.4	57.0	29526.8	6.1	45.3	24424.2
>100%	42.9	100.0	169331.2	54.6	100.0	215776.8
	100.0		394605.0	100.0		394605.0

DISCUSSION

Delineation of drought boundaries for 1988 using AVHRR data were obtained by both visual and digital analysis. Classification of the province into the three units showed the southwest to have been dry in 1983, 1984, 1985 and 1988 due to lack of sufficient soil moisture as indicated by low acreage of grain crops. Normally, much of the area will be dominated by native pasture and hayland which remained green only for a relatively short time in spring and early summer. By mid July, much of the pasture grasses were matured and beginning to dry. During 1985, this dry zone was much more extensive to the northeast and affected the forage crop growing area. In 1988, the forage areas did not green-up until the June 30 rains came and the temperature dropped. In general, the majority of the province is below 90% of the 1987 yield.

CONCLUSION

Cloud cover has always been a major deterrent in using satellite data. The process of producing a weekly composite of the data proved to be a successful solution to this problem. Crop and forage estimates using rainfall data were compared to field and ground survey information. The results were within a reasonable range of estimation for the 3 year test 1985-87 and were used for the 1988 data. The cibachrome photographs produced from the enhanced NOAA Satellite data were evaluated by crop, and forage and range management personnel as to their potential use as a management tool for evaluating crop, pasture and forage conditions and production in Western Canada.

Acknowledgements

Canada and Manitoba Centers for Remote Sensing, Saskatchewan Crop Insurance, Saskatchewan Agricultural Development Fund New Pastures and Grazing Technology Projects.

REFERENCES

1. Brown, R.J., M. Bernier, G. Fedosejevs, Geometrical and Atmospheric Considerations of NOAA AVHRR, "Proc. of the 8th Int. Symp. on Mach. Proc. of Remotely Sensed Data", Purdue Univ., West Lafayette, Indiana, pp374-381, 1982.
2. Brown, R.J., Personnel Communication, 1988.
3. Epp, H., and J.Polson, "Monitoring the Recent Drought in Southern Sask. with AVHRR and MSS Data", SRC Report E-905-24-D-88, Saskatoon, Sask. Presented 8th Canadian Remote Sensing Symp. Waterloo, 1987.
4. Faultey, R.A., M.A. Bonneau, S.N. Kulshreshtha, "An Economic evaluation of impacts of the 1984-85 Prairie drought", Canadian Hydrology Symp. - Drought - The Impending Crisis, Regina, Saskatchewan, pp367-383, 1986.
5. Fung, K.I. (editor), "Atlas of Saskatchewan", Dept. of Geography, University of Saskatchewan, Saskatoon, Saskatchewan, Canada, 1969.
6. Graetz, R.D. and M.R. Gentle, "NOAA and the Drought of '82", Proc. 3rd Australian Remote Sensing Conference, Queensland, Australia, pp560-563, 1984.
7. Henricksen, B.L. and J.W. Durkin, "Growing period and drought early warning in Africa using satellite data", Int. J. Remote Sensing, Vol.7, No.11, pp1583-1608, 1986.
8. Henicksen, B.L. "Reflections on drought: Ethiopia 1983-1984", Int. J. Remote Sensing Vol.7, No.11, pp1447-1451, 1986.
9. Hougham, A., "Use of NOAA AVHRR Digital Satellite Data for Precipitation and Forest Fire Assessment", Saskatchewan Research Council Technical Report No.198, Saskatoon, 1987.
10. Prout, N., N. Dalezios, M. Manore, J. Tracey, R. Brown, and G. Walker, "Agriculture drought: precipitation deficiency and NOAA AVHRR indices comparisons", Canadian Hydrology Symp. - Drought - The Impending Crisis, Regina, Saskatchewan Canada, pp91-100, 1986.
11. Rao, P.P.N. and V.R. Rao, "An approach for agricultural drought monitoring using NOAA/AVHRR and Landsat imagery", Proc. of IGARSS '84 Symp., Strasbourg, France, pp225-229, 1984.
12. Smoliak, S., A.Johnson, M. Kicher, and R. Lodge, "Management of Prairie Rangeland", Canada Agriculture Publication #1589, Ottawa, 1976.
13. Thompson, D.R. and O.A. Wehmanen, "Application of Landsat digital data for monitoring drought", Proc. of Tech. Sessions, The LACIE Symp. Houston, Texas, pp431-438, 1978.

REGIONAL PADDY MONITORING SYSTEM USING NOAA HRPT DATA

Yoshiaki MATSUMAE, Tsukasa HOSOMURA,
Haruhisa SHIMODA, and Toshibumi SAKATA

Tokai University Research & Information Center
Tomigaya 2-28-4, Shibuya-ku Tokyo 151, Japan

ABSTRACT

NOAA HRPT data are most suitable for regional scale monitorings, e.g., 1000 - 2000 km scale for their wide coverage and reasonable ground resolution (- 1 km) as well as frequent data acquisition capability. For global change detections, NOAA APT data have been already used for many kinds of applications. However, utilizations of HRPT data introduce several kinds of additional problems. In this paper, methods to use HRPT data for regional paddy monitorings are described.

The following three problems should be solved in order to establish operational monitoring system using HRPT data.

1) Higher geometric accuracy is required to co-register HRPT data for clouds elimination.
2) Faster processing algorithms for geometric corrections are required to process large quantity of data within a limited duration.
3) radiometric distortions mainly caused by incident sun light should be eliminated.

Following processings have been adopted to solve these problems. Faster and accurate geometric corrections were performed by three step processings. In order to eliminate radiometric distortions, sun angle corrections were first applied. Residual radiometric distortions were eliminated by the aid of histogram normalization technique. After all the above preprocessings, pixels having the lowest value of NVI (Normalized Vegetation Index) were selected to compose cloud free image from 1 week HRPT data. Those cloud free images have been proved to be sufficient to monitor rice crop growth conditions of basin of Yang Zi river in china.

1. INTRODUCTION

This paper describes about the regional paddy monitoring system using remote sensing technology. Target area of this study is large paddy fields spreading over Hua Zhong and Hua Nan area, i.e., basin of Yang Zi river. This area has a regional scale and it is very difficult to get ground informations in a timely manner. NOAA AVHRR data is optimal for such monitorings. In this study, a system in order to obtain the necessary informations of paddies in this area within certain time interval by using NOAA AVHRR images are examined.

This system is composed of the following steps. NOAA AVHRR images are collected in Tokai University Space Information Center (TSIC) in Kumamoto, Japan. These data are transported to Tokai University Research & Information Center (TRIC) in Tokyo and processed there. Data processing is composed of three steps. First step is preprocessings which include geometric correction and radiometric correction. Geometric correction was conducted to precisely co-register each images. radiometric correction was performed to eliminate the radiometric distortions mainly caused by incidence angle difference of sun light. In the second step, a cloud free image of the target area was generated by combining images within one or two weeks. In the third step, the informations of paddy growing stages were extracted using a combination of band 1 and band 2 of cloud free AVHRR image with the aid of MOS-1 data as ground truth data. In this step, two kinds of processings were examined. One is the false color composite of band 1 and band 2 while the other is the use of vegetation index. In both cases, paddy conditions of regionalscale have been effectively monitored within reasonable time duration.

2. TEST IMAGE DATA

Test image data used in this study are as follows;

DATA SET 1 :	NOAA-9 and NOAA-10 1988.MAY.15 - MAY.24	5 IMAGES
DATA SET 2 :	NOAA-9 and NOAA-10 1988.JUNE.3 - JUNE.6	5 IMAGES
DATA SET 3 :	NOAA-9 and NOAA-10 1988.JUNE.24 - JUNE.30	8 IMAGES

A NOAA images of the target area is shown in Fig.1 which was taken on June 4, 1988. Target area of this study is Hua Zhong and Hua Nan area, i.e., basin of Yang Zi river.

Fig.1 THE OBJECT AREA OF THIS STUDY

3. PREPROCESSINGS

A flowchart of the image processings in this study is shown in Fig.2. Geometric correction and radiometric correction are very important in these processings.

3.1 GEOMETRIC CORRECTION

In this section, geometric correction for AVHRR data are described. In the geometric correction for this study, there are following two problems.

1) Higher geometric accuracy is required to co-register AVHRR data for clouds elimination.
2) Faster processing algorithms for geometric corrections are required to process large quantity of data within a limited duration.

Following processings have been adopted to solve these problems. Faster and accurate geometric corrections were performed by three step processings. At the first step, tangential and earth curvature distortions were corrected using table look up algorithm.

Tangent correction and earth curvature correction was done as follows. The symbols used in this preprocessings are shown in Fig.3. Let P is the object point, N is the satellite nadir, S is the satellite and E is the geocenter θ denote $\angle$ PSN and φ denote $\angle$ PEN. Then relation between scan angle θ and φ can be calculated as follows.

$$\tan\theta = \frac{y}{H + x}$$

$$x = R\,(1 - \cos\varphi)$$

$$y = R \sin\varphi$$

$$\tan\theta = \frac{R \sin\varphi}{H + R\,(1 - \cos\varphi)}$$

$$= \frac{R \sin}{H/R + 1 - \cos\varphi}$$

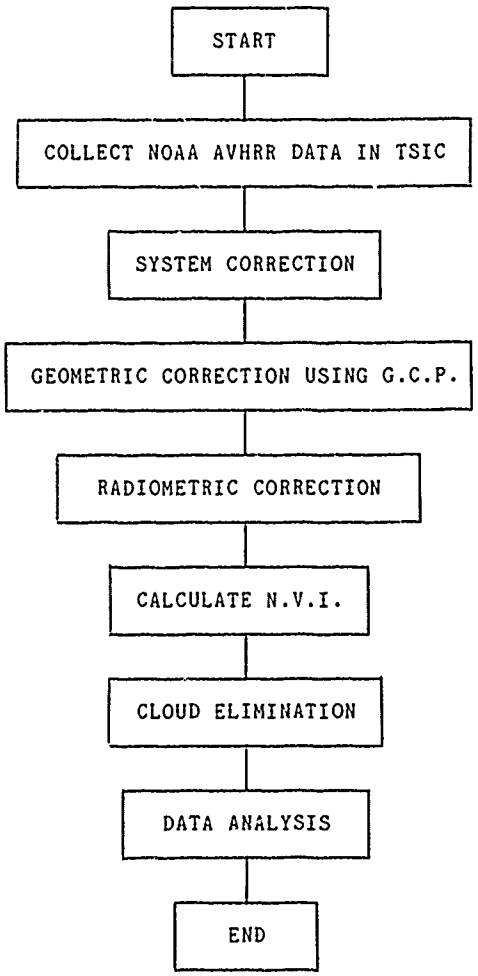


Fig.2 FLOW CHART OF PROCESSINGS

Pixels of AVHRR image are sampled with equiangle interval $\Delta\theta$. So, the ground distance corresponding to changes depend on the location of P. In this preprocessing, corrected image is resampled in constant distance corresponding to $\Delta\varphi$. This preprocessing reduced the computational load in the next step.

In the second step, geometric corrections using orbital elements were accelerated with the aid of scan and pixel functions. One dimensional 3rd and 2nd order polynomials for each function were sufficient to maintain within 1 pixel relative accuracy.

Last step is the co-registration process of images. As most of images are largely covered by clouds, cloud free areas of each image were first selected and a correlation technique was used to determine control points. Images were then superimposed with the aid of these control points. With these techniques, co-registration of image achieved within 1 pixel accuracy.

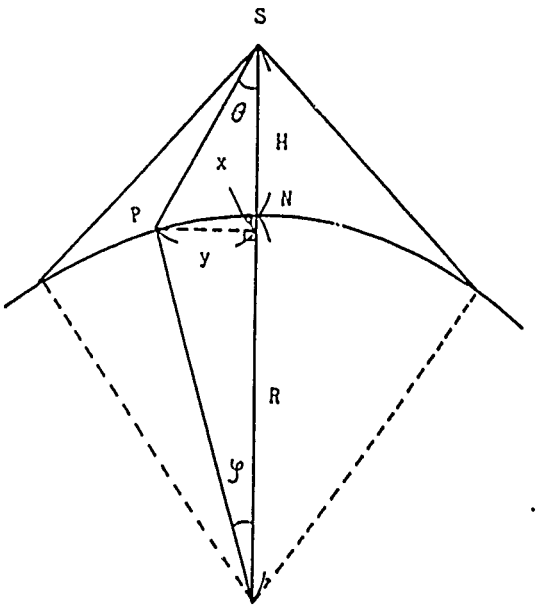


Fig. 3 GEOMETRY OF SCAN ANGLE

3.2 RADIOMETRIC CORRECTION

Radiometric distortions mainly caused by incident sun light should be eliminated. In order to eliminate radiometric distortions, sun angle corrections were first applied to geometrically corrected images. However, there exists brightness differences between different data images mainly caused by atmospheric conditions. These differences were normalized by histogram normalization process using pixels which can be thought to be in the same conditions.

4. CLOUDS ELIMINATION

Fundamental idea to generate cloud free image from images is that the channel 1 and 2 values of cloud parts are larger than those of cloud free area. However, the method using this idea also picks up shadows of clouds. In order to avoid this defect, thresholding was introduced to eliminate shadow area. Thresholding had another defect that water areas like lakes and rivers were sometimes eliminated as shadows. From these reasons, therefore, the method using original data values cloud not applied easily.

In order to eliminate clouds and shadows simultaneously, the method using N.V.I. (Normalized Vegetation Index) was used. N.V.I. can be calculated by the following equation.

$$N.V.I. = (Ch.2 - Ch.1)/(Ch.2 + Ch.1)$$

The N.V.I. is then scaled as follows:

$$\text{Scaled N.V.I.} = 240 - (N.V.I. + 0.05) \times 350$$

As shown in the above equation, NVI has larger values in cloud or shadow area where the difference of Ch.1 and Ch.2 is relatively small. On the contrary, NVI in water area has not so large value because there exist some differences between Ch.1 and Ch.2. Cloud free images could be generated by taking the area which have the smallest NVI value.

5. MONITORING OF THE PADDY FIELDS

Two kinds of cloud free images were generated for monitoring the informations of paddy growing stages. One is the false color composite image of band 1 and band 2. The other is the use of normalized vegetation index (N.V.I.).

After eliminating the clouds, comparisons of paddy fields in three terms (mid-May, early June and late June) were executed. For the interpretation of these images, MOS-1 data were used as ground truth data.

The image taken on mid-May is shown in Fig.4. This is the false color composite (Ch.1-B, Ch.1-G, Ch.2-R) image. In this image, dark tone appears in south area and red tone appears in north area. This shows that rice-planting have started in south area but there are some plants in north area.

Next, the image taken on early June is shown in Fig.5. In this image, dark tone appears in central area and red tone appears in north area and south area. This shows that rice-planting have started in central area and rice grows in south area and there were some plants in north area.

At last, the image taken on late June is shown in Fig.6. In this image, dark tone appears in north area and red tone appears in south and central area. This shows that rice-planting have started in north area and rice grows in central and south area.

From these three images, it was clarified that rice-planting moves from south to north day terms, same tendency cloud be recognized as color composite images.

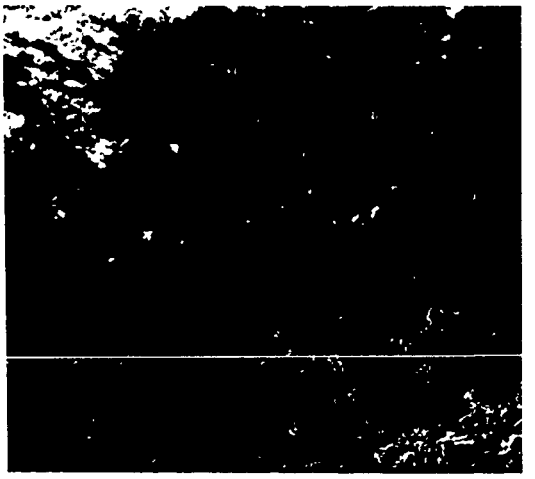


Fig.4 CLOUD FREE COMPOSITE IMAGE (MID-MAY)

Fig.5 CLOUD FREE COMPOSITE IMAGE (EARLY JUNE)

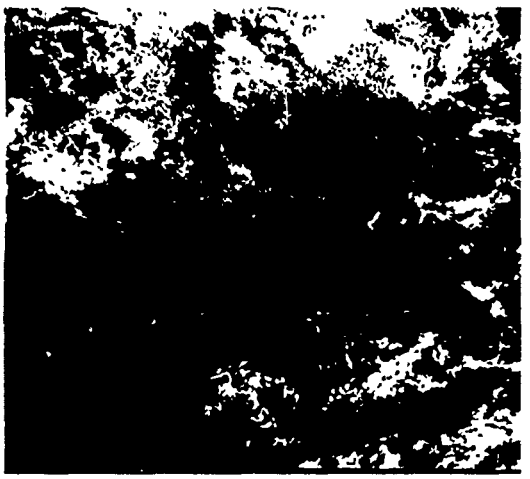


Fig.6 CLOUD FREE COMPOSITE IMAGE (LATE JUNE)

6. CONCLUSIONS

1) The regional paddy monitoring system was established.

2) The precision of system correction was improved by using tangential and earth curverture correction.

3) Radiometric correction was performed by using sun angle corrections and histogram normalization techniques.

4) Tow kinds of cloud free images were generated for monitoring the informations of paddy growing stages. One is the false color composite image of band 1 and band 2. The other is the use of normalized vegetation index (N.V.I.).

MULTI-TEMPORAL LAND USE ANALYSIS FOR SOIL EROSION IN SELECTED PRINCE EDWARD ISLAND (P.E.I.) WATERSHEDS

D.A. Wilson[1], M.L. McCourt[2] and T.M. Humes[1]

[1]Environment Canada, Inland Waters, 45 Alderney Dr., Dartmouth, N.S. B2Y 2N6, Tel. (902) 426-4197, Fax (902) 426-2690;
[2]Land Resource Information Service, P.O. Box 310, Amherst, N.S.

ABSTRACT

Prince Edward Island (P.E.I.) has recently undertaken a series of surface water programs designed to reduce soil erosion within several selected watersheds. The implementation of such programs resulted in a need to establish a cost-efficient means for monitoring and reporting on the effectiveness of these and future land/water conservation programs. To this end, SPOT and Landsat Thematic Mapper (TM) products were examined digitally and visually for their capabilities to provide information on soil erosion parameters such as micro-drainage patterns, land slope and land cover types. This information would be used to either create basic soil loss maps or to provide a yearly input parameter to the cropping management factor in the universal soil loss equation (USLE).

Digital results indicated that it was more effective to classify bare soil, low density crop, high density crop and forest cover on SPOT or TM fall scenes than to try to infer fall bare soil through identification of crop types on a SPOT summer scene. Visual analysis of the fall data provided an accurate alternative to digital classification of the same cover types. A point that illustrates the adequate resolution of and the compatibility between SPOT and Landsat TM imagery.

It is recommended that SPOT panchromatic data be visually interpreted to create potential soil loss maps where unavailable while subsequent updating for the crop management factor would be most efficiently managed through analysis of TM imagery. The choice of digital vs. visual analysis for this updating is dependent upon equipment availability, time factor involved, staff experience and budget constraints.

Keywords: P.E.I., soil erosion monitoring, Landsat TM imagery, SPOT stereo imagery, comparison study.

INTRODUCTION

Soil erosion is declared to be the Island's most serious environmental problem by the authors of the P.E.I. Conservation Strategy (1987). Soil eroded from farm fields, combined with fertilizers, herbicides and pesticides can, on entering a water course, lead to such detrimental effects as siltation of fisheries and wildlife habitat, blockage of ditches and culverts and the over-enrichment or pollution of estuaries with fertilizers and/or other applied chemicals. Thus the management of our land and water resources are inextricably linked as both rely on up to date knowledge of ongoing land use practices.

In 1987 a Canada-P.E.I. Water Resources Management Agreement was implemented under which several programs were designed to study specific aspects of P.E.I.'s water resources.

The purpose of this project was threefold: 1) to compare the contribution of TM and SPOT satellite digital and photographic data in mapping areas of potential soil erosion, 2) to assess the feasibility of using such data to monitor the progress and/or effectiveness of programs implemented under the Canada-P.E.I. Inland Water Resources Management Agreement as well as planned conservation projects and 3) to demonstrate to resource managers within the region the potential and constraints of satellite image technology.

STUDY AREA

This study focuses primarily on the Dunk/Wilmot River watershed complex (Fig. 1). This intensively cultivated portion of the Island which has a maximum elevation of 95m, dominated by orithic podzol soils and falls almost completely into the Canada Land Inventory (CLI) agricultural soils capability class II (Whiteside, 1965). This suggests that, under good management, these soils are moderatly high to high in productivity for a wide range of crops.

The Dunk/Wilmot watershed complex represents the most intensive potato producing regions on the Island. (N. Stewart, pers. comm., 1989). As potatoes are harvested mainly during late fall, the harvested fields are left to overwinter as bare soil. These soils are open to freeze-thaw cycles in late winter and early spring which accelerates the rate of erosion (Canada, 1984). This bare-soil-over-winter problem associated with potato growing is P.E.I.'s greatest single problem in soil conservation, in part due to the potential erodability of orithic podzol soils. Within the upper 42% of the Wilmot watershed alone, 18.1% of the total area was planted in potatoes in 1986 (Marenco, 1987).

Estimates of soil loss on the upper Wilmot River watershed for 1986 were 7.3 tonnes/ha/yr on a total agricultural base of 2,750 hectares. Of this, 638ha of potatoes contributed 36.6% of lost soil while 1,030ha of grain accounted for 32.0% (Marenco, 1987). The remaining cleared land was predominantly hay, pasture, soybeans and rye grass.

Climate checks have shown abnormally high October rainfall of 141.5mm in 1988 and precipitation for October 1986 totalled just 38.4mm (Env. Can.,

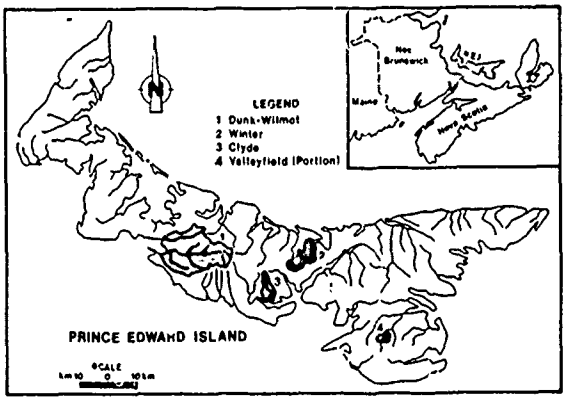


Figure 1: Location map illustrating watersheds investigated.

1988). This suggests a delay of potato harvesting of approximately two weeks in 1988, somewhat offsetting a direct comparison of the similar monthly dates, SPOT 18-Oct-88 and TM 19-Oct-86. The weather window for image acquisition in P.E.I., where determining bare soil between harvest date, snow and spring planting, is relatively narrow.

Three additional watersheds were examined briefly, near the completion of this project. The smallest of the two, the Clyde River, has similar soil capabilities as the Wilmot but is more limited in terms of actual agricultural production due to more extreme topography. The Winter River possesses even greater agricultural limitations than the Clyde River which is reflected in a reduction of land cleared for agriculture. The Valleyfield watershed was also considered as the imagery provided a springtime perspective on bare soil conditions. A marked difference between spring and fall spectral response from bare soil was noted by Wright and Morrice (1988) on Landsat MSS bands 5 and 7.

METHODOLOGY

Data Acquisition

Acquisition of satellite imagery was targeted for July and October 1986 as well as October 1988 (Table 1). The SPOT image acquired during the July growing season was examined for digital crop classification and was augmented by a ground reference map of July 1986 documenting crop type within the upper 42% of the Wilmot watershed as completed by Marenco Engineering.

Fall SPOT (1988) and TM (1986) data were acquired for the purpose of examining the capability of each data set to identify bare soil and low density crops. The intent was to identify a means of determining the varying value for the 'C' factor (cropping - management) as input to the USLE (Wischmeier and Smith, 1978). To this end both digital and photographic products were acquired. Ground reference data for the TM 1986 scene relied on the Marenco report updated by interviews with farm operators visited during the spring of 1989. For the October 1988 scene, a combination of on-site inspection data in October '88 and farm visits comprised the ground reference information.

SPOT coverage of the Clyde River watershed was acquired for 12-Oct-86, again for digital classification of bare soil and low density crop. In addition, a black and white film positive was created from the positive film color transparency of the same scene and used in combination with the 29-Sept.-87 scene for stereo-viewing on a PROCOM-2. Ground reference data for this site included a soil loss map based on the USLE, on which soil loss classes were delineated (roughly 0-0.4, 0.4-1, 1-2 and greater than 2 tonnes/ hectare/year).

Additional SPOT scenes were acquired in the east and west for the eventual creation of a SPOT photographic mosaic of P.E.I., and the Landsat thermal band 6 is linked to continuing effort toward detecting springs, soil moisture and estuary sediment loadings.

Digital Imagery

All digital image analysis was completed on a Dipix, Aries II PDP, 11/73 based system. Geographic information analysis and mapping were performed on a GeoBased geographic information system (GIS) which supports a bi-directional interface with the Dipix system. Work is in progress to establish the same link with a CARIS microvax based GIS. This new interface will facilitate faster transfer of digital thematic data to PEI government agencies which now rely on the CARIS for geographic information.

Digital analysis of SPOT and TM data followed parallel procedures. Spectral bands used for SPOT included XS1, XS2, and XS3 while those chosen for TM were that sensor's complimentary bands TM2, TM3, and TM4. This was done in order to accelerate user familiarity with agricultural land reflectances when assessing the digital or photographic products of either sensor. A three band subset was chosen for budgetary purposes as the price of an additional one or three band subset would double. As one of the primary considerations in this project was the visual assessment of bare soil, a medium infrared TM band was not chosen as no equivalent was available from the SPOT sensor.

Preprocessing of both TM and SPOT data involved a non-scaled merge of scene quadrants as the watershed geographically spanned two magnetic tapes in both data sets. This was done within the Dipix environment through geometrically resampling each half to a larger blank scene and then adding the top and bottom together. Subsequent classifications for all data sets involved creating supervised training areas for input into a maximum likelihood classifier. Cover types chosen for SPOT single date imagery acquired during the growing season included close grown crops, row crops, pasture, residential, bare ground and forest. For the fall imagery from TM '86 and again from SPOT '88 cover types were joined into four basic classes, bare soil, low density crop, high density crop and forest.

One such Dipix classification for the Clyde R. watershed was output to tape in LGSWOG format and sent to Tydac Industries Inc. for transcription onto a 13.4 cm 'floppy' diskette. Others have converted to diskette to utilize readily available microcomputers (Ringrose and Dube, 1986). The diskette was then input into Environment Canada's micro-based GIS (SPANS) for overlay analysis with a soil loss map derived from the USLE. This movement into a micro-based environment is formalized in Canada through a working agreement between Nucor Computing Resources Inc. and CCRS as producer of SPOT and EOSAT products. Nucor was contracted to

Table 1: Imagery used in soil erosion analysis (all processed to minimum level of bulk full scene georeferenced)

SATELLITE	SENSOR MODE	BANDS	FORMAT	DATE IMAGED	LOOK DIRECTION	GENERAL LOCATION
1. LANDSAT 5	TM	2,3,4	CCT/PFCT	19/Oct/1986	Nadir	85% of PEI
2. LANDSAT 5	TM	3,5,4	PFCT	19/Oct/1986	Nadir	" "
3. LANDSAT 5	Thermal	6	PFT	19/Oct/1986	Nadir	" "
.....						
4. SPOT	XS	1,2,3	CCT/PFCT	18/Oct/1988	07.67	Dunk/Wilmot
5. SPOT	XS	1,2,3	CCT/PFCT	25/Jul/1986	-02.60	R.watershed
6. SPOT	XS	1,2,3	CCT/PFCT	01/Sep/1988	10.02	W 3rd - PEI
7. SPOT	XS	1,2,3	CCT/PFCT	02/May/1987	-02.30	Valleyfield R watershed
STEREO-PAIR:						
8. SPOT	XS	1,2,3	CCT/PFCT*	12/Oct/1986	-26.88	up. Wilmot & Clyde R.
9. SPOT	P	1	CCT/PFT	29/Sep/1987	17.52	Up. Wilmot Clyde & Winter R. watersheds

TM - Thematic Mapper
SPOT - Satellite Pour l'Observation de la Terre
XS - Multispectral
* Created black & white film positive transparency for stereo viewing on Procom-2 device.
P - Panchromatic
CCT - Computer Compatible Tape
PFCT - Positive Film Colour Transparency
PFT - Black & White Film Transparency

provide a transcription of a 512 pixel by 512 pixel window of SPOT raw (features) data for display and viewing on GIS (SPANS). If feasible, a yearly update could be performed to monitor areas of bare soil located on land highly susceptible to soil erosion.

Output map products from digital analysis were both digital and photomechanical in nature. In digital map creation, Dipix classifications were output to tape and transferred to a GIS environment for plotting to planimetric base maps. In most cases, bare soil and low density crop were the only thematic classes to be plotted as these were of primary importance to soil erosion monitoring. Area statistics of the mapped results were also compiled.

Photomechanically produced map products of the same thematic information were created to enable a cost comparative analysis with digital output products. In producing photomechanical maps, the land classes of bare soil and low density crop were plotted on to an ACT II ink-jet plotter with two vastly different shading densities. These plots were then overlaid (no scale change was needed) with contoured mylar base maps at scales of 1:50,000 and 1:125,000 and then photographed to produce non-digital, black and white thematic maps.

Visual Imagery

All visual interpretations of false color composites (FCC) were completed using a Aviopret stereoscope, a standard mirror stereoscope and a Procom-2. The Procom-2 is an optical stereo-transfer device suited for thematic change detection and mapping using satellite imagery as it provides a magnification range from 3x to over 140x. This stereo capability and the magnification range facilitates the overlaying of up to two images of varying scales onto a base map/image of a third scale. It also enables three-dimensional viewing through the use of red and green colored lenses. Panchromatic image transparencies must be used to obtain a stereo view from an anoglyph. Thus a black and white positive transparency was created of the 12-Oct-86 colour composite. The panchromatic stereo-pair (Table 1) was used in an attempt to target potential areas of high soil erosion.

FCC's of October scenes TM '86 and SPOT '88 (185 mm format) were registered to each other and, through off/on viewing of individual fields, changing land use patterns were identified. Conventional photointerpretation parameters of tone, shape, texture and association were used in the analysis.

In the Dipix environment the October TM infrared band 4 was resampled to 20 metres and geometrically registered to the October SPOT band 3. With TM band 4 assigned to the red gun, SPOT band 3 assigned green and the blue gun turned off, it was possible to visually interpret changes in cropping patterns on the colour monitor. This rendition was then output to tape for transfer to a colour recorder and film transparency.

On the Dipix CRT monitor band 3 of the SPOT 18-Oct-88 scene was displayed in black and white and "micro-drainage" was manually traced with the system cursor. A 35mm slide was taken and through the use of the Procom-2 was compared to micro-drainage as mapped on a 1:25,000 NTS sheet and on a 1:35,000 black and white aerial photograph taken 02-May-86. This was an attempt to identify a stream "dissection class" parameter to be considered when indexing small river basins for susceptibility to erosion as described by Schumacher (1986). This author was considering 1:65,000 aerial photography when describing the other "readily observable features" in this erosion hazard index; slope topography and land use.

The Procom-2 was also used to determine a general percentage of grain stubble, a conservation measure, that could be detected on the Oct. '88 SPOT scene. The process was based on a limited number of fields checked this spring.

RESULTS

Digital

As shown by Table 2 the highest individual accuracy achieved for crop classification on the SPOT summer scene was for row crops at 84.5%. Due to the poor performance of the remaining classes overall agreement dropped to 59.4%. This falls well below the general guidelines set by Anderson et. al. (1976) whereby a land cover classification should achieve an overall accuracy of 85% to be useful.

Thus it was deemed impractical to either identify dynamic crop rotation patterns or to infer fields to be left as bare-soil-over-winter through identification of crop types on single date SPOT summer imagery. Also the increasing practice of planting winter grains as a conservation practice reduces the reliability of such an inference. This conclusion was not unexpected as Thomas et. al. (1986) and Bernier et. al. (1987) recommend using multi-temporal data when monitoring a diverse agricultural system.

As an alternative approach, both TM and SPOT, data obtained after fall harvest, were examined for their ability to map bare soil on a single date. Land cover classes for use in supervised training were restructured to fulfill the requirements for identifying areas of potential soil loss. As such, cover types included: (a) bare soil, (b) low density crops, e.g., harvested grains or other areas where soil reflectance comprised a major portion of the reflected spectral energy, (c) high density crops such as pasture, hay, idle land, and winter grains and (d) forest cover.

Table 3 shows the overall agreement of SPOT Oct. '88 data to be 74.6%. The primary problem causing this low agreement was related to confusion between low density and high density crops. This problem may have been related to the assessment procedure whereby ground reference was input in blocks associated with field boundaries. As such, all pixels in a field of high density were assigned a ground cover class of high density regardless of whether or not the field contained any low density pixels. If separation of these classes was not required by the resource manager the overall accuracy would rise to 93.4%.

Table 2. Assessment of classification accuracy for original land cover classes. SPOT 25-July-86.

	Classification Results								
Cover Type	Close Grown	Row Crop	Pasture	Resident.	Bare Soil	Forest Cover	Total	% Correct	% Omission Error
Close Grown	13,254	1,129	7,916	483	12	269	23,063	57.5	42.5
Row Crop	1,290	8,756	222	94	-	-	10,362	84.5	15.5
Pasture	3,250	373	6,306	661	-	30	10,620	59.4	40.6
Residential	8	83	20	156	28	-	295	52.9	47.1
*Bare Soil	442	77	278	90	-	-	887	0	100
Forest	2,625	43	1,263	4	-	1,785	5,720	31.2	68.8
Total	20,869	10,461	16,005	1,488	40	2,084	50,947		
%Commission Error	36.5	16.3	60.6	89.5	100	14.3			

Overall Agreement = 59.4%

* As ground truth did not exactly coincide with overpass date it is believed that these areas supported some vegetation at the time of image acquisition causing class confusion.

Table 3. Assessment of classification accuracy for generalized land cover classes, SPOT 18-Oct-88 and TM 19-Oct.-86

	Classification Results													
Cover Type	Bare Soil		Low Density Crop		High Density Crop		Forest Cover		Total		% Correct		% Omission Error	
	SPOT	TM	SPOT	TM	SPOT	TM	SPOT	TM	SPOT	TM	SPOT	TM	SPOT	TM
Bare Soil	1,870	2,635	143	292	-	2	3	4	2,016	2,933	92.8	89.8	7.2	10.2
Low Density Crop	-	102	4,312	743	2,892	81	270	5	7,474	931	57.7	79.8	42.3	20.2
High Density Crop	-	-	716	6	5,736	826	290	5	6,742	837	85.1	98.6	14.9	1.4
Forest Cover	1	1	38	61	3	23	878	937	920	1,022	95.4	91.7	3.6	8.3
Total	1,871	2,738	5,209	1,102	8,631	932	1,441	951	17,152	5,723				
% Commission Error	-	3.8	17.2	32.5	33.5	11.3	39.1	1.4						

Overall SPOT Agreement = 74.6%
Overall TM Agreement = 89.8%

* Due in part to assessment procedure which did not take into account low density pixels occurring in high density fields.

Thematic Mapper data had an increase in overall agreement over SPOT by 15.2% (Table 3). This was primarily related to the improved performance of the low density crop class and may be a result of the larger pixel size having a reduced tendency to identify areas of low density throughout high density fields.

Overall it was evident that both sensors were capable of monitoring bare soil vs areas with some type of vegetation cover including grain stubble. The Thematic Mapper has a definite advantage over the SPOT sensor in its ability to capture 85% of the Island in one pass. A TM sub scene required a third less image processing 'machine time' than did a SPOT sub scene based on a supervised classification of a 1:50 000 map sheet. Although SPOT has an increased revisit capability, scenes acquired over several dates require more extensive ground reference data. This same increased spatial resolution simplified the linking of reference data on a sub field level with SPOT sub scenes.

Of note was the visibly lower return from fields in bare soil which had been recently harvested for potatoes as opposed to the bright return of other fields in bare soil. This discrimination capability suggests the ability to monitor such crop rotation as potatoes, grains and either hay, pasture or idle land using SPOT or TM post harvest

data. A study in Scotland (Wright and Morrice, 1988) noted a marked increase in the spectral return from bare soils in May on Landsat MSS bands 5 and 7 when compared with similar bands in October. When comparing the acreage of land in bare soil for NTS map sheet 11L/5 it was found that the 1988 fall season had 1406 more ha of exposed land. For the entire Tm 1986 scene (92% of the Island) 56 689 ha or 11% were classed as bare soil.

A cost comparison of digital vs photomechanical output products for map sheet 11L/5 showed that for plotting bare soil and low density crop at a scale of 1:50,000 the digital plot could be as much as 75% more expensive. Of this cost 60% was incurred in creating a digital base map, a one time expenditure which would not be repeated in the yearly updates. Although more expensive initially, GIS mapping allows for yearly analysis and updating which is not readily available through photographic products.

Visual

Photo interpretation of the satellite FCC's showed it was possible to accurately differentiate between the four cover types chosen for digital classification. Analysis of image texture, feature shape and association, enable differentiation between bare soil agricultural and bare soil non-agricultural which was not possible in digital image analysis.

Much of the visual interpretation of SPOT and TM imagery was done from the digital image analysis system (DIAS) due to the enhanced clarity of its monitor. It proved easier to detect waterbodies, of less than 20 meters wide, on SPOT XS 3, than on the panchromatic mode having twice the spatial resolution. Additional ground reference information is needed to determine a correlation between small ponds and a base flow from groundwater input.

Visual interpretion of micro-drainage on SPOT XS3 showed that, of the 76 stream branches in the Dunk/Wilmot watershed complex, 60.5% correlated directly with the 1:25,000 NTS map and another 25% appeared through interpolation of contour level giving a total of 85.5% accuracy. Such results could not be duplicated within the Winter River watershed where the increase in shrub and forest and subsequent reduction of agricultural land made visual delineation more difficult. These inconsistent results indicate that a stereographic view is needed to attempt the delineation of sub-field drainage patterns.

Evaluation of SPOT XS and P mode image stereo-pairs on the Aviopret and a panchromatic stereo-anaglyph on the Procom-2 having two projectors was undertaken. Past experience in a nearby area (Wilson et. al. 1988), indicates that these parameters relating to the erosion hazard index (Schumacher, 1986) could be delineated providing SPOT stereo-pairs having the maximum look-angles are available. The Procom-2 offered the advantage of simultaneous interpretation and mapping. With the stereo model having the highest combined look-angles of 17.52 degrees and -26.88 degrees, only elevation differences approaching 20 metres could be perceived, based on attempts to create contours lines from the anaglyph. The reduced quality of the black and white film positive created from the colour October, 1986 scene was a factor in lowering the perceived vertical exaggeration, and thus the capability to determine slope angles. Past experience has demonstrated the correlation between larger look-angles (up to -30.37 degrees) and increased vertical perception. (Wilson et. al., 1988). For a regional targeting program to be successful, smaller contour intervals than those determined here are required.

Given these available look-angles it was possible, from the Procom-2 produced anaglyph, to extend the drainage patterns beyond some of those noted on a 1:50 000 topographic map. Questionable cases were resolved with a mirror stereoscope.

The Aviopret stereoscope offered a good three dimensional, high definition overview of 33 km at 3x, 13km at 8x and 6 km at 15.5x. Due to a "fish bowl" effect, it became difficult to determine drainage direction after 7x, based on the SPOT stereopair (P and XS) listed in Table 1. Perceived stereo view tended to be lost over fields of high spectral reflectance.

As others have noted (Chevez and Bowell, 1988), the complementary rather than duplicative value of merging Landsat TM and SPOT imagery, was illustrated with the P.E.I. study. The bands selected provided a practical and readily comparable multi-temporal data set dealing with soil loss. As well, no difficulty was noticed in merging georeferenced images in the Procom nor matching projected images to base maps of varying scales including 1:10 000.

The banding most noticable, especially over water, on the fall scene (due to the added contrast stretch used, according to CCRS data accompanying CCT) presents less of a problem to visible interpretation than digital methods. The ease of dealing with this banding, horizontal in the TM case and vertical in the SPOT scenes, represents an advantage of employing film image products. The degree of banding is a major factor when considering the creation of an image mosaic.

Visual interpretability of film products proved advantageous, even when setting up trainings areas for digital classification. The Dunk/Wilmot watershed complex proved to be a dynamic test area difficult to classify. The shifting crop calendar and narrow climate windows for image acquisition, favour the direct approach of targeting potential soil loss sites using imagery acquired before planting and after harvest. The method offers managers a simple means of monitoring programs, a need stated by Vessey and Stewart (1986).

Analysis Technique - Advantages

Digital: capable of providing map of user required classifications (graphic and statistical) in an efficient time frame
- with link to a GIS can provide resource managers with powerful analytical capabilities in a a consistent, repeatable format.
- split screen (DIPIX system monitor) provides useful means of comparing images on full pixel resolution in initial analysis.
- use of imagery and microcomputer environment.

Visual: has stereo viewing capabilities with SPOT data.
- computer independent, i.e., no specialized computer personnel required.
- allows human reasoning power with photo interpreters consideration of tone, texture, shape and association.
- use of a transfer device such as the Procom allows for comparative viewing (3 data sets at once), interpretation and mapping of satellite data in one step.
- less expensive than digital data.
- hard copy is immediately available for field use.

Analysis Technique - Disadvantages

Digital: requires access to an image analysis system (and possibly a geographic information system) with specialized personnel.
- digital image analysis does not consider image texture, shape or an association when classifying, i.e. totally dependent on image tone.
- more costly than photographic data.

Visual: may be labor intensive and time consuming.
- limited for analytical capabilities when introducing more than one layer of information unless digitized/scanned and entered into a GIS.

RECOMMENDATIONS and CONCLUSIONS

1. That Landsat (TM) and SPOT (XS and P) pre-planting or post-harvest imagery affords resource managers a practical means of monitoring the effectiveness of present and future land/water conservation programs in P.E.I.

2. That SPOT panchromatic film transparencies (stereopair having positive and negative look angles of 27 degrees or over) be used with a Procom-2 to target and delineate areas of high potential soil erosion for regions not presently mapped with the USLE. Subsequent to this, digital or photographic TM imagery would be used on a yearly basis to monitor soil erosion and crop rotation patterns.
3. That crop density classes should be established using SPOT XS imagery. A means of estimating the long-term reduction in agricultural productivity based on crop density/crop vitality, may result.

4. That May SPOT (XS and P mode) imagery be evaluated for the detection of sub field drainage patterns and evidence of soil erosion over winter.

5. Future research on this project should focus on the existent/non-existent correlation between the TM band 6 (thermal infrared) and soil moisture or estuarine sediment loadings.

ACKNOWLEDGEMENTS

The authors wish to thank Kent MacRae of the PEI Fish and Wildlife Branch and Bryan Pellerin (contractor) for their field documentation. Appreciation is expressed to John Corning and Terry Kennedy of Land Resource Information Service for the linking of imagery to GIS data, John Keefe and Joe Arbour for editorial advice and Maxine Heckbert for manuscript preparation. Finally, the provision of much of the imagery by CCRS as part of the SPOT Technology Enhancement Program, is acknowledged.

REFERENCES

1. Anderson, J.R., E.E. Hardy, J.T. Roach and R.E. Witmer, "A land use and land cover classification system for use with remote sensor data", Prof. Paper 964. U.S. Geol. Surv., Washington, D.C., 28 pp., 1971.
2. Bernier, M., O. Dupont, M. Carignan and J. Cihlar, "Evaluation des donnees de SPOT pour l'agriculture au Quebec", Eleventh Canadian Symposium on Remote Sensing, Waterloo, pp. 155-164, 1987.
3. Canada, "Soil at risk", A Report on Soil Conservation by the Standing Committee on Agriculture, Fisheries and Forestry in the Senate of Canada, pp. 129, 1984.
4. Chavez, P.S. Jr., J.A. Bowell, "Comparison of the spectral information content of Landsat Thematic Mapper and SPOT for three different sites in the Phoenix, Arizona Region", Photogrammetric Engineering and Remote Sensing, Vol. 54, No. 12, 1699-1708, 1988.
5. Environment Canada, Monthly Meteorological Summary, Principal Station Data, Summerside, Atmospheric Environment Service, Ottawa, 1988.
6. Marenco Engineering, Ltd., A Soil Erosion Control Plan for the Upper Wilmot River Watershed, 33 pp., 1987.
7. Province of Prince Edward Island, A Conservation Strategy, 115 pp., 1987.
8. Ringrosa, S., P.O. Dube, "The use of the darkening effect (red absorbancy) to monitor range degradation in the southeast Botswana Kalahari", I.S.P.R.S. Conference, I.T.C., pp. 37, 1986.
9. Schumacher, A.E.A., "Indexing small catchments in Java, Indonesia, with respect to their relative susceptibility to erosion", Tenth Canadian Symposium on Remote Sensing Edmonton, pp. 705-714, 1986.
10. Thomas, H.C., F.L. Scarpace and T.M. Lillesand, "Use of multi-temporal spectral profiles in agricultural land-cover classification", Photogrammetric Engineering and Remote Sensing, Vol. 52, No. 4, pp. 535-544, 1986.
11. Vessey, G., N. Stewart, "Watershed management - a new approach to an old problem", Proceedings of Freshwater Sediment Issues Workshop, Environment Canada, p. 31, 1986.
12. Whiteside, G.B., "Soil Survey of P.E.I.", Canada Department of Agriculture and P.E.I. Department of Agriculture, 70 pp., 1965.
13. Wilson, D.A., M.L. McCourt, M. Poirier and J. Seel, "An evaluation of the applicability of SPOT stereoscopic imagery to the Canada land use monitoring program", SPOT 1 Image Utilisation Assessment Results, Paris, pp. 551-557, 1988.
14. Wischmeier, W.H., D.D. Smith, "Predicting Rainfall Erosion Losses", Soil Conservation Service, U.S. Department of Agriculture, Agriculture Handbook No. 537 Washington, D.C., 1978.
15. Wright, G.G., J.G. Morrice, "Potato crop distribution and subdivision on soil type and potential water deficit, an integration of satellite imagery and environmental spatial data base", Int. J. Remote Sensing, Vol. 9, No. 4, pp. 683-699, 1988.

1986

AUTOMATIC DETERMINATION OF THE TOPOGRAPHIC FACTOR REQUIRED IN THE UNIVERSAL SOIL LOSS EQUATION

D. Benmouffok and D.N.H. Horler

HORLER INFORMATION INC.
Suite 704, 116 Albert Street
Ottawa, Ontario, Canada K1P 5G3

R. Leconte

Canada Centre for Remote Sensing
1547 Merivale Road
Ottawa, Ontario, Canada K1A 0Y7

ABSTRACT

Soil erosion from agricultural land represents a loss of nutrients and fertility. Efforts towards controlling and managing erosion by water require accurate and current data. The Universal Soil Loss Equation (USLE) gives the predicted soil loss rate in tonnes per hectare per year as a function of the rainfall, topography, soil cover and conservation practices. The paper describes the development of an automated methodology for obtaining the topographic factor, LS, which reflects the landform contribution to the potential soil loss calculation. This factor, which regulates the potential erosion from an exposed area, is derived from slope steepness, slope length and aspect which are directly obtained from a Digital Elevation Model (DEM).

Keywords: Soil Erosion, Catchment Basins, Digital Elevation Model, Hydrology

1. INTRODUCTION

Soil losses by erosion are both economically and environmentally costly. They represent the loss of arable soils and nutrients, as well as the pollution of watercourses and lakes by suspended or dissolved matter, especially, nitrogen, phosphorus and many toxic chemical and biological constituents. It was reported (Snell, 1984) that: "A recent study conservatively estimates annual erosion costs in Ontario attributable only to yield, nutrient and pesticides losses to be $68 million... at current rates of soil loss, all of the topsoil on soybean and corn-growing land in south-western Ontario will be gone in 40 to 60 years."

Cihlar (1987) reported that soil erosion is becoming increasingly significant in the maintenance of crop productivity in Canada and is now considered to be a national problem for this and future generations. He added that the location, extend and rate of erosion must be determined before implementing erosion control programs in an attempt to maintain soil erosion rates within tolerable limits.

The commonly used method to compute soil erosion is described in the manual cited above (Snell, 1984). The method is wholly manual, input data are graphically extracted from existing topographic and thematic maps, and output data are qualitative thematic maps. These outputs can help a decision-maker take appropriate remedial measures at the regional level, but do not allow the quantification of sediment volumes reaching streams or the location of most water quality problems. The labour cost of mapping by this method, including production of three maps , was estimated to be $0.12 (1982) per hectare. However, the erosion process is highly dynamic and needs periodic if not continuous monitoring.

Developing from this classical method, new advances have been made. Cihlar (1987) described a methodology which integrates information derived from satellite imagery and aerial photographs to map and monitor soil erosion by water.

Recently, Horler Information has developped a methodology for the automatic delineation of catchment basins from DEM and calculation of the topographic factor, LS. The development of the software tools has been completed and the process tested using DEMs provided by the Canada Centre for Remote Sensing (CCRS) for test areas in Oxford County, Ontario, and near Sherbooke, Quebec.

This paper gives a summary description of the software and illustrates its functionality with a concise example.

2. MODEL DESCRIPTION

The process modelled by the system is to find that part of a watershed basin that acts as catchment for surface water draining to any defined point on the map, and obtain topographic parameters required by the USLE

for each pixel of this area. The USLE gives the predicted soil loss rate as a function of the rainfall, landform topography, soil cover and conservation practices (Wischmeier and Smith 1965):

$$A = R*K*CP*LS$$

where A = predicted annual soil loss
R = rainfall and snow factor
K = soil erodibility factor
CP = control practice factor
LS = topographic factor.

The topographic factor is calculated from the (Wishchmeier and Smith, 1965) formula

$$LS = (L/72.6)^{m}((430x^{2}+30x+0.43)/6.613)$$

where L = slope length in feet
x = sin θ (θ is slope angle).
m = an exponent dependent upon slope

The model developed consists of fourteen programs assembled in three groups.

The overall flowchart of the software is shown in Figure 1.

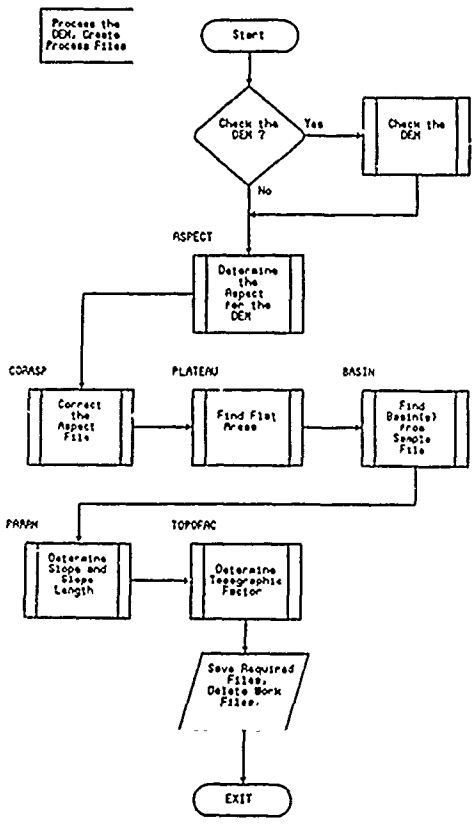


FIGURE 1. Catchment Basins System

The first group (Utilities) contains the modules which allow the user to manage the required data. This is useful because it gives the user the ability to select working zones from larger areas, to simulate different landforms by modifying pixels values or to check visually areas of interest.

The second group (Preprocessing) includes the modules which check and correct the source DEM. The automatic checking is done only on local one-cell extrema in the DEM. Correcting involves smoothing the DEM by filtering or removing one-cell extrema and replacing them by the median value of their eight neighbours. The purpose of this process is to remove gross errors due to lack of interpolation of the DEM generating algorithm. The menu offers a combination of filtering, maxima removing and minima removing.

The third group (Processing) comprises the modules which identify flat areas, delineate watershed catchment basins, construct the drainage net and extract topographic parameters required by the USLE for given samples from the DEMs.

The programs of the Preprocessing and Processing groups are packaged in a batch or can be run individually.

3. APPLICATION

The model was applied to a DEM of the region of Waterville, Quebec, located in the eastern townships approximately 15 km south of Sherbrooke. Figure 2 illustates the topography of the study area. The DEM was produced by the Gestalt Photo Mapper from stereo pairs of aerial photographs taken in 1985 at a scale of 1:50 000. The DEM covers an area of 5 km x 5 km, with a grid size of 50 m.

The area under study is physiographically part of the Appalachian highlands. The Coaticook river, a tributary of the Massawipi river, meanders through the study area. A variety of relief types are present, with very steep slopes bordering the Coaticook river north of Waterville and in the northwest of the area, moderate to steep slopes in the northeast and southeast, and a fairly flat area in the southwest. A total of seven sample points were selected in the area for the analysis. Five of them are located along the river, one inside the flat area and one at a summit (Figure 2).

4. METHODOLOGY

The DEM was first cleaned of one-cell pits and one-cell summits. Cleaning is necessary in order to avoid the generation of small watershed basins characterizing local surface runoff. The cleaning was performed using the "correct" option available in the

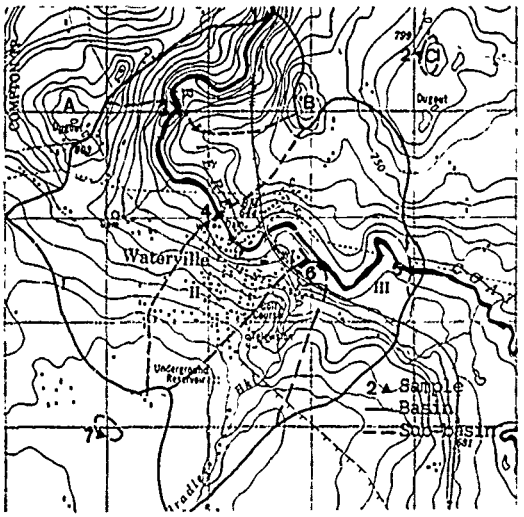


Figure 2. Study area

window of 3 by 3 pixels surrounding them. The resulting DEM was then filtered by a 3 by 3 weighted moving mean to smooth the local relief.

The model was first run using only sample #1, which is located at the lowest point of the Coaticook river. A very narrow basin was produced, the watershed divide surrounding the Coaticook river and its tributaries. The model got trapped in local relief zones, which resulted in a total basin area much smaller than anticipated. A glance at the topographic map revealed that rather than passing through the gorge, the model was stopped by the gorge between the two hills labelled A and B on Figure 2. The aspect file, showing the direction of the surface runoff, had a few cells with north to south flow that resulted in the misplacement of the watershed divide. Local elevations of a few cells resulted in misalignment of the flow directions. The aspect file was edited and local flow paths were interactively modified. Overall and for the seven samples used in the study, the aspect of some fifty cells (out of about one million) was modified to better represent the large scale surface runoff process.

5. RESULTS AND DISCUSSION

The watershed boundaries produced by the system are illustrated in Figure 2. The model produced seven distinct basins. As expected, basin #1 has the largest area, since the sample point is located at the river's outlet. Basin #5 is included in #4, and #4 is included in #3, etc..., the order being related to the relative location of the sample points along the Coaticook river. The basin area #1 correctly passes over the top of hills A, B and C (see Figure 2). However, the overall area is smaller than one would expect from the topographic map. For example, the watershed divide abruptly deviates to exclude the flat area where sample #7 is located. Likewise the divide leaving hill C branches off to the south rather than continuing eastward. A possible reason for this discrepancy is presented in the next section. Finally, it would have been interesting to see if two sample points located in different streams of the same order would have generated two distinct watersheds which do not overlap. Unfortunately, The DEM used here did not include a river network for verifying that particular situation.

Figure 3. Changing the aspect of a single cell

The results obtained here indicate that both the shape and the size of drainage areas are very sensitive to the main direction of the surface runoff, which is well described by the aspect of a cell. The aspect is obtained from elevation data of the neighbouring cells. Thus, it appears that the accuracy and the scale of a DEM are important factors to take into account in automatically delineating watershed boundaries. For example, a DEM which contains gross errors, e.g. one-cell extrema, may modify the pattern of the surface runoff. Also, a fine resolution DEM contains small scale topographic variations that may alter the location of the watershed divide. Consider the example shown in Figure 3.

On Figure 3, the aspect of each cell is represented by an arrow pointing in the direction of the main flow. Changing the aspect of a single cell (shown on Figure 3 (a) and 3(b) by a bold arrow) substantially modified the location of the watershed divide. Consequently, great care should be exercised in generating a DEM and selecting an appropriate DEM scale for the analysis. Coarse grids will eliminate local relief effects and help in delineating large watershed systems. Removing local extrema (one-cell pits and one-cell summits) and smoothing the landforms by filtering also help in obtaining DEMs suitable for further processing. It must be noted that that human intervention is undoubtedly required to produce acceptable results. DEMs generated automatically either from stereo pairs of aerial photographs or from satellite images are produced with a minimum of human interpretation, and thus are more subject to ill-defined local points. For example, tree canopy elevations may be measured rather than the true ground elevations, causing distortion in the actual topography. DEMs generated from topographic maps are less subject to those errors, since

the generation of the map involved the intervention of a photo-interpreter, and a generalization of forms is introduced.

In addition to generating basins divides, the software system also generated other end-products. Among them, the drainage network and the topographic factor files are worth mentioning. The drainage network file gives for each cell the number of cells draining into it. The drainage network gives an overall picture of the preferential flow paths and should closely correspond to the streamflow network. Figure 4 shows a small portion of the flow paths produced from the drainage network file. Numbers are the number of pixels draining into that pixel including the pixel itself. Zero is assigned to pixels which are outside the watershed.

	45	46	47	48	49	50	51	52
2	1	1	1	1	1	13	3866	0
3	2	2	2	2	6	3861	12	4
4	3	3	3	9	3834	4	12	11
5	16	20	3805	3821	4	6	3	11
6	3	3777	3798	3	3	3	5	2
7	1	3773	2	1	2	2	2	4
8	2	3769	3	2	1	1	1	1
9	4	3765	7	6	5	4	3	2
10	3729	3755	2	1	1	3	2	1
11	3666	55	19	18	16	12	11	10

Figure 4. A small portion of the drainage network

Overall, the correspondance between the actual and the simulated networks is satisfactory, but a discrepancy exists in the southern part of the DEM. This was expected, since basin #1 as created by the model does not include all of the Coaticook river.

The LS factor is used in the Universal Soil Loss Equation (USLE) for the calculation of potential erosion. According to the USLE, the steeper and the longer the slope, the higher the potential for erosion, assuming that all other factors, rainfall, land cover, soil conservation practices and soil types, are spatially homogeneous. The results obtained here confirm that high LS values exist in the steeper portions of the study area (north section) and lower values in areas of mild topography. Higher values occurred especially along the Coaticook river, because of the presence of the gorge. The LS file, by itself, does not indicate the severity of erosion, but gives an appreciation of the relative importance of the topographic factor in the overall erosion process. Factors such as the ground cover types and conditions are also important in estimating the erosion potential of an area.

	45	46	47	48	49	50	51	52
2	2267	1453	2017	2916	3026	353	1276	-99
3	2172	2531	3443	1957	750	1301	2531	3561
4	2078	3026	1668	499	1301	3100	3288	4251
5	1957	1640	-3	1401	4514	4469	3561	3561
6	2298	699	485	3641	5248	2809	1668	2172
7	1796	699	1285	4417	3443	1301	2172	4875
8	2172	433	2068	4417	2601	2078	3803	6081
9	2267	750	2478	-2	4578	2417	2601	4026
10	1543	-3	2298	3470	4208	4603	3470	3724
11	-3	717	2017	2601	2922	3328	2922	3293

Figure 5. Topographic factor (LS)

On figure 5, the LS factor is expressed in per thousand. The software takes into account the definition domain of the USLE which is between 3 and 20 percent for slopes, and up to 120 m for slope length; -2 and -3 are codes which indicate different situations which do not meet the definition domain and stand respectively for slope above the maximum and slope below the minimum allowed

The approach initiated in this study could be further expanded to include other parameters of importance in the erosion process. An obvious area of further development of the work would be the development of a soil erosion modelling package using the present software as a starting point. Other applications are also envisaged, including pollutant and chemical transport, geochemical prospecting and hydrological modelling.

ACKNOWLEDGEMENTS

The authors thank [illegible] and Dr L. St-Laurent, both [illegible] for their contributions to this [illegible] Horler Information acknowledges [illegible] vided by the Unsolicited Proposals [illegible] of the Department of Supply and Se[illegible]anada, by CCRS and by the Water Reso[illegible] Branch of Environment Canada.

6. REFERENCES

1. Cihlar, J., "A methodology for mapping and monitoring cropland soil erosion", Can. J. Soil Sci., Vol. 67, pp433-444, 1987.

2. Snell, E.A, "A manual for regional targeting of agricultural soil erosion and sediment loading to streams", Environment Canada, Lands Directorate, Environmental Conservation Service, Working paper No.36, 1984.

3. Wischmeier, W.H. and Smith, D.D., "Rainfall erosion losses from cropland east of the Rocky Mountains", Agricultural Handbook No.282, U.S. Department of Agriculture, Washington, D.C., 1965

1990

APPLICATION OF REMOTELY SENSED AND GEOGRAPHIC INFORMATION SYSTEM DATA, TO QUANTITATIVE ASSESSMENT OF LANDSLIDE DAMAGE.

C M Trotter, P R Stephens, N A Trustrum, M J Page, K S Carr and R C De Rose.

Division of Land and Soil Sciences
Department of Scientific and Industrial Research
Private Bag
Palmerston North
NEW ZEALAND

ABSTRACT

In March 1988, Cyclone Bola caused severe landslide damage over a 12,000 km^2 area of the North Island of New Zealand. In submitting disaster compensation claims, landholders were required to estimate the amount of landslide damage to pasture. This damage was to be further categorised as either temporary or permanent. As part of the compensation programme, an independent damage assessment was also undertaken for more than 100 properties throughout the region. The assessment obtained total bare ground area from classified panchromatic SPOT imagery. However, the landslide scar and debris-tail components (equating to long-term and temporary damage to pasture respectively) could not be resolved on the basis of radiance levels. These components were instead determined for representative landforms in the region by a combination of photo-interpretation and digitisation classification, of 1:5000 scale aerial photographs. An estimate of the average ratio of scar to debris-tail area for an individual farm was then obtained by intersection of a regional landform map and a farm boundary overlay, using a geographic information system.

Key words: remote sensing, SPOT, image processing, GIS, landslides.

INTRODUCTION

In March 1988, Cyclone Bola caused damage in excess of $NZ120M to the North Island, New Zealand. Damage to the north-east of the North Island - from East Cape to Tutira (see Figure 1) - was particularly severe. In this region, rainfalls of up to 900 mm were recorded over a three day period, resulting in widespread flooding and siltation of lowlands, and extensive damage to road and rail transportation systems. Much of the affected hinterland is pastorally farmed hill country, which suffered severe landslide erosion.

In providing disaster-recovery assistance to landowners, the Government sought not only to compensate for the direct effects of the damage incurred, but also to provide some compensation for loss of future earnings foregone as a result of such damage. For hill country landowners, this meant that an important component of submitted damage claims was the visual estimation of the area affected by landslides, and the categorisation of that area as representing either long-term or temporary production loss. Temporary damage to pasture occurs in those areas covered by landslide debris-tails, with recovery to former productivity levels generally taking place within 1 to 3 years. Productivity on the landslide scar (source) area, however, remains severely depressed over a much longer term, often of more than 20 years duration (Lambert *et al.*, 1984).

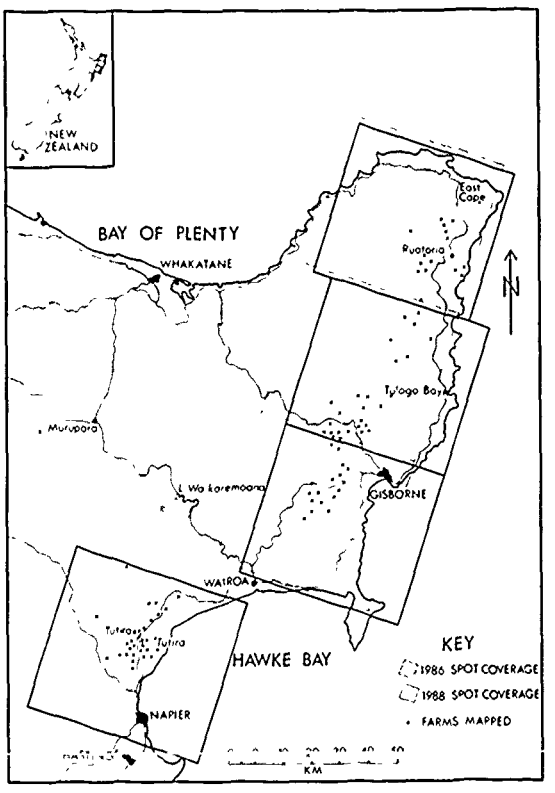


Figure 1. Locality map of the region damaged by Cyclone Bola, showing coverage of SPOT satellite scenes and location of farms assessed for landslide damage.

In view of the well-known inaccuracies in visual estimation of landslide area by inexperienced observers (farmers), some 100 properties spread throughout the region were nominated for independent, quantitative damage assessment. The principal purpose of this assessment was to provide data for overall quality control during processing of damage claims. An important secondary issue became the provision of objective data in those instances where the claims submitted by individual landholders appeared to be excessive. It must be emphasised that the approach adopted to providing quantitative damage data was strongly constrained by the requirement that damage claims be processed rapidly. This meant that use had to be made either of data that could be obtained readily, or of data that was already in existence - pre-disaster data that could be correlated with the post-disaster situation.

AN APPROACH TO DAMAGE ASSESSMENT

The landslide damage information required by those responsible for administering compensation payments was the amount of both temporary and long-term pasture damage, on a farm-by-farm basis. Compensation was to be paid only for the damage due to Cyclone Bola. That is, any damage assessment scheme using remotely sensed data had to take account of the relatively high levels of historical landslide damage existing in some parts of the region. A block diagram of the approach taken to deliver the required information is shown in Figure 2. Such an approach requires the following information: data for assessment of pre- and post-Bola landslide extent, farm boundaries for delineation of the area of interest, a relationship between landform type and the ratio of landslide scar to debris-tail area, and a landform map of the affected region.

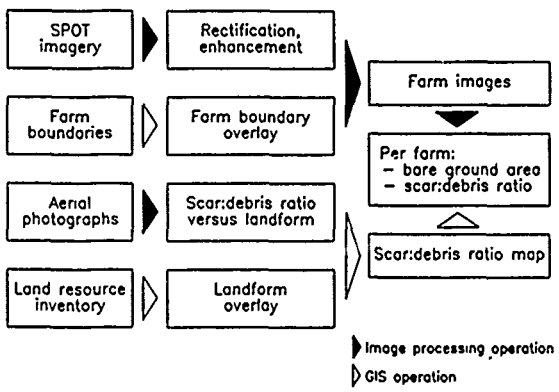


Figure 2. Block diagram of the approach taken to assess landslide damage using a combination of remotely sensed and geographic information system (GIS) data.

Digital imagery from the French SPOT satellite was chosen as being the most suitable data source for the assessment of landslide areal extent. Although storm damage assessment using remote sensing from aircraft was well established in New Zealand (Stephens *et al.*, 1981), use of SPOT data possessed a number of advantages over conventional aerial photography. Principal among these was the ability to provide rapid coverage of the relatively large area involved (about 12,000 km^2) - a factor made even more important by the restricted amount of time available for obtaining coverage of the area, prior to low winter sun angles making acquisition of useful imagery impossible. Other important features of SPOT included the on-board tape recorder (which compensated for New Zealand's lack of an earth resource satellite receiving station), and the across-track viewing capability of the sensors (which increased the probability of timely information being obtained). The potential of SPOT imagery for use in landslide damage assessment had been established in a previous investigation (Bellis *et al.*, 1987).

The coverage of SPOT scenes used for assessing landslide damage is depicted in Figure 1, with details of the individual scenes given in Table 1. Pre-Bola damage was assessed using multispectral SPOT scenes acquired in 1986, which although not covering all of the region, did include the area subject to the most serious landsliding in the decade prior to Bola. This was the area south-west of Gisborne, which suffered widespread landsliding in July 1985. While the 1986 SPOT scenes were not acquired for the purpose of landslide mapping, and were thus not at the higher 10 m spatial resolution, the larger landslides that occurred in the 1985 storm were discernible on this imagery.

Date	No. of scenes	Spatial resolution	Look angle, direction	Sun elevation
Pre-Bola:				
18 Mar 86	3	20 m	1.8°, W	42°
Post-Bola:				
26 Mar 88	2	10 m	17.4°, W	38°
26 Mar 88	1	10 m	21.1°, W	37°
5 Apr 88	1	10 m	7.2°, W	37°

Table 1. SPOT satellite scenes used in landslide damage assessment.

A number of vertical aerial photographic surveys were also used to assist in the interpretation and classification of the SPOT imagery. These included 1:25,000 scale panchromatic photography of some of the areas more severely affected by the July 1985 landsliding (acquired in December 1985). Also, the following aerial photographs were acquired during March and April 1988: 1:25,000 scale panchromatic coverage of areas west of Gisborne and Tolaga Bay, 1:30,000 scale natural colour coverage of all major roads and selected watersheds, and 1:15,000 scale colour infrared coverage of Tutira (see Figure 1 for location of these areas).

In assessing landslide damage, it was convenient to extract from the SPOT scenes those sections of the image that related to the individual farm properties - termed "farm images" in Figure 2. The farm boundary information necessary to allow this extraction was obtained from 1:50,000 scale cadastral maps. Most property boundaries were digitised directly from the maps and built into a farm boundary coverage, registered to the New Zealand Map Grid. The system used for this task consisted of a DEC microVAX-II computer running the Environmental Systems Research Institute geographic information system (GIS) software ARC/INFO, together with a Summagraphics digitiser and Tektronix display station. This system was also used to rasterise the polygon information in the farm

boundary coverage. The result of the vector-to-raster conversion was a farm-mask image, which could then be overlaid onto, and digitally intersected with, the (rectified) SPOT imagery to produce the individual farm images.

Earlier experience in using SPOT imagery had indicated that it was not possible to resolve the landslide scar and debris-tail components on the basis of radiance levels. However, it was known from both extensive mapping and other studies of erosion in the New Zealand landscape, that different landforms exhibit a response to storm damage that is primarily a function of the landform components lithology, slope-angle and vegetative cover (Trustrum and Hawley, 1985; Harmsworth *et al.*, 1987) Furthermore, for landsliding on pasture-covered slopes, it had long been observed that different lithology/slope-angle groupings appeared to exhibit landslides with different ratios of scar to debris-tail area.

Attempts to quantify any possible relationship between landform type and the ratio of landslide scar to debris area had not previously been made, due primarily to severe rainstorms failing to impact on a wide range of landforms. Cyclone Bola, however, delivered high intensity rain to all of the major landforms in the region. A study of aerial photographs of the region confirmed that a useful relationship between the ratio of the scar to debris-tail area and landform type could be derived for the Bola event. (The derivation of this relationship is discussed later.)

The major reason for seeking to establish a relationship between the ratio of scar to debris-tail area and landform type was that a landform coverage of the entire cyclone-affected region already existed. Intersection of the farm boundary coverage with this information allows the calculation of an average ratio for the scar to debris-tail area for individual farms. The landform coverage of the region, in both map and electronic form, is part of the 1:63,360 scale New Zealand Land Resource Inventory (Eyles, 1983). Although not ideal for farm-scale studies, the errors induced by using this somewhat small-scale information were not expected to be any greater than those encountered in the other components of the damage assessment approach.

CLASSIFICATION OF LANDSLIDES

Processing of SPOT images was undertaken using an ANTARES image display unit linked to a DEC microVAX-II computer, together with EPIC software (McDonnell *et al.*, 1984). SPOT images were edge and contrast enhanced, and then rectified to the New Zealand Map Grid using an unweighted least squares fit to a third-order polynomial mapping. Co-ordinates of ground control points were provided by reference to terrain and cultural features on 1.50,000 and 1:63,360 topographic maps.

Once rectified, the co-registered SPOT and farm-mask images were intersected to generate the digital images of individual farms. The intersection process consisted of constructing a new image according to the following simple algorithm: if a pixel in the farm-mask image had the required value, then the radiance level of the corresponding pixel in the co-registered SPOT image was written to the new image, otherwise the radiance level of the pixel in the new image was set to zero.

Classification of the farm images was not always a straightforward task. The low sun elevation at the time of acquisition of the 1988 SPOT imagery (see Table 1) caused such a wide range of hillslope irradiance levels that the radiance levels of vegetation and landslides (bare ground) were not distinct over the entire image. However, the radiance levels of landslides could be distinguished from vegetated ground on land of similar slope-angle and aspect. Even so, the typical degree of topographic variation, within the area covered by an individual farm image, meant that it was necessary to classify the image as a series of small areas of interest, essentially at the hillslope scale. A similar approach was also necessary when classifying the 1986 SPOT imagery, although the higher sun elevation at the time of image acquisition resulted in less overlap between landslide and pasture radiance levels, on all hillslopes.

Classification and analysis of the individual farm images was undertaken using Media Cybernetics IMAGE-PRO software operating on a 10 MHz, IBM compatible PC-AT, together with a MATROX MVP-AT image processing board. As the IMAGE-PRO/PC system required images of only 512 by 480 pixels, the farm images were first split into sub-images of this size - and written to PC compatible files - using the EPIC/microVAX system, and then transferred to the PC via a serial line. Compared to the microVAX-based system, the PC-based system allowed a greater degree of flexibility in the interactive classification of the farm images, a feature that was essential to the efficient hillslope-scale classification of these images. The splitting of the processing/rectification and classification/analysis tasks between two computer systems also had the advantage of increasing the overall throughput of farms assessed for damage.

A combination of interactive parallelepiped classification, and visual interpretation of the computer-displayed farm images, was used to determine landslide area. On well illuminated slopes, classification of landslides was achieved fairly readily. However, in broken terrain, the presence of poorly illuminated slopes made classification more difficult, and under these circumstances accurate classification required that analysts had good ground truth for the areas being classified - from either field observation or well-exposed aerial photographs.

An assessment of the error in the derived landslide area was made by comparing the landslide damage apparent on well-exposed - enlarged and contact - aerial photographs, with classified farm images. The classified farm images were visually checked on a landslide-by-landslide basis against the aerial photographs, and in some cases the aerial photographs were themselves also digitised and classified by an independent analyst. As a result of this process, the error in the landslide area was assigned a value of 20% when SPOT data was classified without reference to some form of ground truth, or 5% when ground truth was available. This 5% error component arose largely from variation in the ability of analysts both to precisely classify landslides, and to differentiate between landslides and highly reflective, vegetated ridge and spurs. There were 6 analysts involved in the damage assessment, and to reduce variation due to analyst interpretation, all classified farm images were checked by one other analyst.

In one situation the error limits determined above were not appropriate. This was in the case of farms which had suffered only a small amount of landsliding, at a large number of sites - with damage totalling perhaps just 1%-2% of the farm area. In such a circumstance,

the spatial resolution of the SPOT imagery became the limiting factor, with considerable over-classification tending to occur. For this reason, the minimum damage assessment error was set at 1% of the total farm area.

Although the errors involved in the quantitative assessment of landslide damage are appreciable, they are quite acceptable when judged against the figures for damage that appear in farmers' claims. A comparison between SPOT-derived and visually assessed landslide damage is presented in Figure 3. On average, the landslide damage claimed by landowners, per farm, was 3.7 times that detected on the SPOT imagery. Only one tenth of the claims were for damage that equated to that detected on the SPOT imagery - with no claims being for damage that was significantly less than the SPOT-derived value. This result emphasises the difficulties that claimants had in providing accurate visual assessments of damage.

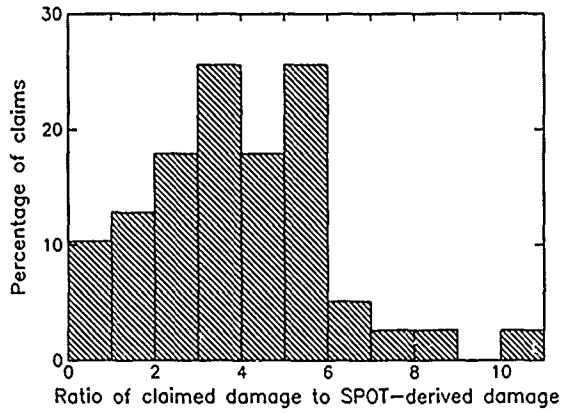


Figure 3. Comparison of landslide damage determined from analysis of SPOT digital data with damage as claimed by landholders.

LANDSLIDE SCAR AND DEBRIS-TAIL COMPONENTS

Classification of SPOT imagery gave a figure for landslide areal extent, on a farm-by-farm basis. However, as expected, radiance levels were such that the landslide scar and debris-tail components could not be separately classified. Instead, these components were estimated using an approach that relied on correlating the ratio of scar to debris-tail area, to the landform components lithology and slope-angle. An examination of aerial photographs taken immediately after Cyclone Bola indicated that such a correlation was likely to exist. The advantage of this approach was that if a reliable correlation could be found from case studies of damage to common landform units, then existing regional landform data could be used to extrapolate the results to the entire cyclone affected region.

The 3 variables thought to exert the dominating effect on debris-tail area for a given amount of scar material were lithology, the angle of the hillslope, and rainfall. Landslide sites were selected from the natural colour and colour-infrared aerial surveys mentioned earlier. Site lithology was determined from the New Zealand Land Resource Inventory, and categorised as belonging to 1 of 3 rainfall classes (<400 mm, 400-600 mm, >600 mm), and 1 of 2 hillslope-angle classes (21°-25°, 26°-35°). The landslide scar and debris-tail components were separated by photo-interpretation of stereo pairs, printed at a scale of 1:5,000. Transparent sheets laid over the photographs were manually shaded in different colours for the scar and debris-tails, digitised and classified to yield ratios of scar to debris-tail area.

An analysis of 1150 landslides was completed, resulting in the data presented in Table 2. The most surprising aspects of the analysis were the apparent lack of any systematic variation in the ratio of scar to debris-tail area in response to rainfall, and the failure of slope-angle to appear as a important factor on other than the jointed mudstone lithology. For example, in the >600 mm rainfall class, volcanic ashes showed ratios of scar to debris-tail area of 32:68 and 31:69 for the two slope-angle classes. At constant slope angle, jointed mudstones exhibited ratios of 20:80 and 22:78 in moving between rainfall classes of <400 mm and >600 mm. While the study was not exhaustive - the available time and photography preventing an analysis of a complete lithology/slope-angle/rainfall matrix - the results given in Table 2 appear to be accurate to within 15%. This error figure readily covers the variation seen in the ratios of scar to debris-tail area, for any given lithology/slope-angle grouping.

Lithology/slope-angle group	Ratio of Scar to debris-tail area
Siltstones	20:80
Jointed mudstones	20:80
Steep, jointed mudstones	25:75
Volcanic ashes	30:70
Sandy siltstones	35:65
Sandstones	40:60

Table 2. Derived relationship between lithology/slope-angle groupings and the ratio of the landslide scar to debris-tail area.

Having established the above relationship, the GIS was used to create a map of the region keyed to the ratio of scar to debris-tail area, using the landform information in the New Zealand Land Resource Inventory as base data. Farm boundary data were then digitally intersected with this map to provide an estimate of the average figure for the permanent and temporary pasture damage components, for particular farms.

CONCLUSIONS

Analysis of panchromatic SPOT imagery, of an area which suffered extensive cyclone damage, indicates that satellite data can be successfully used to provide a timely, quantitative assessment of landslide damage, at the individual farm scale. A comparison of the amount of landslide damage as determined from SPOT imagery, with visual estimates submitted by damage claimants, shows that most claimants over-estimated the amount of damage present. Of the 100 farms analysed, only 10% of claims were consistent with the SPOT-determined damage figure. On average, the landslide damage claimed was 3.7 times that detected on the SPOT imagery.

The principal difficulty encountered in the use of SPOT imagery for damage assessment was image mis-classification, resulting from the high degree of variation in radiance from landslides on steep slopes, exacerbated by the low sun angle at the time of image acquisition. Analysis of the extent of mis-classification places the landslide damage data as being accurate to 20%. The provision of damage assessment data at this level of accuracy can be seen as satisfactory when judged against the damage claims submitted by landowners. Given the discrepancy between actual and claimed damage, it is probable that more rapid damage assessment techniques can be developed which will considerably speed up the processing of damage claims while meeting a reduced, but still acceptable, accuracy.

Landslide scar and debris-tail components were not able to be separated in SPOT imagery, on the basis of radiance levels. These components were instead estimated from a derived relationship between the ratio of the landslide scar to debris-tail area, and landform. Analysis to date suggests that this relationship is a function only of landform lithology and slope-angle, and is insensitive to rainfall intensity. This is potentially a very useful relationship for use in storm damage assessments, given that landform and slope-angle information often already exist for affected regions. Further investigation is required to ensure that the relationship remains rainfall independent over all components of the lithology/slope-angle/rainfall matrix.

REFERENCES

Bellis, S.E., P.R. Stephens, I. McIndoe, K. Muscroft-Taylor, G. Ward, P.F. Newsome, 'SPOT-1 over New Zealand: Analysis of SPOT imagery for land use investigations", in: SPOT-1 image utilization, assessment, results. Paris, CNES, pp675-679, 1987.

Eyles, G.O., 'The distribution and severity of present soil erosion in New Zealand', New Zealand Geographer, Vol.39, No.1, pp12-28, 1983.

Harmsworth, G.R., G.D Hope, M.J Page, P.A Manson, "An assessment of storm damage at Otoi in Northern Hawkes Bay , Soil Conservation Centre, Aokautere, Publication No.10 (Division of Land and Soil Sciences, DSIR, Palmerston North, New Zealand), 76pp, 1987.

Lambert, M.G., N.A. Trustrum, D.A. Costall, Effect of soil slip erosion on seasonally dry Wairarapa hill pastures', New Zealand J. Agricultural Research, Vol.27, pp57-64, 1984.

McDonnell, M.J., D. Pairman, A.D.W. Fowler, 'Overview of PEL image processing capability", Int J. Remote Sensing, No.5, pp883-886, 1984.

Stephens, P.R., D.L. Hicks, N A Trustrum, "Aerial photographic techniques for soil conservation research", Photogrammetric Enginneering and Remote Sensing, Vol.47, No.1, pp79-87, 1981.

Stephens, P.R., N.A. Trustrum, J.R. Fletcher, "Reconnaissance mapping of erosion caused by Cyclone Namu, Solomon Islands", Asian-Pacific Remote Sensing Journal, Vol.1, No.1, pp57-64, 1988.

Trustrum, N.A., J.G. Hawley, 'Conversion of forest land use to grazing - a New Zealand perspective on the effects of landslide erosion on hill country productivity', FAO RAPA Report 1986/3, pp73-93, 1985

USING SATELLITE INFORMATION AS A TOOL FOR RANGE MANAGEMENT

Jeff Whiting and John Polson
Saskatchewan Research Council
Saskatoon, Saskatchewan, Canada S7N 2X8
and
Duane McCartney
Canada Agriculture
P.O. Box 1240
Melfort, Saskatchewan, Canada S0E 1A0

ABSTRACT

NOAA/AVHRR (Advanced Very High Resolution Radiometer) data has been used to assist in the yearly management of range and pasture lands in Saskatchewan. The paper looks at optimising carrying capacity, maximize sustained productivity, and identifies range readiness (when is the range ready for cattle in the spring and the season ended in the fall) on a province-wide basis.

A method has been developed for monitoring drought and grassland conditions using NOAA satellite data. NOAA imagery from weekly 1988 data and July or August of 1985, 1986 and 1987 were geometrically corrected and images produced for digital analysis. Cibachrome photographs were produced from the computer enhanced imagery. Actual forage crop yield data were used to produce a production map and compared to maps obtained from NOAA data. Results indicated that boundaries obtained from NOAA data compared favorably with those obtained from actual yield estimate.

Weekly composite satellite images were obtained in 1987 and 1988. The 1987 data was used as a comparison (as a per cent of 1987). Range readiness and conditions were provided for pasture managers and ranchers on a weekly basis. Training seminars were provided to interested range managers.

INTRODUCTION

Project Objective: To provide a means of using Satellite data to assist in the yearly management of range and pasture lands in Saskatchewan. The potential benefits of the project will include methods to:

(1) Optimize carrying capacity of Saskatchewan rangeland.
(2) Reduce damage caused by over grazing.
(3) Help maximize sustained productivity.
(4) Provide a method that pasture managers and ranchers can use to assist in managing pasture areas on a long term basis.

Background and Rationale: The grazing and hay lands of Saskatchewan, a Province in Western Canada, extend from the wide open short grass prairie of the south to the northern rough-land bush pastures of the Aspen Parkland (Populus tremuloides). These areas total 7,853,000 ha.

This study is part of a continuing effort directed towards the development of operational remote sensing analysis techniques for the estimation of grazing conditions in both native and seeded pastures in Saskatchewan.

Multi-spectral remotely sensed data from satellites has proven useful in assessing crop vigour. In most cases, where drought assessment or monitoring was required, large land areas were involved and high temporal resolution was preferred. AVHRR data from the NOAA series of satellites, with two passes per day over the critical areas, low spatial resolution at 1 km^2 for large area coverage and two channels appropriate for vegetation analysis was found to meet these needs.

Major species within the northern areas are bromegrass (Bromus inermis), intermediate wheatgrass (Agropyron intermedium), creeping red fescue (Festuca rubra), Kentucky bluegrass (Poa pratensis), alfalfa (Medicago sativa) and young aspen and balsam poplar (Populus balsamifera).

Native vegetation on Mixed Prairie zone is composed of mixtures of western wheatgrass (Agropyron smithii), northern wheatgrass (A. dasystachyum), blue grass (bouteloua gracilis), and sagebrush (Artemisia cana). Range improvement has included breaking and reseeding to wheatgrass, Russian wildrye (Elymus junceus), crested wheatgrass (Agroyron cristatum) and alfalfa (Medicago sativa).

Intera Ltd. assessed the possibility of using satellite data some 6 years ago in Alberta's short grass prairie. This company with a contract from D. McCartney, Canada Department of Agriculture (CDA), Melfort Research Station, conducted additional studies in an effort

to transfer this technology to the mixed grass area of Saskatchewan. Both projects used the weight-estimates-by-plot calibration method. The Alberta study was reasonably successful in Alberta and was adopted by private and provincial range managers.

Another recent study conducted at the Saskatchewan Research Council (SRC) developed techniques using NOAA data to determine and delineate drought in Saskatchewan.

The Saskatchewan Research Council (SRC) recently completed a study for the Melfort Research Station. A new method was developed to demonstrate the assessment of range condition, carry-over and overgrazing using satellite information. The photographs provided by the new method are superior to those presently available from either the Alberta Centre for Remote Sensing or the Canada Centre for Remote Sensing.In general these photographs provided an assessment of high and low carrying capacity but we still require a link between the photograph and what the rangeland manager needs. Results of the recent study for CDA also pointed out the need to treat at both the low and high end of the biomass range separately rather than a more general analysis in order to produce the most sensitive enhancements for individual pastures.

Guidelines for pasture management are required if a province wide methodology is to be provided to rangeland managers. SRC therefore proposes to continue to obtain detailed calibration at both Pathlow and McCranney Pastures in order to provide a method to address the problem of year-to-year variation. Existing data sets were collected from some ten percent of pasture in the province for coarse verification.

Project Description: The specific objectives of the proposal project are:

1) Obtain weekly satellite images of agriculture areas of Saskatchewan for the 1988 growing season.
2) Using previously developed techniques process the images to provide a weekly photo of forage growth in Saskatchewan.
3) Ground check satellite data.
4) Using satellite data from previous years ie. drought, wet and normal, construct a map which identifies areas of low, normal and high yield.
5) Provide and interpret information on range readiness, conditions and moisture levels for pasture managers and ranchers.
6) To provide an information seminar to interested range managers.
7) To provide the framework for additional research for the development of more detailed satellite information.

METHODOLOGY

Computer compatible tapes of NOAA AVHRR data were obtained for June to August of 1983 to 1987 and April 15 to October 15, 1988. The data was normalized by applying a quadratic correction curve to the data using the minimum radiance (clear water at nadir) and the brightest radiance values such as clouds. Normalized data from the visual and infrared channels were rationed to form a Normalized Difference Vegetation Indices (i.e. 350 x [Channel 2-1]/[Channel 2+1]) for specific region of the province. An agricultural district boundary map was overlain on the NOAA data.

Digital image analysis was performed on a DIPIX computer system. The daily Normalized Difference Vegetation Index (NDVI) was summarized by pixel image (representing 1 sq km) into a weekly composite in order to eliminate cloud cover. The maximum daily NDVI for each pixel was used in order to have the most recent data.

The Normalized Difference Vegetation Index for each week was then contrast stretched using Dipix Aries II computer software. The contrast stretch was performed on the bases of an equation. This equation calculated first an estimate of monthly rainfall and from that the amount of forage in kg ha^{-1} (Smoliak et al., 1976). The weekly 1988 data was divided by the maximum forage yield in 1987 at each point. Historical yield differences for specific areas between the normal and dry years for forage were used to derive ratios.

Range readiness was calculated by plotting the NDVI for each test pasture on a weekly basis. The threshold NDVI chosen at the time corresponding to the average date on which cattle were put on the range and when they were taken off.

The yearly data for 15 years (1965 - 1987) was compared to 1987 for calibration. The agricultural district boundaries and rural municipalities were digitized and used to extract an average NDVI value for each region.

RESULTS

Cibachrome colour prints of enhanced AVHRR imagery were produced from the computer enhancements to visually define the boundaries of the areas affected. Colour prints of enhanced imagery were provided with three classes, namely, dry, transitional and equal to or greater than 1987 for the chosen dates in 1988.

The 1985 imagery showed an extensive zone of drought extending from the southwest towards the northeast. During 1986 and 1987, these same drought areas supported mainly healthy crops and other vegetation. For forage the spread of the three years to the 15 year mean was larger with a greater spread at the higher end.

Analysis of the NOAA data compared to the ground data was done using ratios of 1985 to 1987 and 1986 to 1987.

For 1988, the NOAA analysis for forage showed that Saskatchewan averaged 57% less than normal (Table 1).

TABLE 1

AREA OF THE PROVINCE AT OR BELOW
LAST YEAR FOR FORAGE USING SATELLITE DATA

% OF 1987	AREA %	CUMULATED %	AREA (km^2)
0-39	--	--	--
40-49	18.9	18.9	74728
50-59	6.5	25.5	25906
60-69	7.5	33.0	29616
70-79	8.1	41.1	32250
80-89	8.4	49.6	33238
90-99	7.4	57.0	29527
>100	42.9	100.0	169331
Total	100.0		394605

DISCUSSION

Delineation of drought boundaries for 1988 using AVHRR data were obtained by both visual and digital analysis. Classification of the province into the three units showed the southwest to have been dry in 1983, 1984, 1985 and 1988 due to lack of sufficient soil moisture as indicated by low acreage of grain crops. Normally, much of the area will be dominated by native pasture and hayland which remained green only for a relatively short time in spring and early summer. By mid July, much of the pasture grasses were matured and beginning to dry. During 1985, this dry zone was much more extensive to the northeast and affected the forage crop growing area. In 1988, the forage areas did not green-up until the June 30 rains came and the temperature drop.

CONCLUSION

The study deals specifically with the use of computer assisted enhancement and analysis of NOAA AVHRR satellite data at 1 km resolution as tools for the development of hardcopy products to provide definition of biomass differences within grasslands in Saskatchewan.

Special emphasis is placed on the development of techniques which can be used to digitally enhance NOAA data for use in pasture and rangeland management. Drought and vegetation biomass maps are produced and early spring and late fall grazing conditions are monitored and related to nistorical data for specific areas in the province.

References

1. Smoliak, S., A. Johnson, M. Kicher, and R. Lodge, "Management of Prairie Rangeland", Canada Agriculture Publication #1589, Ottawa, 1976.

Acknowledgements

Canada and Manitoba Centers for Remote Sensing, Saskatchewan Crop Insurance, Saskatchewan Agricultural Development Fund New Pastures and Grazing Technology Projects.

An Assessment of TM Imagery for Land Cover Change Detection

Tung Fung

Department of Geography
The Chinese University of Hong Kong
Shatin, NT, Hong Kong

Abstract

The purpose of this paper is to assess the information content and accuracy of Landsat TM digital images in land cover change detection in the Kitchener-Waterloo area, Ontario, Canada. Change detection techniques of image differencing, principal components analysis and tasseled cap transformation were applied to yield 12 images. These images were thresholded according to the highest Kappa coefficients of agreement. The thresholded images were checked in terms of the producer's accuracies of specific land cover changes. The results show that the Kappa coefficients of the 12 images range from 32.2 - 63.2. Different images do detect different types of changes. Images associated with changes in the near infrared reflectance or greenness detect crop type changes and changes between vegetative and non-vegetative features. Images related to changes in the visible reflectance or brightness are able to detect changes due to rural to urban land conversion. The mid-infrared bands and change in wetness images do not provide additional information about land cover changes in this study area.

Keywords : TM, change detection, Kappa coefficient, producer's accuracy

Introduction

Change detection is a major application of remotely sensed data. With a repetitive acquisition of imagery, it is possible to determine the types and extent of changes in the environment. Many digital change detection techniques have been developed for this purpose. These include image overlay, image differencing, image ratioing, principal components analysis, spectral/temporal classification and post-classification comparison (Jensen, 1986). Techniques such as image differencing, image ratioing and principal components analysis involve transformations of the original spectral bands so as to enhance the land cover changes. Earlier studies applying image differencing with Landsat MSS data showed that the red band is useful to detect urban development (Toll et al., 1980) while the near infrared band may detect changes within agricultural land (Fung & LeDrew, 1987). Ingebritsen & Lyon (1985) indicated that principal components generated from MSS imagery normally yield two components accounting for change in brightness and change in greenness.

The TM imagery provides higher resolution data for remote sensing studies. Crist and Cicone (1984) indicated that at least three tasseled cap vectors, namely brightness, greenness and wetness, are inherent in the TM data. Wetness is especially related to the inclusion of TM5 and TM7, the mid-infrared bands. The ability of new spectral information for change detection is yet to be tested.

The objective of this paper is to assess the information content and accuracy of digital Landsat TM imagery for land cover change detection in the Kitchener-Waterloo area. Change related images are generated from techniques of image differencing, principal components analysis and tasseled cap transformation. A thresholding technique is used to differentiate change from no-change areas. The thresholded images are subjected to accuracy assessment. For each image, the producer's accuracies of different land cover change categories are computed to assess their ability to detect specific changes.

Study Site and Data Description

The study area consists of the twin-cities of Kitchener and Waterloo in southern Ontario, Canada. The total area is about 280 sq km. Two Landsat-5 TM images dated August 3, 1985 and July 21, 1986 of Path 18, Row 30 are available for the analysis. A subarea of 500 x 700 pixels was extracted. Both images were radiometrically calibrated and converted to reflectance values to alleviate any differences in solar elevation (Markham and Barker, 1986). To geometrically register the two images, the 1986 image was used as the master reference image. Twenty-nine ground control points were identified to generate polynomials for resampling based on a nearest neighbour resampling routine. The standard errors were below 0.2 pixels along the X- and Y-directions.

Major land cover changes are associated with the rapid urban development of the cities, especially along the western fringe of the cities. Despite a difference in only one year, a visual interpretation of the images depicts that significant urban development is found. There are other changes related to agricultural land cover changes in the surrounding rural landscape. Ten types of changes are identified:

(1) change from corn to grain;
(2) change from corn to pasture;
(3) change from corn to bare soil;
(4) change from pasture to grain;
(5) change from pasture to corn;
(6) change from bare soil to grain;
(7) change from bare soil to corn;
(8) change from bare soil to pasture;
(9) change from construction sites to residential land
(10) change from agricultural land to construction sites.

Categories (1) to (8) are changes in the agricultural land whilst (9) and (10) are related to urban development in this area.

A stratified systematic unaligned sampling method was used to generate sample sites for accuracy verification purposes. Each sample consists of a 3 x 3 block of pixels. These sample pixels were verified based on a field study performed in August 1986 as well as aerial photographs.

Methodology

Three change detection techniques were used to generate 12 change images. Image differencing was used to subtract the 1985 images from the 1986 images. Six differenced images (DIT1, DIT2, DIT3, DIT4, DIT5 and DIT7) were yielded for each of the TM spectral band (except TM6).

Fung and LeDrew (1987) suggested the use of standardized principal components for change detection. Thus, a principal components rotation was carried out based on the merged 12 TM bands. The third (PC3), fifth (PC5) and sixth (PC6) components are associated with changes. PC3 is related to change in brightness. PC5 represents change in near infrared reflectance mainly. PC6 shows changes in the contrast between the mid-infrared bands and the visible bands (Table 1).

To explore the inherent spectral data structure of TM imagery, a multidate tasseled cap rotation was performed. Tasseled cap coefficients derived by Crist and Cicone (1984) were used with spectral bands of different dates assigned to opposite signs. The vectors were then subjected to a Gram-Schmidt orthonormalization. Three tasseled cap change images were produced, namely change in brightness (ΔB), change in greenness (ΔG) and change in wetness (ΔW).

Each of the 12 change images was thresholded to delineate change from no-change areas. A method of generating optimal threshold levels for change detection suggested by Fung and LeDrew (1988) was adopted. An iterative process was done by means of selecting threshold values at $\pm N$ standard deviation from the means. In the first iteration, a N value of 0.1 was chosen. An increment of 0.1 was added in each subsequent iteration until the N value reached 2.0. For each iteration, an error matrix together with its Kappa coefficient of agreement and overall accuracy were computed. The threshold value yielding the highest Kappa coefficient was chosen as the optimal threshold levels for that particular image. Table 3 illustrates the levels at which the 12 images were thresholded together with the corresponding Kappa coefficients and overall accuracies.

The thresholded images basically contain binary information, change and no-change. To further assess the information content of each of them, the producer's accuracy for the 10 land cover change categories are produced (Table 4). Producer's accuracy is computed as the number of correctly identified samples of a particular category divided by the total number of reference samples for that category.

Results and Discussion

For the differenced TM images, the three visible bands (DIT1, DIT2 and DIT3) provide very similar information. They are effective in detecting changes due to rural to urban land conversion. The producer's accuracies of change from agricultural land to construction sites for these three images are over 85% (Table 4). Change from construction sites to residential land is over 80% accurate in these differenced visible images too. However, crop type changes cannot be accurately detected. With the highest Kappa coefficient among the three, it can be judged that DIT3, the red band, is more useful for change detection than the other spectral bands in the visible spectrum.

The DIT4 (near infrared) image detects changes from corn to bare soil, from pasture to grain, and from bare soil to pasture accurately (100%). However, it is not able to detect all the crop type changes, e.g. the accuracy of changes between pasture and corn is less than 80%. It cannot detect the change from bare soil to grain (19.75% accurate) as the near infrared reflectance of mature grain crops and bare soil are both low. Changes due to rural to urban land conversion are also not enhanced in this image.

The two differenced mid-infrared images (DIT5 and DIT7) do not provide any overwhelmingly additional information as compared to the visible and near infrared bands. While they can detect changes involved in rural to urban land conversion, they do not yield higher producer's accuracies for these changes as compared to the visible bands. For the detection of crop type changes in agricultural land, they are no better than DIT4. These two bands are originally designed to detect soil and plant moisture (Tucker, 1979). But in terms of land cover change detection in the study area, changes in the physical content of objects is perhaps more significant than changes in moisture content.

For the two multispectrally transformed images related to change in brightness, the Kappa coefficient of PC3 (53.9) is higher than that of ΔB (46.8). Both images

have producer's accuracies exceeding 80% for rural to urban land cover changes. PC3 has a higher producer's accuracy (95.51%) in detecting changes from agricultural land to construction sites. However, ΔB is the better in detecting changes form construction sites to residential land (91.67%). They can detect changes from corn to bare soil accurately as well (88.89-100%). The producer's accuracies for changes from corn to grain and from corn to pasture are over 90% for ΔB, but PC3 cannot detect these changes.

The tasseled cap change in greenness (ΔG) produces the highest Kappa coefficient (63.2) among the 12 images and is able to detect most of the crop type changes and vegetation to non-vegetation changes or vice versa. The producer's accuracies for changes from corn to bare soil and from bare soil to pasture are 100%. However, changes from bare soil to grain cannot be detected as the mature grain crops have a similar low greenness value as the bare soil. The producer's accuracy for change from bare soil to corn is not satisfactory (65.26%) as corn may not grow to full canopy in the July season in 1986. Similar to DIT4, it does not account for changes among non-vegetated features such as change from construction sites to residential land (44.44%).

PC5, an image related to change in near infrared reflectance, has a lower Kappa coefficient than ΔG and DIT4. Although the ability to detect specific changes is very similar to ΔG, the producer's accuracies of most categories are lower than those of ΔG.

For the tasseled cap change in wetness (ΔW) image, the producer's accuracies for changes from corn to bare soil, from pasture to grain, and from bare soil to pasture are 100%, same as those from ΔG. The producer accuracy of change from corn to grain is 85.19%. However, those of the other changes are mediocre, ranging from 29.63% to 70.79%.

PC6 is not effective in detecting most of the changes. It has the lowest Kappa coefficient among the 12 images as well. Only the change from construction sites to residential land has a producer's accuracy of 83.33%.

Conclusion

On the whole, although DIT5, DIT7 and ΔW represent new spectral dimensions in the TM data, their information on land cover changes is no better than those acquired from the change in brightness, change in visible or change in greenness images. Most of the land cover changes involve changes in brightness and greenness mainly and thus are captured by the first two tasseled cap variables. While wetness is sensitive to moisture changes, this is less significant in this study.

Not one of the images is able to detect all types of change. Rural to urban land conversion changes can be accurately detected by the change in brightness (ΔB) and change in red reflectance (DIT3) images. For the detection of crop type changes, the tasseled cap change in greenness (ΔG) provides the best information. The change in near infrared reflectance image (DIT4) can also detect crop type changes, but producer's accuracies of these change categories of DIT4 are lower than those of ΔG. Thus, even though there is a total of 12 bands in a multidate TM data set, it can be reduced to two vectors (DIT3 and DIT4, or ΔB and ΔG) for land cover change detection in the study area.

References

Crist, E.P. & R.C. Cicone, 1984, 'Comparisons of the Dimensionality and Features of Simulated Landsat-4 MSS and TM Data', Remote Sensing of Environment, 14, 235-246.

Fung, T. and E. LeDrew, 1987, "Application of Principal Components Analysis for Change Detection", Photogrammetric Engineering and Remote Sensing, 53(12), 1649-1658.

Fung, T. and E. LeDrew, 1988, "The Determination of Optimal Threshold Levels for Change Detection Using Various Accuracy Indices", Photogrammetric Engineering and Remote Sensing, 54(10), 1449-1454.

Ingebritsen, S.E. & R.J.P. Lyon, 1985, 'Principal Component Analysis of Multitemporal Image Pairs', International Journal of Remote Sensing, 6(5), 687-696.

Jensen, J.R.,1986, Introductory Digital Image Processing, Prentice-Hall, Englewood Cliffs, New Jersey.

Marker, B.L. and J.L. Baker, 1986, "Landsat MSS and TM Post-Calibration Dynamic Ranges, Exoatmospheric Reflectances and At-Satellite Temperatures", Landsat Technical Notes, EOSAT, August, 1986, No.1, 3-8.

Toll, D.L., J.A. Royal & J.B. Davis, 1980, 'Urban Area Update Procedures Using Landsat Data', Proceedings of American Society of Photogrammetry, RS-E-1-17.

Tucker, C.J.,1979, 'Red and Photographic Infrared Linear Combinations for Monitoring Vegetation', Remote Sensing of Environment, 8, 127-150.

Table 1 The Standardized Principal Components

	1985						1986					
Images	TM1	TM2	TM3	TM4	TM5	TM7	TM1	TM2	TM3	TM4	TM5	TM7
PC3	-0.20	-0.25	-0.33	0 26	-0.33	-0.35	0.29	0.30	0.32	-0.11	0 28	0.35
PC5	-0.03	-0.09	-0.02	-0.67	-0 23	-0.08	0.13	0.12	-0.03	0.65	0.18	-0.03
PC6	-0.35	-0.26	-0.17	-0.03	0.43	0.22	0.28	0.34	0.27	0.07	-0.39	-0.35

Table 2 The Tasseled Cap Variables

	1985						1986					
Images	TM1	TM2	TM3	TM4	TM5	TM7	TM1	TM2	TM3	TM4	TM5	TM7
ΔB	0.21	0.19	0.34	0.39	0.36	0.13	-0.21	-0.19	-0.34	-0.39	-0.36	-0.13
ΔG	-0.20	-0.17	-0.38	0.51	0.06	-0.13	0.20	0.17	0.38	-0.51	-0.06	0.13
ΔW	0.11	0.14	0.23	0 24	-0.50	-0.32	-0.11	-0.14	-0 23	-0.24	0.50	0.32

Table 3 The Optimal Threshold Levels, Kappa Coefficients and Overall Accuracy of the Images

	DIT1	DIT2	DIT3	DIT4	DIT5	DIT7	PC3	PC5	PC6	ΔB	ΔG	ΔW
Threshold Levels	1.4	1.0	0.9	1.1	1.1	0.7	0.8	0.9	0.8	0.8	0.8	0 8
Kappa (x 100)	33.7	42.7	57.5	52.5	46.9	47.7	53.9	47.8	32.2	46.8	63.2	51.9
Overall Accuracy	76.7	78.4	83.3	82.2	79.6	78.6	81.3	78.8	71.1	77.4	85.0	80.6

Table 4 The Producer's Accuracy of Different Land Cover Changes of the Thresholded Enhanced Images

	Land Cover Changes									
Images	C ↓ G	C ↓ P	C ↓ BS	P ↓ G	P ↓ C	BS ↓ G	L ↓ C	BS ↓ P	CC ↓ R	A ↓ CC
DIT1	0.00	0.00	100.00	22.22	0.00	2.47	25.26	44.44	80.56	88.76
DIT2	16.67	0.00	100.00	72.22	11.11	9.88	45.26	66.67	88.89	98.88
DIT3	48.15	0.00	100 00	94.44	11.11	39.51	65.26	100.00	83.33	97.75
DIT4	87.04	62.96	100.00	100.00	77.78	19.75	41.05	100.00	13.89	34.83
DIT5	33.33	66 67	100.00	50.00	40.74	28.40	32.63	86.11	83.33	76.40
DIT7	33.33	44.44	100.00	100.00	33.33	53.09	50.53	75.00	80.56	97.75
PC3	29.63	22.22	100.00	100.00	35.19	51.85	61.05	94.44	80.56	95.51
PC5	96.30	100.00	5.56	100.00	77.78	29 63	33.68	100.00	75 00	50.56
PC6	51.85	77.78	61.11	33.33	59.26	32.10	42.11	47.22	83.33	78 65
ΔB	94.44	96.30	88.89	61.11	51.85	35.80	37.89	72.22	91.67	87.64
ΔG	88 89	77.78	100.00	100.00	83.33	25.93	65.26	100.00	44.44	85.39
ΔW	85.19	51.85	100.00	100.00	55.56	29.63	51.58	100.00	44.44	70.79

Note: C = Corn
G = Grain
A = Agricultural Land
R = Residential Land
P = Pasture
BS = Bare Soil
CC = Construction Sites

UTILITYOF SPATIAL FILTERING TECHNIQUES FOR DETECTING SOIL EROSION ON LANDSAT THEMATIC MAPPER IMAGES OF IRAN

M. NAJAFI DISFANI

Department ot Geography and Topographic Science, University of Glasgow
Glasgow G12 8QQ, U.K.

ABSTRACT

Neither supervised nor unsupervised classification was successful for identifying gullies and gullied areas. Maps of dissected land in sample areas were obtained by applying the Laplacian convolution function of the IAX general purpose image processing software package to the enhanced Landsat Thematic Mapper digital image, band 3. Noise within vegetated areas was eliminated by using the vegetation indices image as a mask. To produce a map from the image of dissected lands, the image was smoothed and density slicing was applied to the smoothed image to categorise the dissected lands into five levels. Gullies, which are more important than badlands because they are dissecting the arable lands, can be brought out from Thematic Mapper digital data by use of the Laplacian convolution mask.

1.0 INTRODUCTION

Active gullies not only dissect the land and split it into different segments by headward erosion, but they also expand the gully width by pushing back the flanks of the gully into adjacent flat lands. The other problem of gully erosion is that it removes the soil down to alluvial fans, flood plains and finally reservoirs.

According to previous studies the following phenomena are sources of sediment: badland, gully, sheet, rill and bank erosion, different kinds of mass movement and finally accelerated erosion in arable lands. (Disfani, 1983; SOGREAH, 1973). In the SOGREAH reports (SOGREAH, 1973) badland areas have been considered as the major sources of sediment, but in fact we found both active and inactive gullies on aerial photographs of badlands. SOGREAH assumed the same spatial erosion rate for all sorts of gullies within completely different physiographic units. We reached the conclusion that parts of the gullied areas are old and inactive, appearing on the photos like series of individual domes, while other parts are still active, some of active gullies following and re-excavating the lines of old gullies. Therefore not only does the spatial rate of erosion differ between badland areas in different physiographic units, but it may also differ within a badland area according to the state of activity of the gullies.

The study area, the basin of the Sefid Rud reservoir, includes mountainous area in both the Zagros mountains in the west and the Alborz mountains in the north of Iran. The reservoir was constructed to generate hydroelectricity and supply water for irrigation, but it has gradually failed owing to the high rate of siltation. The reservoir basin with an area of 57,000 Km2 is situated in a semi arid region, having a continental climate, subjected to extreme seasonal variation of temperature and precipitation, as well as uneven spatial distribution.

The problem of siltation in the Sefid Rud reservoir and the loss of fertility of the land within a basin with about 3 millions population is the reason for the establishment of this research project which uses Thematic Mapper images and 1:50,000 and 1:20,000 black and white aerial photographs. The aim is to investigate the utility of Landsat Thematic Mapper digital data assisted by computer analysis to study soil erosion.

2.0 BASIC DATA AND EQUIPMENT

Compared with Landsat multispectral scanner (MSS) higher spatial and wavelength resolution of the Thematic Mapper (30m in 6 bands) improved the possibility of identifying gullied areaS. Multispectral SPOT (20m) and panchromatic SPOT (10m) images would have been even better if they had existed when the study began. In this study Landsat Thematic Mapper digital data from July 1985 was used together with 1:20,000 and 1:50,000 black and white aerial photographs, 1:50,000 and 1:250,000 topographic maps and the 1:250,000 geology map.

IAX, the general purpose image processor software package on the IBM 3090 mainframe computer of Glasgow University was used for image enhancement and filtering techniques.

Fig. 1 Effect of histogram equalization on Thematic Mapper band 3.

Fig. 2 Result of high pass filtering

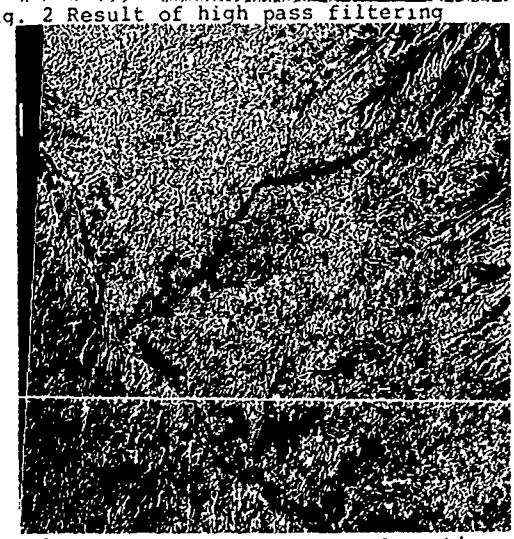


Fig. 3 result of edge and line detection

Fig. 4 Dissected land image derived from combination of VI and figure 3.

Fig. 5 Map of dissected land produced from TM

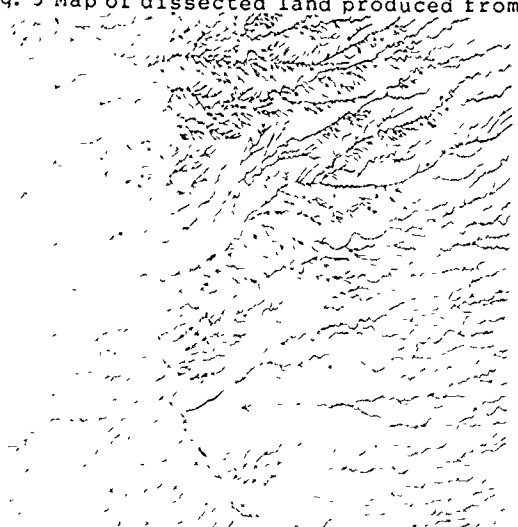


Fig. 6 Map of dissected land produced from aerial photographs.

IAX is an interactive image processing system for the professional end user. It can be used to process a wide variety of image data including binary, grey scale, and colour images. IAX is an acronym for Image Application eXecutor (IBM, 1987). In IAX, number, vector, image, and character string data formats are manipulated. It is a friendly algebraic image processor: data structures in these formats are held as variables and are processed by simple expressions. The DIAD-32 software package was used for classification purposes. DIAD (Digital Image Analysis and Display) runs under the MS/DOS operating system on an IBM-compatible personal computer.

3.0 DETECTING DISSECTED LAND BY SPATIAL FILTERING

Despite the presence of much literature on line and edge detecting techniques, no published work applying this to soil erosion on Landsat data could be found, so this is believed to be the first such attempt.

3.1 HISTOGRAM EQUALIZATION

The first stage which was carried out for band 3 of the Thematic Mapper image was histogram equalization (fig. 1). Histogram equalization is often used for increasing contrast between features for visual interpretation. Normally it is not recommended to apply this to the original image and then use it for further processing such as classification or change detection (Jensen, 1986), but here it is deliberately applied to the original band 3 to reduce the noise and unwanted edges and lines in the dark tail of the histogram, mainly vegetation, and the light tail, the non-eroded areas, and also to improve the visual appearance of edges and lines on the processed image.

3.2 HIGH PASS FILTERING

The next step is applying the high pass filtering to the image. In the conventional edge detection technique the first step is low pass filtering (Paine, 1987) . Applying the low pass filter to the original image removes the high and increases the low frequency data so that the image becomes gentle and smooth. In this instance, the result was that faint edges, evidence of the gullies, were removed and highly eroded areas detected as non eroded areas. Therefore the low pass filtering was rejected and replaced with high pass filtering. In the result the faint edges and lines were highlighted (fig. 2).

3.3 EDGE AND LINE DETECTION

The next step is detecting the edges and lines. There is no intention to use the detected edges as boundaries of polygons but they will be used as evidence of gullies. When using the edge and line detecting technique for detecting dissected lands one needs to take into account that a gully might appear as 3 edges if its width is more than one pixel or as one line if it is just one pixel or less than one pixel in width on the Thematic Mapper image. Therefore an algorithm should be chosen which has the ability to detect both edges and lines. The existing edge and line detecting filters such as the Sobel, the Robert, compass, the Laplacian convolution masks and the directional line detecting technique were evaluated. The Sobel and the Robert operators were found to be powerful edge detecting techniques, but the Laplacian convolution mask was found to be the best for detecting the badland and gullied areas because it has the ability to detect faint edges as well as coarse edges. Not only does it detect both edges and lines, but it also gives stronger weight to the lines than the edges. Only edges and lines in gullied areas were of interest for detecting the dissected lands, but all other artificial and natural lines and edges were also detected.

3.4 THRESHOLDING

The result of applying the Laplacian function on the Nikpay area appears on the screen as black, white and gray pixels. In this image the black pixels are non-eroded land, white pixels are eroded and gray pixels are transitional between eroded and non eroded. To change the transitional pixels to either eroded or non eroded and also for printing the image as hardcopy the thresholding function of IAX was applied to the edge detected image (fig.3).

3.5 MASKING

In order to mask out the noise within the vegetated areas caused by edges of plots of different crops the vegetation index (VI) was added to the detected image. If we simply add the VI to the edge detected image, problems will arise, because the non eroded areas on the edge detected image are black pixels and vegetated areas (non gullied) on the VI image are white so that we will have part of the non eroded area as black and part as white pixels. To solve the problem it was necessary that one image, either vegetation index or the edge detected image, should change to negative form. The best way was to change the edge detected image and add it to the vegetation index image. In the derived image (fig. 4) black pixels are evidence of gullies and the white pixels are non dissected lands. In this image it is possible to find out the relative proportion of dissected and non dissected land globally and / or within the regions of interest by applying the THIST (Type the HISTogram of an image) function of the IAX image processor.

3.6 SMOOTHING

Although it is possible to measure the relative proportion of dissected and non dissected land from fig. 4, and visually also dissected and non dissected land are interpretable, the dissected and non dissected land have not been categorised so far. To have the dissected and dissected land as a map with certain categories, the first

step is to smooth the image. To obtain the smooth image a low pass filter was used.

3.7 DENSITY SLICING

Two ways have been tested for producing the map of dissected lands from the smoothed image. In the first method one of the strongest edge detecting techniques, the Sobel operator was used on the smoothed image of dissected lands. In the result boundaries were detected and eroded and non eroded areas outlined.

In the second method for categorising the smoothed image, the density slicing function of IAX was used to split the dissected land into different levels of severity. Fig. 5 is the outcome of density slicing into 5 levels. We concluded that the second method gives a better result.

3.8 ASSESSING THE RESULT

To verify the accuracy and degree of correlation between image figure 5 which is produced from TM satellite data by means of filtering on the one hand, and figure 6 which is produced by means of aerial photograph interpretation on the other hand, they were superimposed on a light table. The result of comparison was satisfactory. The only disagreement between these two relates to areas containing remains of old gullies or evidence of mass movements. By applying this method, not only is it possible to detect dissected lands, but also it is possible to classify them according to the frequency of occurrence.

4.0 SUPERVISED AND UNSUPERVISED CLASSIFICATION

In order to compare the utility of the spatial filtering methods which were the main topic of the research with the conventional classification approach to feature extraction, attempts were also made to detect erosion by classification methods.

Both supervised and unsupervised classification methods were applied on the Thematic Mapper images of the study area. It was found that neither of them is successful for identifying badlands and gullied areas and their severity. The reason behind this is that the brightness values of badland areas is determined by the parent material, and similar brightness values from the same parent materials, not eroded, were found elsewhere in the image. The second reason is that badlands have formed on a sequence of different parent materials, including conglomerates and red beds so different parts of the badlands have different reflectances and therefore were classified in different classes. The third reason is that a wide gully may be classified as three separate classes representing the two flanks of a gully and the flat space between two gullies. The failure of classification methods for classifying badlands and their severity confirms that it is worthwhile to use the filtering techniques already described.

5.0 CONCLUSION

It was found in previous work that among erosion features gullies are recognizable on Thematic Mapper data. Detection of gullies and gullied areas by means of classification, whether supervised and unsupervised, was not successful in this study area. We came to the conclusion that the application of a Laplacian mask on the enhanced band 3 image could detect dissected lands. When aerial photographs and Thematic Mapper data are compared, the advantage of aerial photographs was that gullies actively cutting headwards were detectable, but on the Thematic Mapper data distiguishing between active and non active gullies was impossible. Aerial photographs are a very good tool to detect all kinds of erosion features (sheet, rill, and gully), but in my study area Thematic Mapper data can provide as much detail of soil erosion as is included in previous soil erosion maps made from aerial photographs. The Sobel and the Robert operators were found to be very strong edge detectors, but the ability of the Laplacian convolution mask for detecting faint edges and lines was greater.

6.0 ACKNOWLEDGMENT

I would like to thank my supervisor, Dr. A.Morrison for his constructive comments during the preparation of this paper.

REFERENCES

Disfani, Najafi M., 1983, Geomorphology and geodynamics of slopes in Taleghan valley. Persian text. M.Sc. Thesis, Beheshti (National) University of Iran, Tehran, Iran.

IBM, 1987, IAX image processing program, description and operation manual, document number SU59-7003-00, IBM UK Scientific Centre, Winchester.

Jensen, R. John, 1986, Introductory digital image processing. Prentice-Hall, Englewood Cliffs, New Jersey.

Paine, S. H. and Lodwick, G. D., 1987, "Edge detection and processing of remotely sensed digital images". Twenty- first International Symposium on Remote Sensing of Environment, PP 805-814, Ann Arbor, Michigan.

SOGREAH, 1973, Ingenieurs Conseils, project defence et restauration des sols du bassin versant du Sefid Rud.

THE STATISTICS CANADA CROP CONDITION ASSESSMENT PROGRAM

K.D. Korporal, R. Dobbins, N.M. Hillary

Remote Sensing Unit, Agriculture Division, Statistics Canada
12th floor, Jean Talon Building
Ottawa, Ontario
K1A 0T6

ABSTRACT

The Remote Sensing Unit of Statistics Canada's Agriculture Division has developed a new program to assess crop conditions in the Prairie Provinces. Preliminary response from the agricultural community, based on the prototype products produced during the crop growing season of 1988, revealed significant interest for such a program. For 1989, timely, quantitative and objective crop condition reports will be produced at the Crop Reporting District and Rural Municipality level on a weekly basis using low resolution NOAA satellite data. The products include maps, graphs, tables and visual images which can be integrated with historical crop condition information, conventional farm survey, administrative, census, soil and weather data. The crop condition reports are planned for weekly public release during the 1989 crop growing season.

Key Words: crop condition assessment, Crop Information System (CIS), remote sensing/geographic information system integration, NOAA AVHRR, Statistics Canada.

INTRODUCTION

Recent dry and drought years in the Canadian prairie provinces of Alberta, Saskatchewan and Manitoba have emphasized the need for new sources of timely, objective and quantitative information on crop conditions. In response to this information requirement, the Remote Sensing Unit of Statistics Canada's Agriculture Division has developed a cost effective remotely sensed based crop condition assessment program. Daily NOAA AVHRR images, composited weekly through the Crop Information System (CIS), provide the remotely sensed data necessary for the operational program.

The CIS data of approximately 1 km resolution are aggregated and averaged over a Crop Reporting District or Rural Municipality. This information is best suited to organizations and agencies dealing with similar aggregated units of analysis - it does not fulfill individual farmer needs for crop condition information. However, for time-series analysis and to highlight areas of vegetation stress, this information can help to focus resources and effort on specific areas requiring further study.

BACKGROUND

A prototype series of crop condition reports was developed in the summer of 1988 and introduced to selected organizations within the agricultural community. Response to the map and graph products was very positive; consequently, the crop condition reports are now planned for weekly public release, on a subscription basis, during the 1989 crop growing season. The products include maps, graphs, tables and visual images which can be integrated with historical crop condition information, conventional farm survey, administrative, census, soil and weather data. The reports provide timely, quantitative and objective information on crop conditions based on a Crop Reporting District level or on smaller geostatistical areas such as Rural Municipalities. This useful, relatively inexpensive analytical product is not intended to replace other sources of information which have served crop analysts; rather, it should complement these sources.

Weekly NOAA Composite Image

The Crop Information System (CIS) is a joint development of the Canada Centre for Remote Sensing, the Manitoba Remote Sensing Centre, the Canadian Wheat Board and others including Statistics Canada.

The Manitoba Remote Sensing Centre in Winnipeg, Manitoba produces weekly NOAA AVHRR CIS outputs on a subscription basis. The CIS involves extensive processing, including geometric correction, of daily NOAA imagery. Up to seven daily images are combined in a compositing process to produce a weekly composite image for two channels and a Normalized Difference Vegetation Index (NDVI) file. This compositing process significantly reduces cloud cover problems associated with the daily images. The

NDVI is a quantitative index emphasizing differences between stressed and unstressed vegetation and is calculated using two of the composited NOAA channels. This NDVI has been shown to be a good indicator of vegetation stress (Brown, et al., 1989; Hougham, 1987). A more thorough discussion of the CIS process and the NDVI is available in Brown, et al. (1989).

Image Processing

Each week, the NOAA composite imagery is transferred on tape or electronically to the remote sensing facility at Statistics Canada and is processed on the ARIES III image analysis system. Average NDVI values are calculated for the 40 Crop Reporting Districts and the almost 500 agricultural Rural Municipalities. An agricultural overlay mask is used in the calculations so that forest, urban and water areas are eliminated from the processing. Cloud-covered pixels are also excluded from the calculation of the mean NDVI statistics.

For internal Agriculture Division use, various enhanced visual images are produced based on the needs of the Crops Section analysts and statisticians, e.g., areas of extreme stress are highlighted for identifying additional sample locations. A colour image is also produced for the Crop Condition Assessment Program; the image data are classed and displayed showing the statistical difference between a current three-week running NDVI average and a seven-year (1982-1988) NDVI expected normal.

In addition to the visual images, the generated statistics are transferred to the ARC/INFO geographic information system (GIS) in a procedure explained in Korporal, et al. (1989). Extensive integration of data capture, display, manipulation and analysis techniques, mostly in the GIS environment, is used to enhance the spatial characteristics of these data.

Crop Condition Mapping

The average NDVI statistics transferred from the ARIES system are processed and an INFO database file is created. Seven years of historical NDVI averages are also accessible in INFO. Each week, difference and percentage change statistics are calculated against the same Julian week in the previous year and the same Julian week of the seven-year normal.

The resulting maps display time-series comparisons by Crop Reporting District for both recent and historical crop condition information (Figure 1). More detailed maps, based on smaller geostatistical areas such as Rural Municipalities, are also available in order to permit the user to better focus subsequent analysis (Figure 2). Maps could also be produced for user-defined geostatistical areas.

Precipitation And Temperature Mapping

Daily and weekly summaries of temperature and precipitation data are captured by an on-line PC microcomputer from a service available from the Atmospheric Environment Service, Environment Canada, in Winnipeg, Manitoba. These data are processed by an automated editing program and transformed to an INFO format data file.

Two types of meteorological map products are produced. The first averages weekly percentage of normal precipitation since April 1 by Crop Reporting District. The second map type plots individual variables such as the percent of normal precipitation, total precipitation and the temperature from normal at the location of each of 145 Prairie meteorological stations. A reference map locating all meteorological stations is also provided.

Graphic Display and Tabular Data

For quantitative analysis, it is desirable to have both graphic and tabular representation of the NDVI information. Unfortunately, ARC/INFO does not have a graphic output capability. It is therefore necessary to re-process the NDVI data and transfer a copy of all files to the PC LOTUS 1-2-3 or SAS/GRAPH environment. These data are then displayed by plotting the mean Crop Reporting District NDVI values against time in weeks. Crop Reporting Districts in areas of similar geography (soils, climate, crops, etc.) can be plotted together to check the integrity of the data as the crop growing season progresses. More interestingly, on an individual Crop Reporting District basis, previous years and the seven-year normal can be compared graphically to the current year, also by week (Figure 3). Weekly tabular data summaries also are produced using the report functions in INFO, LOTUS and SAS.

FUTURE DIRECTIONS

The immediate research and development goal is to produce early season yield estimates. A cooperative research program is underway with CCRS and other CIS subscribers. the preliminary research results are very promising. It is hoped that early yield estimates will be an operational component of the Crop Condition Assessment Program for the 1991 crop season.

Concurrent research focusses on examining possible correlations between the NDVI and precipitation. Combined with soil moisture content and soil type, this information also may be useful for crop condition and yield predictions.

The Crop Condition Assessment Program, now in its first operational year, will continue to evolve and expand. In addition to the processing related to the subscriptions, the program will be further integrated with traditional crop surveys in order to provide spatial analysis and map output capabilities. Standardized map products are being considered for Crop Reporting Districts and other geostatistical units in provinces other than the Prairies. Other developments will focus on better user interfaces using the ARC/INFO macro language and menu capability.

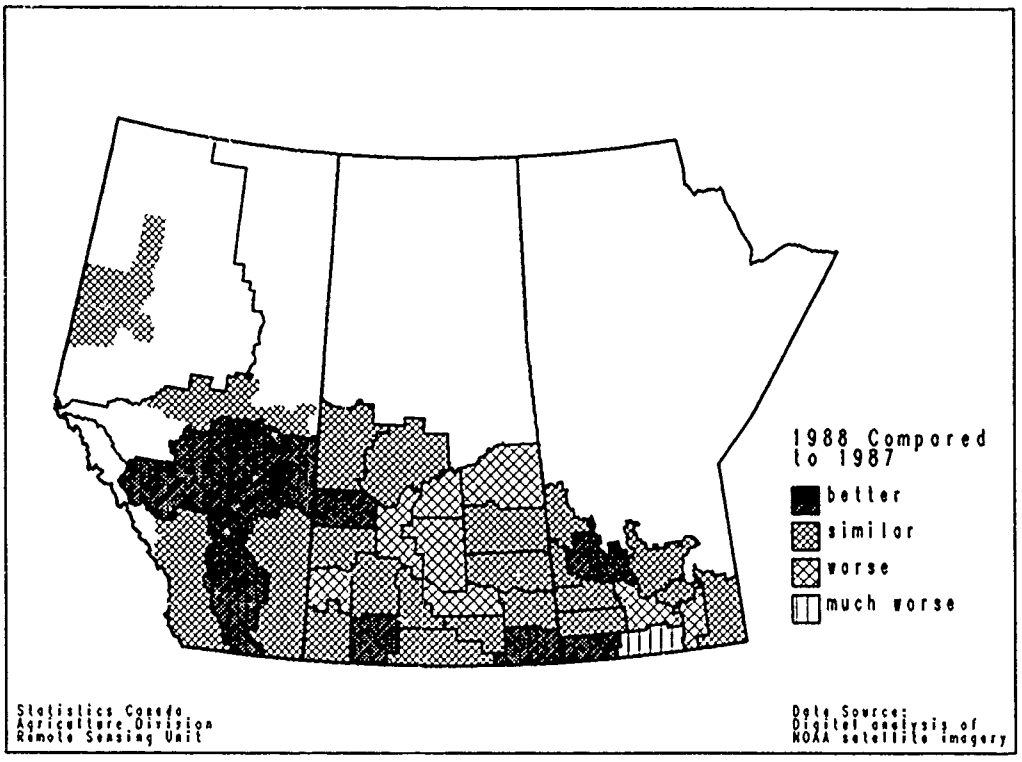


Figure 1 : Comparison of Crop Conditions, Week of July 11-17 (Julian Week #28), Prairie Provinces by Crop Reporting District

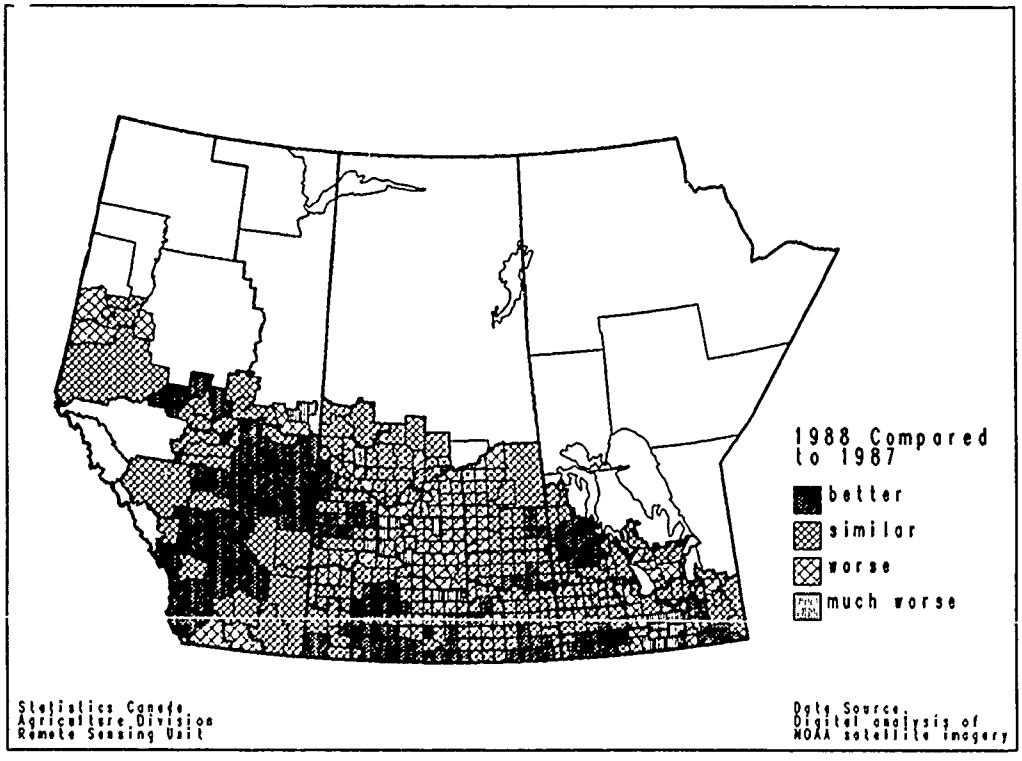


Figure 2 : Comparison of Crop Conditions, Week of July 11-17 (Julian Week #28), Prairie Provinces by Rural Municipality

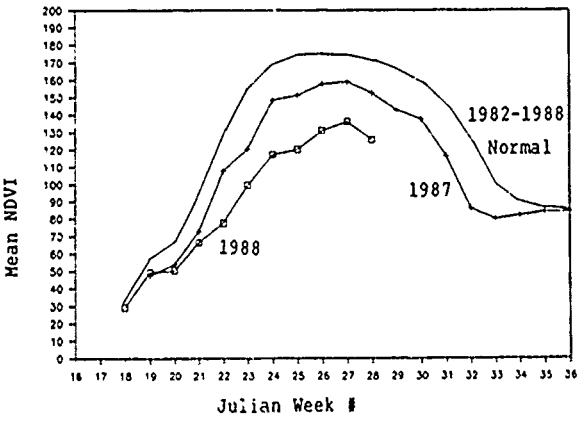


Figure 3 : Normalized Difference Vegetation Index, Manitoba Crop Reporting District 7, 1987, 1988 and 1982-1988 Normal*

* 1982-1988 Normal not yet calculated; approximate curve is graphically presented here.

CONCLUSION

Objective, timely, quantitative and accurate information on crop conditions is of increasing importance to the Canadian agricultural community. The Crop Condition Assessment Program provides the infrastructure and spatial data display to provide this information. The program will grow in importance as geostatistical analysis and digital map products become integral tools for agricultural analysts and statisticians.

ACKNOWLEDGEMENTS

The authors wish to acknowledge the Canada Centre for Remote Sensing, especially R. Brown, M. Manore, H. Press, D. Mullins and J. Gairns, for their cooperation in the research and development of the Crop Condition Assessment Program.

REFERENCES

Brown, R.J., M.J. Manore, K.D. Korporal and H.E. Press. 1989. GIS in the Analysis of Satellite Data for Vegetation Monitoring: The Crop Information System, *Challenge for the 1990s - GIS National Conference*, Ottawa, February 27 - March 3.

Code, G.O. 1986. The Use of Remote Sensing Techniques by Statistics Canada in the Statistical Assessment of Land Use Data, *Statistical Assessment of Land Use: The Impact of Remote Sensing and Other Recent Developments on Methodology*, Theme 3, Series E, pp 189-200, Luxembourg, September 29 - October 1.

Dobbins, R. and H. Epp. 1987. *Crop and Summerfallow Area Estimation in Saskatchewan Using Landsat MSS Data for 1985*, Saskatchewan Research Council Technical Report No. 196, Saskatoon, April.

Hougham, A.M. 1987. *Use of NOAA AVHRR Digital Satellite Data for Precipitation and Forest Fire Assessment*, Saskatchewan Research Council Technical Report No. 198, Saskatoon.

Korporal, K.D., R. Dobbins, N.M. Hillary. 1989. The Statistics Canada Geographic Crop Monitoring System, *Challenge for the 1990s - GIS National Conference*, Ottawa, February 27 March 3.

Korporal, K.D. and M.J. Manore. 1986. An Application for the Testing and Use of the Standard Data Transfer Format, *Proceedings of the Tenth Canadian Symposium on Remote Sensing*, pp. 867-873, Edmonton, May 5-8.

Korporal, K.D. 1985. *The Delineation of Census Urban Areas by Computer-Assisted Mapping and Remote Sensing*, Unpublished Master's Thesis, Department of Geography, Carleton University, Ottawa, 117 pp.

Korporal, K.D. 1983. A Demonstration Transfer of Remotely Sensed Data Utilizing the Standard Format for the Transfer of Geocoded Polygon Data, *Proceedings of the Sixth International Symposium on Automated Cartography*, Ottawa, pp. 607-615; and in *Cartographica*, Vol. 21, Nos. 2 & 3, pp. 53-61.

Prout, N.A., J. Sutton, J. Wessels, M. Manore and R.J. Brown. 1986. Technical Feasibility to Mobilization For Operations: The NOAA Crop Monitoring Case, *Proceedings of the Tenth Canadian Symposium on Remote Sensing*, pp. 787-796, Edmonton, May 5-8.

Prout, N.A., G.J. Wessels, G.K. Walker and R.J. Brown. 1986. Routine Production of Crop and Drought Monitoring Products from NOAA-AVHRR Data, presented at the *20th International Symposium on Remote Sensing of the Environment*, pp. 1221-1230, Nairobi, Kenya, 4-10 December.

Ryerson, R., J. Harvie, J-L Tambay and R. Plourde. 1983. The Use of Landsat, Ground Data and a Regression Estimator for Potato Area Estimation, *Users' Manual 83-2*, 42 pp, Canada Centre for Remote Sensing, Ottawa.

OPERATIONAL GIS: PAST, PRESENT AND FUTURE

Frank Hegyi

Inventory Branch, Ministry of Forests

1319 Government St, Victoria, B.C. V8W 3E7

Canada

ABSTRACT

During the past 10 years, Geographic Information Systems (GIS) have progressed from a research and development state to that of operational implementation. Parallel with this development, capabilities of GIS systems advanced through various stages, such as cartographic applications, computer aided designs and mapping, polygon-based overlays and data management, relational data base structures, rasterized overlays and data management, and integration with satellite image analysis systems. At the same time, users of systems or clients of GIS vendors increased their requirement from technical capabilities to cost-effective operational performance specifications.

In this paper the author highlights some of the main problems, challenges and achievements of GIS applications from the development stage to the current operational implementation phase, as well as projects future trends and opportunities. The vision of the future includes expansions to resource management, land use evaluation applications and global monitoring, resulting in significant impacts on current management practices and economic policies.

SPECIFICATIONS FOR GIS AND IMAGE ANALYSIS DATA: A GOVERNMENT PERSPECTIVE

Peter Friesen and Mark Sondheim

Surveys and Resource Mapping Branch
British Columbia Ministry of Crown Lands
533 Superior Street
Victoria, British Columbia V8V 1X5
Tel: (604) 387-9355, 387-9352 FAX: (604) 387-3022

Abstract

GIS is in the process of changing from a specialized development accessible to a select few, to a common business and technical tool widespread throughout society. As well, image analysis now holds the promise of providing a large volume of data applicable within the GIS environment for a variety of applications.

In order to facilitate ease of use and integration of georeferenced data, the government of British Columbia has taken on the responsibility of developing specifications for transferral and archiving of such data.

This paper places the issue of specifications in a business context and then provides evaluation criteria against which different formats are judged. A proposal for British Columbia follows, dealing with the whole gamut of georeferenced data, including various vector and raster representations. The paper concludes that existing formats are not the most suitable for British Columbia and that the option of migrating to a format such as the one proposed is justifiable.

I Scope and Objectives

British Columbia consists of roughly 95 million hectares, 90% of which is owned by the provincial or federal government, with management delegated through the Canadian constitution primarily to provincial authorities. Within this land mass - larger than Washington, Oregon and California combined - reside less than 2½ million people. The trend in B.C. has been to maintain a civil service small in terms of both absolute numbers and as a proportion of the total population. Consequently, the ability of comparatively few people in the B.C. government to manage very large areas is seen as a critical issue.

In order to gain efficiency, both within and across specific provincial agencies, there has been a broad thrust beginning several years ago in GIS. This push has concentrated primarily on data collection and operational use for specific applications, the more sophisticated work has tended to be project specific and handled by in-house specialists. More recently it has been recognized that GIS holds the potential to benefit a full spectrum of users dealing with many day-to-day business operations associated with: water management, waste management, wildlife and fisheries management, forestry and geological applications, cadastral administration and other activities. These business operations include both typical administrative functions as well as traditional analytical functions normally carried out with GIS systems.

This thinking leads to two major issues. (1) provision of standardized databases spanning the province, and (2) development of a GIS infrastructure throughout the government, applicable on a variety of hardware platforms, across network, and in regional as well as central (Victoria) offices. For both issues it is clearly imperative that specifications be put into place which are applicable to any data which may be used in GIS work. Note that digital data from imagery is considered as simply another data source for GIS, whether in an enhanced pictorial form, or as a classified product.

The government of British Columbia has taken on the responsibility of defining a vendor independent non-proprietary format to allow users to freely exchange data between existing and future GIS (and image analysis) systems. This format should also be suitable for the archival of large spatial databases in a compact form to facilitate the retrieval and distribution of the data as well as to minimize the impact of system upgrades or replacements.

II Evaluation Criteria

In order to assess the adequacy of existing or proposed formats to meet our needs in British Columbia, a list of evaluation criteria was first devised. Twelve criteria have been defined as follows.

a. - **Data Volume**
Despite the decrease in cost and increase in efficiency of digital storage systems, data volume still remains a concern of any database. Spatial databases are typically large and there is an increasing demand for the import and export of this data in the microcompute. environment. Ideally, a suitable transfer and archival format should have the ability to adapt to the amount of ancillary (i.e. file description) information required. For example, headers need not carry redundant and unused fields.

b. - **Hardware Platform**
Spatial data will in many cases be transferred between a variety of minicomputer and microcomputer environments and so the format should be supportable on a wide range of platforms under different operating systems.

c. - **Media**
The format should be suitable for transfer on a number of media including magnetic tape, floppy disks, CDROM, and computer networks.

d. - **Archival Suitability**
The format should be suitable for archival. This means that the spatial data for a region or project area must be self-contained such that source, producer, updating, completeness, accuracy, and referencing information is retained as part of the data set. It also must be in a compact form.

e. - **Integration with Traditional Relational Structures**
The format must have the ability to relate individual spatial objects through the use of a primary key. This would allow the maintenance of a link between the spatial attributes and a relationally structured set of business attributes.

f. - **Appropriate for Use in Production and Research Environments**
The format must be suitable in a production environment as well as a research environment. There must be sufficient detail to allow for large project areas and varied themes to be encoded. However, it should also be realistic to clip subsets for research purposes, without the encumbrance of unwieldy data structures.

g. - **Appropriate for All Spatial Entities**
The format must have the capability to encode all spatial entities. These entities include such elements as points, lines, areas, labels, text, and symbols.

h. - **Ability to Encode a Variety of Spatial Data Types**
The format must be able to encode, archive and transfer both vector based and raster based data.

i. - **Ability to Encode both 2D and 3D Spatial References**
The format must be able to carry a height *(z)* value as part of a *(x, y, z)* coordinate triplet for each spatial reference.

j. - **Apparent Seamlessness**
Although the format can be file based, it should have the ability to give the appearance of seamlessness. This means that there should be the capability for the selection of a window of data of any size or shape. Ideally, a large watershed spanning several mapsheets and UTM Zones should be able to be handled.

k. - **Ability to Support Projections and Datums Used in B.C.**
The format must have the ability to encode spatial references in the projections and datums most commonly used in mapping in the province of British Columbia. This requires a method of storing polyconic, geographic and UTM coordinates using the NAD27 or NAD83 datums.

l. - **Ease of Migration from Formats Currently Used within B.C.**
If possible to accomplish without compromise, the specifications should be such that migration to them is reasonable both technically and organizationally.

III Spatial Format Alternatives

Seven published formats were examined including ISIF, TIGER, CCSM, DLG, MACDIF, MOEP, and SDTS. Some of these are currently used operationally, whereas others are undergoing testing. These formats are briefly reviewed below.

ISIF (Intergraph Standard Interchange Format)
This format was developed by Intergraph to act as a common mechanism for transmittal of graphics and associated data between systems. ISIF can be created in either an ASCII or binary format with no header information, resulting in reasonable data volumes for a single file. It can be used on virtually any hardware platform and is simple enough that with a basic understanding of Intergraph, software creation should not be a problem. Any kind of standard media can be accommodated, as these are simple ASCII and binary files. The format cannot be integrated with traditional relational systems; features are not uniquely identified to allow the use of primary keys or indexing. It is strictly file based with 4 byte integer Cartesian coordinates based on an origin that is peculiar to the system on which it was created. ISIF is more appropriate for transfer rather than archiving, due to its Intergraph dependence and lack of historical header data. It is however useful for the translation of all spatial element types, although topology is not taken into account. It is suitable only for vector data, with no support for rasters.

TIGER (Topologically Integrated Geographic Encoding and Referencing)
TIGER was developed by the U.S. Department of Commerce as a digital map data base that automates the mapping and related activities required to support the census and survey programs of the Census Bureau. It is difficult to estimate the data volumes for this format as it is specific to census data. It is suitable for use on any hardware platform and media. It was designed as both an archive and distribution format. It can be integrated with relational database structures due to its reliance on table lookups and primary keyed structure. It will only code linear spatial elements; thus, it is not suitable for area or cartographic entities. It is a truly seamless database using geographics as its spatial referencing system. It does not support raster or surface structures, only lines with census related topology. TIGER was developed with very modern concepts; however, it is oriented strictly toward use of census data.

CCSM (Canadian Council on Surveying and Mapping)
CCSM was developed by the Canadian Ministry of Energy, Mines and Resources to provide a machine and language independent national standard for the exchange of digital topographic data. This format was designed as a magnetic tape based transfer format. All data is coded in ASCII with extensive header and

related attribute information. This typically results in very large data files. The software available for translation to and from different GIS systems is VAX based. Because it is a sequential media based format, it is not suitable for incorporation with traditional relational structures. The format was designed exclusively as a production transfer format and is less suitable for research applications. CCSM is appropriate for all spatial entity types and can carry some topological characteristics; however, it does not support rasters or grid based surfaces. As it is a file based, sequential storage system, extracting an arbitrary window of data can be awkward. CCSM includes a flexible and well documented mechanism for classifying feature categories using a ten character-number combination.

DLG (Digital Line Graph)

DLG was originally developed by the U.S. Geological Survey as part of the U.S. National Mapping Programs' Digital Cartographic Data Standards. The enhancement to DLG that we have been reviewing was created by Universal Systems Ltd. for use as a transfer format to and from their CARIS system. The USL enhancements include increased topology, 3D capability, and a separate attribute file structure. Data volume is not a problem with this format, in a binary format with limited header information, files are typically very lean. It is suitable for any hardware platform or media storage, although it may not be appropriate as an archival format due to the limited header information. Without enhancement there is no way of identifying object structures, and so linkage with traditional relational structures is not possible. It would however be suitable for production and research environments, as a means of transferring spatial attributes. It would also be difficult to obtain seamlessness in this format due to its file based nature. A vector based format, DLG does not support raster structures.

MACDIF (Map and Chart Data Interchange Format)

MACDIF was developed primarily by the Ontario Ministry of Natural Resources and the Canadian Hydrographic Service as a comprehensive telecommunications oriented, interchange standard, for both topographic map and hydrographic chart data. This format is the most complete to be reviewed so far. It has extensive header information and spatial entity capability. This makes it suitable for all applications of spatial data with the exception of raster data. The data volumes are large because of its extensive header information; however, carrying this information makes it suitable as an archival as well as transfer format. It is not restricted by any hardware platform or media. As intended, it is particularly suited to transmission over telecommunication systems. MACDIF is seamless in concept using a coordinate system that unifies all input projections and datums. It is file based but has the capability of uniquely identifying spatial objects; these objects could in turn be linked to a relational structure. It also has the ability to handle all types of spatial and cartographic elements, although rasters are not addressed.

MOEP (Ministry of Environment and Parks)

Based on an enhancement of a format developed by the Province of Alberta, this format is intended primarily as a means of transferring 1:20 000 digital topographic and cadastral map data generated by the Terrain Resource Information Management (TRIM) program. It has been well received in British Columbia for this purpose. It has very limited header information and with binary compression its data volumes are very small relative to most proprietary GIS formats. It is not restricted by hardware platforms or media. It will handle all types of spatial data with the exception of raster. MOEP is well suited to research environments but its limited scale and projection capabilities make it unsuitable in its present form to varied production environments. This format does carry the ability to introduce unique object identifiers which in turn could link to a relational structure. It is file based with only the UTM projection supported at this time. This makes the concept of seamlessness not suited to this format. The simplicity and compactness of the format makes MOEP a suitable candidate for extensive enhancement. Various GIS vendors operating within British Columbia now support MOEP.

SDTS (Spatial Data Transfer Specification)

Developed in the U.S. by three related groups: (1) the National Committee for Digital Cartographic Data Standards (NCDCDS), (2) the Standards Working Group of the Federal Interagency Coordinating Committee on Digital Cartography (FICCDC-SWG), and (3) the Digital Cartographic Data Standards Task Force (DCDSTF). This format specification covers vector and raster data, and supports a variety of zero, one, and two dimensional elements. It is the most complete spatial data format of those reviewed. SDTS breaks down the essential high level components of a spatial format into a set of modules that can be viewed in a semi-relational fashion. These modules contain such things as: (1) global information containing attributes describing cataloguing, identification, security, spatial referencing, spatial domains, data dictionary, and transfer statistics, and (2) data quality describing lineage, positional and attribute accuracy, logical consistency and completeness. The spatial data itself can be described in three ways; these are referenced as the vector form, the relational form, and the raster form. This format gives a blueprint for the archival and transfer of spatial data, not restricting it to any hardware platform, media, or recognized digital standard. Its semi-relational table structure makes it ideally suited to both the research and production environments. SDTS gives the plans for encoding of any spatial data type or structure.

IV A Proposed Format for British Columbia

BC-SAIF (British Columbia Spatial Archive and Interchange Format)

A proposed format for the transfer and archival of digital spatial data for use in the province of British Columbia would use concepts and extensions of the formats reviewed above, particularly SDTS. It would make extensive use of the relationally structured high level information as outlined by the SDTS format. The high level information for a specific project or map area would be maintained as a set of separate files.

The vector format would be an object oriented extension of the MOEP format with implicit topology and the choice of the two projections (UTM and polyconic) relevant to the area covered by the province. The individual spatial component feature coding would use the principles as developed by the CCSM, but the unique object identifiers would be generated through a combination of primary information relevant to the theme and the spatial area of the object itself. This would keep the vector data in a sequential file structure separate from the header information. This data would be in a simple form that is easily understood and converted while maintaining a seamless unique identifier as a link to a relational attribute structure. The raster data sets would use the format as outlined in the SDTS. An accompanying header would contain all relevant information to build the raster image from the spatial component file. The spatial raster file itself would have the option of either a simple sequential binary list of raster values or a run encoded form. There also would be an optional accompanying cell value file to contain attribute information for each pixel.

BC-SAIF would satisfy the evaluation criteria listed above. The set of relational files for a particular area of interest contain all information needed for that area including lineage, positional accuracy, attribute linkages, cataloguing information, and spatial referencing. The files can be archived and interchanged as a complete set, or any subset of data such as just the spatial references or just the header information can be extracted for research purposes. The multi-file structure of the format also allows for the mixing of file types; all header information can be held in ASCII while all spatial reference data is in a compressed binary form. Note that BC-SAIF can be considered as a variant of SDTS.

V Discussion

To satisfy our twelve evaluation criteria a blend of the above specifications utilizing the qualities best suited to the interchange and archival requirements in British Columbia can be developed. BC-SAIF, a merging of concepts under the guidelines specified in the STDS, best suits the purpose of the format for the following reasons:

(1) The extension of MOEP for the storage of vector references facilitates the migration of existing map products.

(2) The use of CCSM structured feature coding expedites the migration of federal mapping products. The coding structure also provides a method for uniquely categorizing features to allow the grouping of like entities.

(3) The use of STDS concepts gives companies with a presence in B.C. the opportunity to become familiar with the developing U.S. format, while at the same time meeting the immediate needs of many B.C. customers.

(4) The resulting multi-file structure allows ready integration with relational databases, as these files fit the Binary Large Object (BLOB) concept for which there is ongoing development by relational database vendors. Each spatial object is a variable length BLOB with a unique identifier serving as the foreign key linking the object to a set of business attributes.

(5) The SDTS concept allows the existence of both vector and raster spatial data sets under the same specification. Both spatial types can utilize the same format header files.

It is our intention to develop BC-SAIF for use within a GIS prototype being established by our Branch. We will be working with vendors and other government agencies during this process, as appropriate. The long term objective is to ensure that, through a common specification, data archive and interchange needs throughout the province can be achieved.

VI References

The American Cartographer: The Proposed Standard for Digital Cartographic Data, Volume 15, No.1, American Congress on Surveying and Mapping, January 1988.

Specification of the Map and Chart Data Interchange Format: MACDIF, Version 2, Canadian Hydrographic Service, Department of Fisheries and Oceans, Government of Canada, March 1988

TIGER File Extracts, U.S.Department of Commerce, Bureau of the Census, Washington, D.C., June 1 1988.

National Standards for the Exchange of Digital Topographic Data, Volume I/Part III - Standard EDP File Exchange Format for Digital Topographic Data, Canadian Council on Surveying and Mapping, Federal Department of Energy Mines and Resources, Ottawa, Ontario, January 1989.

DLG Format Definition User Guide, Document No. USL-88-03-31C, Universal Systems Ltd., Fredericton, New Brunswick, 1988.

Specifications and Guidelines for 1:20 000 Digital Mapping, Release 3.0, British Columbia Ministry of Crown Lands, Surveys and Resource Mapping Branch, Victoria, British Columbia, April 1988.

INTEGRATING SATELLITE IMAGERY WITH OTHER SPATIAL DATA SETS FOR MONITORING THE DOWNSTREAM IMPACTS OF DAM CONSTRUCTION IN NIGERIA

Paul G. Pilon and Philip J. Howarth

Earth-Observations Laboratory
Institute for Space and Terrestrial Science
Department of Geography, University of Waterloo
Waterloo, Ontario, N2L 3G1

Abstract

Difficulties obtaining and importing spatial data into geographically co-registered databases can hamper attempts to monitor large-scale resource developments. In northern Nigeria, although the adverse downstream effects of large-scale dam construction are acknowledged, the exact extent, location and nature of these impacts is largely unknown.

Spatial data sets from northern Nigeria were compiled, co-registered and analyzed using a TYDAC SPANS Geographic Information System. The data sets included. soil, vegetation and topological information obtained from existing maps, land use and land-use change information derived from analysis of satellite imagery; and, questionnaire and field survey data.

Significant correlations between many of the data sets indicate that the effects of dam construction are not equally distributed in downstream areas. Impacts are more detrimental in areas farther from the main river courses. Such information can help identify specific sites where new wells or pumps should be located to assist farmers.

Keywords: Information integration, Nigeria, Satellite imagery, GIS, Impact assessment

Introduction

In developing countries, such as Nigeria, large-scale resource development often proceeds without adequate attention being given to environmental considerations. In northern Nigeria the failure to acquire relevant resource information prior to constructing two large-scale dams has undermined the potential benefits of these projects. Post-dam assessments indicate decreasing agricultural productivity and increasing environmental deterioration associated with the construction of dams in this region (Adams, 1985; Turner, 1984; Wallace, 1981). This is very much related to the fact that environmental factors and traditional agricultural systems were not adequately examined during the project planning stage.

The Environmental Setting

The area of interest (Figure 1) is located in a semi-arid region of northern Nigeria, a region where harsh climatic conditions place extreme limitations on agricultural production. Water, a very scarce commodity, is concentrated along a few narrow river valleys and their floodplains (fadama). These river valleys form axes along which dense human settlement and agricultural activity have developed. Water concentrated along these floodplains during the wet season allows rice and other flood-tolerant crops to be grown. During the dry season, onions, peppers, tomatoes, tobacco and other high value crops can be grown by either relying on residual soil moisture or by using simple irrigation methods. In addition to providing a valuable source of water, the rivers also carry large amounts of sediment which are annually deposited across the floodplain during the rainy season, thereby renewing the fertility of floodplain soils.

Over centuries, the farmers who cultivate these floodplains have developed very sophisticated levels of adaption to the flooding regimes of these rivers. Consequently, an advanced agro-ecological system has developed at a micro-scale which is sensitive to seasonal and annual changes in climate and the physical environment. By the mid 20th century, however, population pressure, a series of droughts and the increasing threat of desertification caused considerable interest in the possibility of enhancing the agricultural potential of the region using modern irrigation methods. Consequently, several large-scale dams were built across the major rivers. The rationale was that the controlled release of water stored in reservoirs would allow a predictable flood during the wet season and provide water for irrigation during the dry season.

Impacts of Dam Construction

Dams constructed across the Sokoto River at Bakolori in 1979 and the Rima River at Goronyo in 1983 have altered the natural flooding regimes of these rivers. Water needed to fill the reservoirs during the

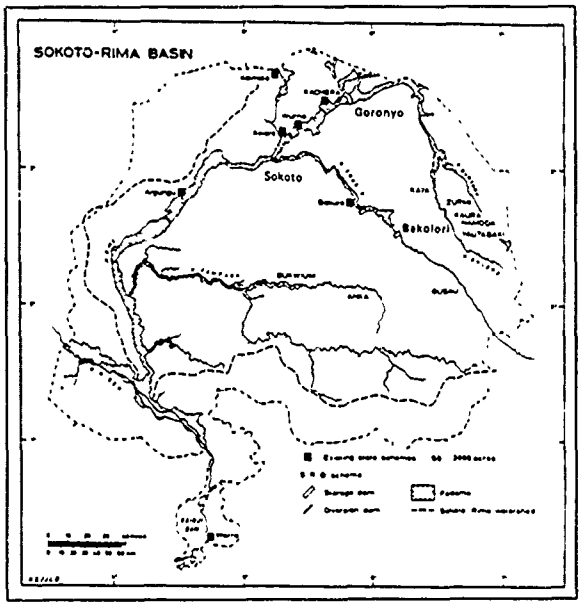


Figure 1: Map of the study area

wet season and to supply offstream irrigation schemes has reduced the maximum flood level in downstream areas (Figure 2). As a result, large portions of the downstream floodplain are no longer inundated during the wet season. By reducing natural ground water recharge during the wet season, floodplain pools dry up more quickly during the dry season and the level of water in wells is lowered. In some cases wells have dried up altogether. Although the flow of water in those sections of the river channels that were previously seasonal is now perennial, it is only available to farmers who have fields adjacent to the rivers and who have the means of transporting the water to their fields (e.g., by gasoline-powered pumps).

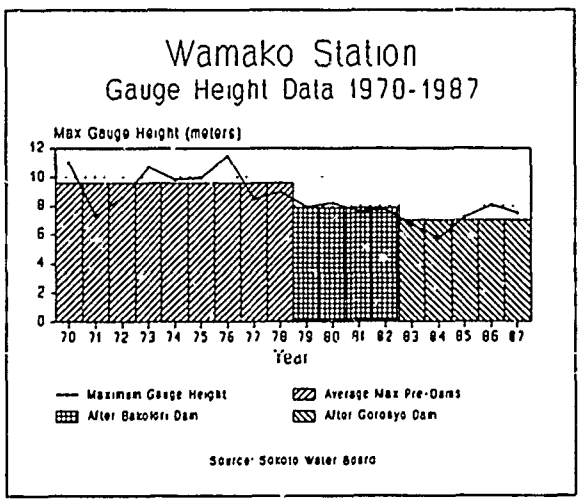


Figure 2: Gauge height data for the Rima River

A large amount of sediment is trapped in the reservoirs reducing sediment deposition downstream. This has not only reduced the fertility of floodplain soils, but has also increased the erosive capacity of the rivers. Due to reduced levels of sedimentation, farmers find that they have to use more manure or fertilizers on their farms to obtain the same level of crop production. As a consequence, by permanently altering natural floodplain processes, the agricultural productivity of downstream floodplain areas has been significantly reduced and in some cases floodplain land has been taken out of cultivation altogether.

Very little can be done to reverse the effects of these dams. Remedial measures aimed at optimizing the use of the remaining floodplain water resources must therefore be considered. For such measures to be successful, however, information on the precise location and nature of downstream impacts is required to determine precisely the amount and duration of flood needed during the wet season or where diversion canals, bore holes or pumps need to be located.

Any such assessment would require appropriate multitemporal information regarding conditions before and after construction of the dams. Although information about the current situation may be obtainable, the lack of any baseline study means that there are little data on pre-dam conditions. It is under such conditions that remote sensing techniques using satellite imagery can offer an efficient method of providing information that is neither technically nor economically feasible to obtain by conventional ground-survey methods.

In this paper, procedures for obtaining useful spatial data from multidate and multiresolution satellite imagery and for integrating these data in a geographic information system with other existing and derived data sources are described.

Data Inputs and Map Creation

The data input into the SPANS geographic information system were diverse (Table 1). Soil and baseline data were hand digitized on a Gentian digitizer using the TYDAC SPANS digitizing package TYDIG. Multidate Landsat MSS and TM, and SPOT XS and P images were initially geometrically corrected on a Dipix ARIES III Image Analysis System. Multidate enhancements and image classifications were conducted and the resultant images were transferred to a mainframe VAX for further data processing using an "enhanced classification" procedure developed by Pilon et al. (1988). The raw imagery, original classified images and the change results were imported into SPANS as raster data.

Point data defining the field-specific location of administered questionnaires and field survey data, and the locations of wells and pumps were hand digitized. By employing a systematic field-survey approach, interpolation between point data sources was possible.

This allowed thematic maps representing spatial variations in flooding, crop productivity, etc. to be produced from the questionnaire and field-survey data.

Table 1: Data inputs to the GIS

EXISTING DATA SOURCES

- Soils map
- Baseline data from topographic maps (roads, rivers etc.)

REMOTE SENSING DATA SOURCES

- Pre-dam and post-dam land use/cover information
- Location of land use/cover change
- Vegetation density mapping (using vegetation indices)
- Location of changes in vegetation density

DERIVED DATA SOURCES

From baseline information

- (e.g., buffer zones around roads, rivers)

From questionnaire/survey data

- Field-specific responses regarding changes in flooding, cropping, crop productivity, fertilizer use.
- Buffer zones derived from well and pump locations

Vector and point data could also be converted into thematic maps by creating corridors or buffer zones around points or lines. This was particularly important for this study because the distance from the rivers, wells or pumps was important in assessing the downstream impacts of dam construction.

Data Analysis

The downstream impacts of the Bakolori and Goronyo dams have been recognized for some time. This is evidenced in the testimony of farmers who, since the construction of the dams, have continually requested greater wet-season water release from the dams. The use of multitemporal satellite data provides an efficient way of providing appropriate information needed to properly address these concerns.

Initial analysis using satellite imagery focused on the effects of the Bakolori project. Landsat MSS images were obtained for anniversary dates before and after construction of the Bakolori dam (December 1975 and 1984). Using fairly simple multidate enhancement and image classification procedures, generalized information on the location and nature of agricultural land use/cover change in floodplain areas downstream of the dam were obtained. It was concluded from the initial investigations that significant changes had indeed occurred in downstream areas which could not simply be attributed to variable climatic conditions, or traditional cropping techniques, such as crop rotation or bush fallowing.

Although change was apparent, one could only infer that these changes in agricultural land use/cover were directly related to the effects of dam construction because other important information on changes in flooding, crop productivity and the location of wells and pumps was not available. Consequently, it was decided that a GIS approach would allow more comprehensive evaluations to be made of the impacts of dam construction in these downstream areas.

Once the required data layers had been compiled, co-registered and input into the TYDAC SPANS Geographic Information System, overlay analysis was used to identify those areas which were most adversely effected by dam construction. This involved weighting the attributes in each data layer with regard to their level of influence.

Preliminary Results

The results of the overlay analysis indicates that the impacts of dam construction are not equal in all downstream areas. The impacts vary laterally as one moves farther from the main river courses. Table 2, for example, indicates that since the construction of the Bakolori and Goronyo dams (approximately 6-10 years) over 50% of the area studied is no longer flooded during the wet season. Of those areas that have not flooded in over 5 years 80% are located at distances of more than 500 meters from the main river courses.

Table 2: Last flooding from main rivers

SINGLE MAP ANALYSIS

Map : LAST FLOODING FROM MAIN RIVERS (years)

Window: 00 - Universe

Class	Legend	Area (%)	Cum Area	Area (km sq)
1	< 2 years	2.15	2.15	0.12605
2	02 - 04	6.83	8.99	0.40017
3	04 - 06	9.68	18.67	0.56704
4	06 - 08	16.85	35.52	0.98685
5	08 - 10	12.89	48.41	0.75468
6	10 - 12	10.84	59.25	0.63484
7	12 - 14	9.30	68.55	0.54465
8	=> 14	5.88	74.43	0.34431
9	Upland Areas	25.57	100.00	1.49755
Total of 9 classes		100.00		5.85614

Table 3 is the result of comparing four data layers. These are: 1) distance from rivers; 2) distance from wells and pumps; 3) years since last flood; and, 4) change in crop productivity. Individual attributes for each data layer were assigned a weighting factor between 0 and 10. For example, areas within 100 meters of a river were assigned a weighting factor of 1 while those areas over 1000 meters were assigned a weighting factor of 10. Index level 6 in Table 3, for example, represents those areas more than 600 meters from a river which are also more than 600 meters from a well or pump, which have not flooded in 8 to 10 years and have experienced decreased crop productivity. The higher the index value for a given area, the greater the need for remedial measures such as installing new wells or pumps.

Preliminary evaluations of the satellite derived data also confirm these results. Information on land-use change indicates that dry season cultivation which was widespread across floodplain areas in the past is now concentrated along the major rivers. Not only has the total area of cultivated floodplain been reduced, many floodplain areas show conversion to lower value crops which are characteristically grown only on poorer upland soils.

Table 3: Areas requiring new wells or pumps

SINGLE MAP ANALYSIS

Map: Areas requiring new wells or pumps

Window: 00 - Universe

Class	Legend	Area (%)	Cum Area	Area (km sq)
1	Index 10	0.65	0.65	0.03803
2	Index 9	4.88	5.53	0.28580
3	Index 8	6.75	12.28	0.39544
4	Index 7	7.51	19.79	0.43978
5	Index 6	11.95	31.74	0.69970
6	Index 5	12.90	44.64	0.75560
7	Index 4	17.48	62.12	1.02371
8	Index 3	7.27	69.39	0.42568
9	Index 2	4.71	74.11	0.27600
10	Index 1	0.32	74.43	0.01887
11	Index 0	25.57	100.00	1.49755
Total of 11 classes		100.00		5.85614

Conclusions

Had the effects of the Bakolori and Goronyo dams on downstream floodplains been properly assessed and been included in initial project evaluations, they should have been enough to prevent or alter development of these projects. That they were not is a considerable problem for the entire region, and is indicative of some of the problems with the project-planning process in many developing countries.

Post-dam remedial measures must optimize the use of the remaining water resources in those areas which are most adversely effected. The types of information obtainable from remote sensing/GIS data integration which have been outlined may be useful for identifying specific areas where diversion canals, new wells or the installation of pumps may be required to assist local farmers

Acknowledgements

This research is supported by a Centre of Excellence grant from the Province of Ontario to the Institute for Space and Terrestrial Science. The authors' field studies were supported by IDRC, Ottawa.

References

1. Adams, W.M., "The Downstream Impacts of Dam Construction: A Case Study from Nigeria", *Trans. of the Inst. of Br. Geog.*, Vol. 10, pp. 292-302, 1985.

2. Pilon, P.G., Howarth, P.J., Bullock, R.A. and Adeniyi, P.O., "An Enhanced Classification Approach to Change Detection in Semi-Arid Environments", *Photogram. Eng. and Remote Sensing*, Vol. 54, No. 12, pp. 1709-1716, 1988.

3. Turner, B., "Changing Land Use Patterns in the Fadama of Northern Nigeria", in Scott, E.P., (ed.), *Life Before the Drought*. (London), pp. 149-170, 1984.

4. Wallace, T., "The Kano River Project, Nigeria: The Impact of an Irrigation Scheme on Productivity and Welfare", in Heyer, J., Roberts, P., and Williams, G. (eds.), *Rural Development in Tropical Africa*. New York: St. Martins Press, pp. 281-305, 1981.

REMOTE SENSING/GIS DATA INTEGRATION FOR GEOLOGICAL INVESTIGATION IN THE SOUTH MOUNTAIN BATHOLITH NEAR HALIFAX NOVA SCOTIA

S. M. Shupe[1] and M. S. Akhavi[2]

[1] Remote Sensing Consultant
[2] Senior Instructor, Remote Sensing Program
Nova Scotia College of Geographic Sciences

ABSTRACT

A portion of the South Mountain Batholith (SMB) west of Halifax was selected as an area suitable for use in an integrated remote sensing/geographic information system (GIS) investigation in a humid climate with an aim towards extracting structural and lithologic information. Thematic Mapper (TM) imagery over this area was processed and enhanced to allow for detailed lineament mapping and the detection of possible outcrop distribution. "Above-threshold" radioelement contours were digitized and then transferred from a GIS to an image analysis system (IAS) where they were superimposed upon the TM imagery to indicate areas where more evolved granites are likely to be located. Through an interface of the IAS and GIS, a final output map delineating lineaments, potential outcrop occurrences, and areas of "above-threshold" radioelement values was produced. Examination of the GIS database in conjunction with satellite imagery assists in adding to the knowledge of the existing geology and would serve as a compliment to a ground-based mineral exploration program.

Key Words: integration, GIS, IAS, Geology, Halifax Pluton

INTRODUCTION

A remote sensing/GIS data integration study was carried out over an area of granitoid rocks with variable glacial deposits and vegetative cover. The data integration resulted in the creation of an interpretive map that met the following objectives:

1. Map lineaments.
2. Map areas of probable outcrop.
3. Suggest regions of different granite phases.
4. Outline areas for potential mineral exploration.

The Study Area

The study area is approximately 600 km^2 and is located in the northeast portion of the South Mountain Batholith, (SMB) from Halifax to St. Margarets Bay (Fig.1). This area contains Devono Carboniferous granitoid rocks of the SMB which intrude Cambro-Ordivician metasedimentary rocks of the Meguma Group. Meguma Group quartzites (Goldenville Formation) and slates (Halifax Formation) contact SMB granites to the north and

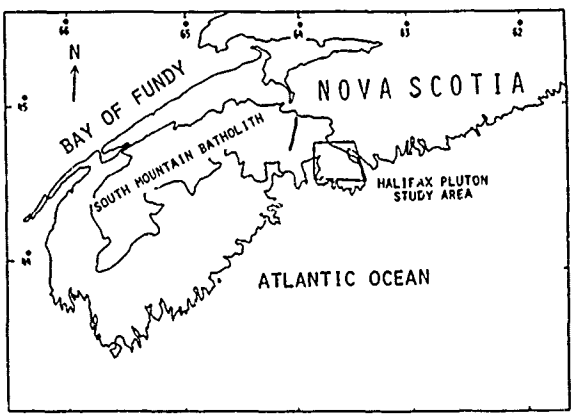


Figure 1 - Location of the Study Area

northeast parts of the study area. The peraluminus rocks of the SMB comprise a cogenetic suite ranging from granodiorite to monzogranite (Mckenzie and Clarke, 1975). The topography of the area is low and hilly with numerous lakes. Glacial till covers most of the area.

METHODOLOGY

The following procedure was used:

1. Enter lake and coastline boundary coordinates from the topographic map into the GIS to create a digital basemap.
2. Enter coordinates for selected "above-threshold" eU, eTh, and eU/eTh contour values from airborne gamma-ray spectrometric contour maps into the database.
3. Transfer the GIS data to an IAS and resample the satellite imagery to fit the GIS data.
4. Enhance satellite data using contrast stretches, principal component analyses, and spatial filtering.
5. Extract lineaments from enhanced satellite images using slide projections, digitize and enter into the database.
6. Highlight probable outcrop areas using band-

ratio images and principal component analyses. Classify and transfer probable outcrop themes back into the GIS database.

7. Generate composite interpretive lineament-outcrop map from final GIS database.
8. Interpret geology using the interpretive map and satellite imagery with and without radioelement contour overlays.

INTERPRETATIONS AND DISCUSSIONS

Interpretation of the dataset was initially carried out by Shupe (1987) then followed up by Shupe and Akhavi (in press). The present investigation is presented in response to the recent publication of an updated geological map (MacDonald and Horne 1987). A comparison of the remote sensing/GIS database interpretations with the new geologic map (Fig. 2) has resulted in a superior demonstration of the usefulness of the integration process.

Structure

Examination of the interpretive map (Fig. 3) reveals that a significant number of lineaments were detected from the enhanced Thematic Mapper data. The prominence of a number of these features on the satellite data indicates that fracturing in the grannites is likely the cause. For example, criss-crossing lineaments in the southeast are interpreted as conjugate fractures, similar to patterns observable in aerial photographs of exposed granitic rocks. A significant NW trending lineament located to the NE of Big Five Bridge Lake is clearly visible as a tonal boundary on the satellite images between two areas of different physiographies. This hints that a lithological difference may be present and that the lineament might be a portion of the contact.

Several lineaments are noted for their extensiveness, such as the one through Long and Fraser Lakes, and the one below Govenors Lake. These are interpreted as possible faults. In the NE of the study area above Bedford Basin, a lineament pattern is observed on the TM imagery which is the result of folding in the metasedimentary rocks. The trace of fold axes parallel to this structure can be presumed to possibly exist in some adjacent lineaments with similar trends.

Lithology

Observation of the location of the "above-threshold" radioelement contours suggests that granites of increased differentiation and different composition are present. Furthermore, the significant NW trending lineament within the eU = 3 and eU/eTh = 1 contours hints that a contact exists here. An imaginary line curving northward from the southeast of Big Five Bridge Lake to just southeast of Frasers Lake is the approximate boundary of concentrations of lineaments in the western half of the study area. This evidence suggests that this boundary represents an approximate termination of a lithologic unit and that the fractures might be confined to this unit.

Data for Mineral Exploration

Information can be derived from the interpretive map which may be useful for mineral exploration. Lineaments can be of use in locating fractures or contacts where mineralization is localized, eU and eU/eTh values can indicate more differentiated granites (Ford and O'Reilly, 1985), and areas of probable outcrop can point out sites in the field where lithologic conditions can be investigated. Based on the interpretive map (Fig.3) the location most conducive for exploration would be in the western half of the study area within the overlapping eU = 3 ppm and eU/eTh = 1 contour. Several fair-sized lineaments and areas of probable outcrop also occur here to support exploration.

Discussion

The structural and lithologic interpretations based upon the interpretive map and satellite imagery have been assessed with respect to a recent geologic map (MacDonald and Horne, 1987) which spans the study area. This map incorporates lineaments derived from air photos which correspond to a number of lineaments from the interpretive map. Some have no correspondence. However, most significantly, some of the most distinct lineaments from the enhanced satellite images do not appear on the geologic map suggesting that some major structural features may be unrecognized.

On an earlier geologic map of the area (Keppie, 1979) essentially all of the area below Halifax was mapped as a single monzogranite. The remote sensing data suggest otherwise. The recent geologic map supports such observations. Two leucomonzogranite bodies, probably connected at depth, have been mapped to lie within the anomalous eU and eU/eTh contours. Furthermore, their southern termini, as located on the geology map, quite closely lie on the significant NW lineament located NW of Big Five Bridge Lake. This lineament marks the boundary between two areas of differing tonal and textural patterns. In addition, a change from monzogranite to leucomonzogranite occurs on the geology map and the western contact roughly follows the trace of lineament concentration terminations mentioned above. Mapped geologic boundaries are in these cases approximate, being based on boulder distributions. In this study it appears that enhanced satellite images can provide valid insights into the subsurface geology and can be an aid in mapping.

CONCLUSIONS

A remote sensing/GIS data investigation was undertaken to derive information for the study of geology over a granitoid area. Interpreted information from a composite GIS produced basemap and enhanced Landsat TM imagery was compared to a geologic map published post to the initial investigation. Information derived from this study was found to have significant similarities to the published information collected from geological studies that were based on field work. This suggests that further refining of the geologic map is possible. In conclusion, this study indicates that geological work can benefit from remote sensing/GIS data integration techniques.

ACKNOWLEDGEMENTS

The authors are thankful to Mr. John Wightman, Principal of the Nova Scotia College of Geographic Sciences for his support and encouragement. The paper benefitted from a review by our colleague Mr. Edward Wedler. We are thankful to Ms. Marion Daniels for typing this paper.

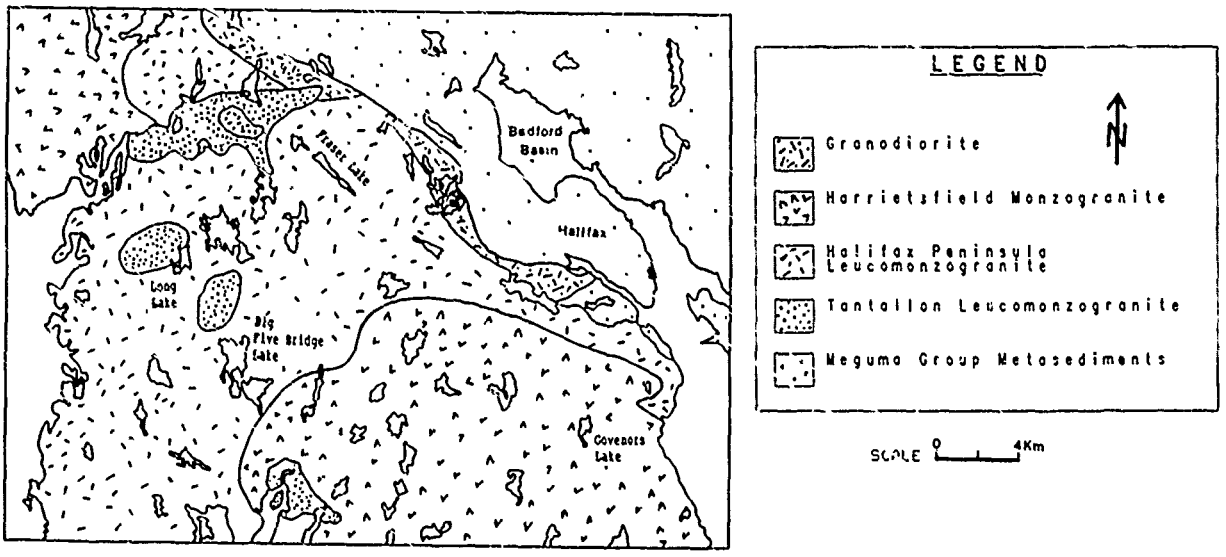


Figure 2 - Geological Map of the Study Area
(Generalized after MacDonald and Horne, 1987)

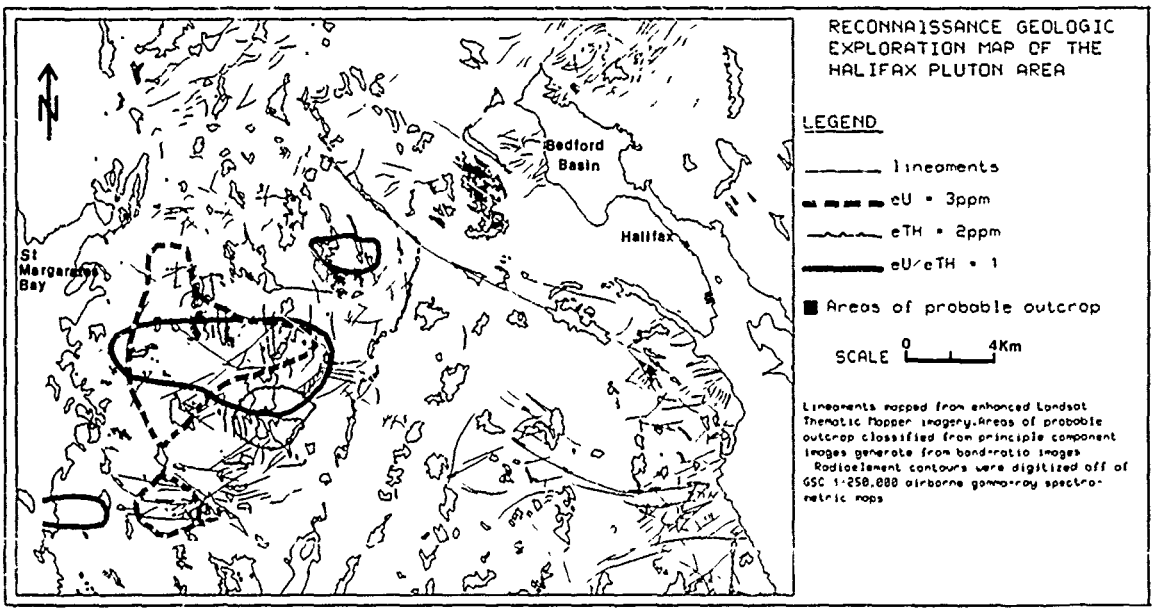


Figure 3 - Reconnaissance Geologic Map of the Halifax Pluton Area

REFERNECES

1. Ford, K. L. and O'Reilly, G. A., "Airborne Gamma-Ray Spectrometric Surveys as an Indicator of of Granophile Element Specialization and Associated Mineral Deposits in the Granitic Rocks of the Meguma Zone of Nova Scotia, Canada", Geological Survey of Canada, Special Paper, p. 1134-1233, 1985.

2. Keppie, J. D., "Geological Map of Nova Scotia", Nova Scotia Department of Mines and Energy, 1979.

3. MacDonald, M. A. and Horne, R. J., "Geological Map of Halifax and Sambro", Nova Scotia Department of Mines and Energy, Map 87-6, 1987.

4. McKenzie, C. B., and Clarke, D. B., "Petrology of the South Mountain Batholith, Nova Scotia", Canadian Journal of Earth Sciences, v. 12, p. 1209-1218, 1975.

5. Shupe, S. M., "Reconnaissance Mineral Exploration and Lineament Mapping in the South Mountain Batholith near Halifax, Nova Scotia using Landsat Thematic Mapper and Airborne Radiometric Data", unpublished diploma report, College of Geographic Sciences Library, Lawrencetown, Nova Scotia, 44 p, 1987.

6. Shupe, S. M. and Akhavi, M. S., "Integration of Remotely Sensed and GIS Data for Mineral Exploration: Halifax Pluton Area, Nova Scotia, Canada", Geocarto International, (in press).

MEETING GIS DATABASE REQUIREMENTS FOR NATURAL RESOURCE MANAGEMENT THROUGH THE INTEGRATION OF DIGITAL TOPOGRAPHY AND SATELLITE IMAGERY

Malcolm Gray, Evert Kenk, and Rostam Yazdani

British Columbia Ministry of Crown Lands
Surveys and Resource Mapping Branch
553 Superior Street
Victoria, British Columbia V8V IX5
Telephone: (604)387-1146 FAX: (604)387-3022

ABSTRACT

In order to be an effective tool for resource management, Geographic Information Systems (GIS's) require access to information that is not only comprehensive, standardized, digital, and current, but is also available Province-wide. Much of the existing provincial resource mapping fails to meet one or more of these criteria. In an attempt to overcome this lack of suitable data, a method is being prototyped that uses digital topography and satellite imagery to provide a geo-referenced resource database of ground cover, present land use and topographic features. This product is called a Baseline Thematic Map. Technical obstacles which are being addressed include; registration issues (including relief displacement), raster and vector format conversions (particularly as related to object definition and retention of attributes), and topographic segmentation of digital elevation models. Results have shown that an economically valuable database can be produced that is standardized, reliable, up-to-date and suitable for use by resource managers in their day-to-day decision making.

INTRODUCTION

The management of British Columbia's natural resources is becoming an increasingly complex task. Reasons for this trend include: a realization that our future prosperity depends upon present decisions, a shift from resource inventories to site specific management, integrated and coordinated management of multiple resources, and a requirement to more strictly monitor and regulate environmentally harmful substances and activities. The economic consequences of resource management decisions are very significant for British Columbia's economy.

Adequate knowledge of the land and its resources, including the capability to analyse this information is a prerequisite of proper resource management. With the advent of Geographic Information Systems (GIS's), the resource manager is provided with an essential and powerful tool which has the potential to efficiently meet the requirement for information analysis. For such a tool to be effective in a manager's ongoing activities, it requires access to resource information that is not only reliable, comprehensive, standardized, digital, and up-to-date, but is also available Province-wide. In this regard an examination of the resource mapping carried out to date in the Province leads to the conclusion that much of it is not suitable for use in a GIS.

In addition it has become increasingly difficult to justify new natural resource mapping using traditional methods for two reasons: time and cost. In an attempt to overcome these limitations for the successful implementation of GIS for resource management, an efficient, cost effective method using presently available data sources to provide suitable geo-referenced natural resource map data is being prototyped. The final objective is a product capable of justifying a production level implementation.

The method involves the production of three natural resource data bases, specifically; ground cover, present land use and topographic features. After integration into a single database the resultant product is called a Baseline Thematic Map. The databases are derived from an integration of digital topography and satellite imagery by a series of procedures involving image analysis, modelling and GIS processing.

EXISTING PROVINCIAL NATURAL RESOURCE DATABASES

From a Province-wide perspective only forest cover (at 1:20,000) and stream and river networks (at 1:50,000) have extensive coverage and are available in a digital form. Digital soils maps (at 1:50,000) cover about 25% of the Province. For limited areas (including East Vancouver Island, the Lower Fraser Valley, and the Okanagan Valley) a greater number of resource themes are available digitally.

Much of the existing resource mapping was done following different methodologies and at a variety of scales. For most themes the majority of the Province is not mapped and the mapping that is available is typically 10 to 20 years old. For many themes the information is still not digitized. In some cases the base maps used to geo-reference the data are not of high quality, and as well different bases for the same area have often been used for different themes.

The length of time to complete a traditional resource mapping project is usually two to three years or more. Typical costs per hectare for a single theme mapping project range from $0.30 to $0.46 for most reconnaissance level 1:50,000 mapping. In the case of broad reconnaissance 1:250,000 mapping this figure ranges from $0.02 to $0.03 per hectare. Single theme detailed mapping at a scale of 1;20,000 costs $1.85 TO $9.00 per hectare, depending upon ease of access and intensity of mapping. These costs include laboratory

analyses, all data coding and entry, map digitizing, and editing For a single, typical, National Topographic Series (NTS) 1:50,000 mapsheet the costs range from $26,000 to $40,000 and from $30,000 to $80,000 for a NTS 1:250,000 mapsheet. The costs for a British Columbia Geographic System (BCGS) 1:20,000 mapsheet vary from $26,000 to $126,000 Economies are achieved on multiple theme projects, with costs per theme on a two theme project averaging about 85% of those given above and for a three theme project about 75% of those figures.

These high costs provide opportunities for other methods to fulfil the information requirements of Provincial resource management, especially if the mapping can be accomplished in a timely manner.

GOVERNMENT INITIATIVES

The economic potential of GIS technology to support planning and decision making has been realized by the Government of British Columbia. As a result a number of significant initiatives are underway to implement this technology in the day-to-day business of the Province. Three such initiatives are of particular relevance to the case at hand. Firstly, the Ministry of Crown Lands is developing a corporate Land Information Strategic Plan to guide the development of an infrastructure to support the exchange of data between government ministries, Crown corporations and the private sector Secondly the Ministry of Forests is designing a Timber Supply Analysis System using GIS technology to support the resource management planning for Timber Supply Areas. Thirdly the Ministry of Environment is planning a GIS implementation strategy that will meet the business goals and objectives of the Ministry's administration, monitoring and planning roles.

SPECIFIC RESPONSE TO THE REQUIREMENT FOR NATURAL RESOURCE DATABASES

The task then is to produce a usable resource database in a cost effective and timely manner. Past pilot projects under taken by the Surveys and Resource Mapping Branch have shown that classification of satellite imagery can meet traditional map accuracy specifications (Kenk et. al., 1988). Results reported in the literature indicate that by including topographic information the number of classes mapped can be increased while maintaining map accuracy (Chibula et. al., 1987) Improvements in digital classification through the use of ancillary data (including digital elevation data) can be achieved in three ways: incorporating those data either before, during or after classification, through stratification, classifier operations or post classification sorting (Hutchinson, 1982).

The method being prototyped involves the production of three themes: ground cover, present land use and topographic features. The area chosen for this initial work is in the southern interior of the Province, specifically NTS mapsheets 82 E and 82 L/3.

The ground cover database provides the generalized type, extent and distribution of present vegetation cover. The main criteria for classification is vegetation structure and composition (i.e. physiognomy). The initial classification from the satellite imagery is refined using the relationship between specific vegetation distribution and topography. The ground cover theme is extended to include some detail on water bodies, wetlands, rock outcrops and the like.

The present land use theme provided is based on a modified form of the "Land Use Classification in British Columbia" (Sawicki et. al., 1986). The level of classification is restricted to land area-based uses versus site uses. Site uses of the land are generally not extensive enough to be identified on satellite imagery. It should be noted that a combination of spectral classification of satellite imagery and interpretation is used to produce the present land use classification.

The topographic features database is derived by segmenting the digital elevation model (DEM) into significant landscape units. These units are described in both raster and vector format. Watershed boundaries are delineated. Attributes that are attached to these units are the maximum, minimum, and mean values for: elevation, slope and aspect. Also included are. area, surface irregularity, and surface shape. Potential solar radiation received is also calculated. It is anticipated that these attributes will significantly add value to the final Baseline Thematic Map product for many users.

The satellite imagery utilized is presently Landsat TM, which has good spectral information for the identification of ground cover and for interpreting present land use. SPOT data has been used on a limited basis and its higher resolution is useful for identifying features such as new roads. Although not presently being used, airborne scanner data could also be utilized where appropriate but provincial coverage is limited.Two versions of the product are being prototyped, one suitable for use at 1.250,000 and the other at 1:20,000. These two versions correspond to the two available topographic data sets. The digital elevation model for the 1:250,000 version will be supplied from scanned 1:250,000 NTS mapsheets. Energy Mines and Resources Canada has scanned more than half of the sheets covering B.C. and expects to complete this work by the end of this year. The Terrain Resources Information Management (TRIM) program (the creation of 1:20,000 digital base mapping for B.C.) will supply the digital elevation models for the 1:20,000 version. Presently there are about 500 of the 7000 map sheets completed with a current production rate of approximately 700 map sheets per year.

METHODOLOGY

In the following discussion the emphasis is placed on those tasks where critical technical limitations had to be overcome for the process to proceed. These include. relief displacement correction, raster to vector format conversions particularly with regard to attached attributes, and topographic segmentation of digital elevation models. Figure 1 presents an outline of the methodology.

For the 1.20,000 version the new digital mapping from the TRIM program will be the base map of choice. For the 1:250,000 version product, scanned 1:250,000 NTS mapsheets will provide the base map. The datum for the prototype work is NAD 83 (North American Datum 1983). It is envisioned that all provincial mapping will migrate to NAD 83 over the next five years. Transformations to other datums (NAD 27) and projections are well defined for those users desiring them.

The first manipulation of the input data is to co-register the

different data sets (imagery, DEM's, and other ancillary map data). Obviously the precision requirements are greater for a scale of 1:20,000 compared to 1:250,000. Because some of the input data will be used to produce both versions of the final product it is logical to register the data sets to the higher precision requirements. If this is done at the outset then scale transformations will not result in location changes due to different registration standards. To avoid any systematic subpixel misregistration all rasters follow the same convention about the origin with respect to the reference UTM grid. UTM coordinates ending in 00 (i.e. every 100 metres) align with the boundary between pixels. Pixel sizes are chosen in a nested hierarchy as follows: 6.25, 12.5, 25, 50, and 100 metres. The actual size is dependant upon the data utilized and the mapscale produced. For example, at 1:20,000 scale mapping, the following pixel sizes are used: 25 m for Landsat TM, 12.5 m for SPOT MLA, and 6.25 for SPOT PLA.

For areas of high relief Landsat TM imagery has geometric distortions due to the panoramic view of the sensor. At the edge of the 185 kilometre wide image swath the look angle is 7.47 degrees, causing a horizontal displacement of 30 m for every 230 m change in ground elevation (Wong et. al. 1981). Many parts of B.C. have relief in the order of thousands of metres. This is an unacceptable amount of distortion for 1:20,000 mapping. Correcting for relief displacement is accomplished using a DEM. The elevation precision of the scanned 1:250,000 topographic data is approximately +/- 120 m. As well as the vertical precision of the data, the accuracy with which positions can be located horizontally effects the precision of the derived elevation. Currently we are investigating the precision attainable utilizing different sources of topographic data for relief displacement correction.

Another concern arising from rectification and relief displacement is the change in image radiometric values due to resampling. If these changes are shown to be significant to the classification results there is the option of rectifying the imagery after classification. This is less desirable than pre-classifcation rectification because it makes the use of ancillary map data for training site selection and accuracy evaluation more difficult.

Image classification for ground cover identification proceeds from field work to selection and refinement of training sites on the satellite imagery. Single ground cover classes are described by more than one set of training sites so illumination and ecological differences due to topographic position are accounted for. After classification with these seperate training sites, the results are merged into a single class type.

Image classification for present land use involves digital classification as well as interpretation of the imagery. Some of the interpretation depends upon pattern recognition and adjacency considerations. This is accomplished partly through the use of the Procom projection device with the results being digitized into the existing digital base.

Post-classification processing is of two main types: context filtering to transform the raw image classification into a more acceptable map like product and modification of the resulting classification through the use of ancillary (primarily topographic) data. Context filtering makes use of the similarity between classes as well as length of common boundary and minimum area rules for amalgamating groups of pixels. This kind of filtering results in a cartographically acceptable map and is also required to reduce the number of homogeneous areas (polygons) as many GIS's are limited in the number of polygons that they can handle. A record is kept of the percentages of the original classes that were amalgamated to form the single class output polygons. Modification of classification results utilizing ancillary data takes place both in the image analysis domain and the GIS domain. Ancillary data is used to construct masks within the image analysis system then Boolean logic is employed to separate classes that are spectrally similar. Equivalent results can be achieved through GIS processing, with potentially a wider selection of attributes available for post-classification accuracy improvement.

Topography plays a dominant role in many biological and geomorphic processes. Terrain segmentation as implemented is based on: surface morphology, hierarchical watershed boundaries, and potential solar radiation received at the surface. Segmentation based on morphology begins with irregularly spaced points, from which a triangulated irregular network (TIN) is created. Units are defined through the amalgamation of neighbouring triangles, based on similarity with respect to slope and aspect. Watersheds are formed by the examination of hydrologic flow across the triangles. Radiation based segments are created through merging of contiguous triangles based on criteria examining the results of shading and shadowing values sampled at set intervals from sunrise to sunset over specific days of the year.

At some point the raster products require conversion to a vector format to be compatible with most current GIS format requirements. Currently our capability is limited in the size of raster supplied as input and the number of resultant polygons output. Presently a single value attribute can be attached to the vector database created; there is a requirement for the transfer of multiple attributes, for example the record of the original composition of a context filtered polygon.

FINAL PRODUCT

GIS capabilities are used to merge the ground cover, present land use, and topographic feature themes into a single integrated database, the Baseline Thematic Map. Final editing takes place at this point. For those GIS's that support 256 colours or more it is proposed that a realistically coloured 8 bit image could be supplied to give the user a better visual context to aid in map use.

DISCUSSION

This method for providing a highly standardized Province-wide resource database is being actively prototyped now. This product has the potential to meet many of the information requirements of Provincial resource management. If the results of the prototype effort merit a production level implementation it is realistic to expect the private sector to become involved in the production of Baseline Thematic Maps.

REFERENCES

1. Cibula, W. G., and Nyquist, M. O., "Use of Topographic and Climatological Models in a Geographical Data Base to Improve Landsat MSS Classification for Olympic National Park", Photogrammetric Engineering & Remote Sensing, Vol. LIII, No. 1, pp 67-75, Jan. 1987.

2. Hutchinson, C.F., "Techniques for Combining Landsat and Ancillary Data for Digital Classification Improvement", Photogrammetric Engineering and Remote Sensing", Vol. XLVIII, No. 1, pp 123-130, Jan. 1982.

3. Kenk, E., Sondheim, M., and Yee, B., "Methods for Improving Accuracy of Thematic Mapper Ground Cover Classifications", Canadian Journal of Remote Sensing, Vol. 14, No. 4, pp 17-31, May 1988.

4. Sawicki, J., and Runka, G., "Land use Classification in British Columbia", B.C. Ministry of Environment Manual 8, 1986.

5. Wong, F., Orth, R., and Friedmann, D. E., "The Use of Digital Terrain Model in the Rectification of Satellite-borne Imagery", Proceedings of the Fifteenth International Symposium on Remote Sensing Of Environment, Ann Arbor, MI, pp 653-662, May, 1981.

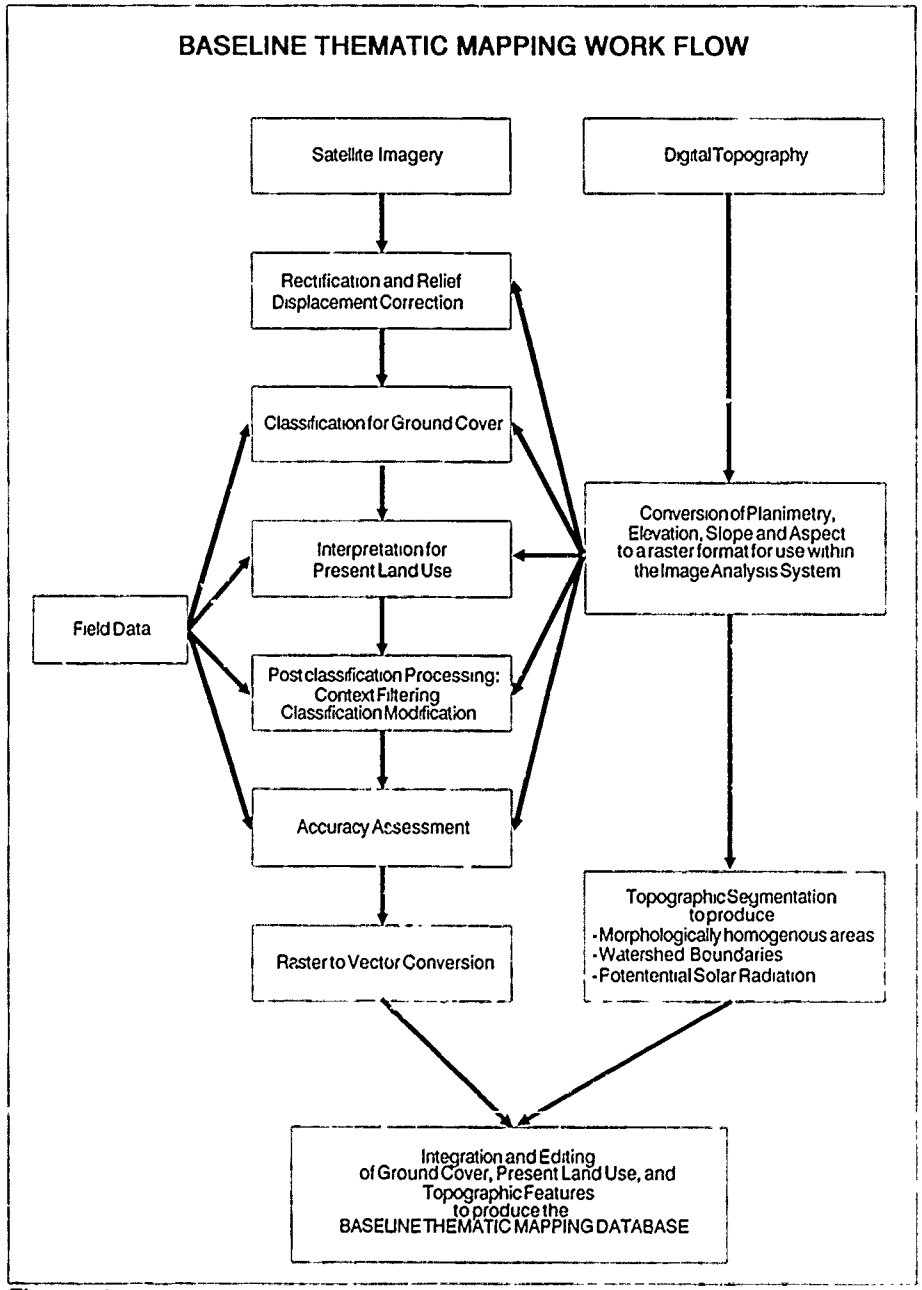


Figure 1.

THE INTEGRATION OF REMOTE SENSING DATA WITH A GEOGRAPHIC INFORMATION SYSTEM TO DEVELOP FOREST COVER PARAMETERS FOR FOREST WILDLIFE IMPACT ASSESSMENT IN SOUTHERN ONTARIO

C.L. Wagner, R.N. Pierce, and S.W. Hounsell

Land Use and Environmental Planning Department
Ontario Hydro
Toronto, Ontario

ABSTRACT

Image analysis of three Landsat TM scenes was performed to map the forest cover within a portion of agriculturally dominated southwestern Ontario. This information was transferred in raster format to a GIS where it was vectorized and used in subsequent calculations. The measurements of the area and edge length of each forest stand within sample areas were calculated and used as parameters in a habitat impact prediction model. This model measures the effects of woodlot fragmentation on bird species inhabiting the woodlot. This study demonstrates the use of satellite imagery and GIS integration in a large scale operational environmental project.

KEYWORDS: Bird habitat, Environmental modeling, Forest classification, GIS integration, Landsat TM

INTRODUCTION

The construction and maintenance of high voltage electric transmission facilities can result in impacts to the natural environment. Ontario Hydro conducts impact assessment studies for proposed facilities. One component of these environmental assessments is the prediction of potential impact on forest wildlife habitat.

The clearing of transmission line rights-of-way results in the fragmentation of forested areas. This can have an adverse effect on an ecological grouping of birds that may generally be described as area-sensitive, forest-interior specialists, many of which are long-distance, neotropical migrants. Populations of these species are expected to decline as forest size becomes smaller, forest edge becomes more prominent and as woodlots become more isolated from neighbouring woodlots in the landscape. The fragmentation of forests can, therefore, lead to a decline in forest interior bird species, and in the extreme cases, can lead to regional expirations of these species (Whitcomb et al., 1981, Freemark and Merriam, 1986). In contrast, bird species inhabiting field and forest edges will tend to flourish in fragmented landscapes. In this regard, Temple and Wilcox (1986) maintain that "habitat fragmentation looms on the horizon as perhaps the single most significant challenge to the development of models applicable to wildlife management, if not ultimately to the survival of wildlife altogether".

Ontario Hydro has developed an impact prediction model for assessing the effects of fragmentation on forest bird species in southern Ontario (Hounsell, 1988). This model incorporates the woodlot parameters of shape, area, edge length and distance to neighbouring woodlots. In order to apply the model there was a need to map the forested land within the region. It was the purpose of this study to use suitable remote sensing images and established image analysis techniques for deriving forest cover data. As well, the mapped data was to be transferred to a vector-based geographic information system and used in determining measurements for the habitat impact model. This study was carried out concurrently with another project which incorporates the model by testing hypotheses regarding relationships between the characteristics of woodlots and the number of interior woodland breeding bird species.

STUDY AREA AND DATA SOURCES

The area chosen for this study extends from Toronto west to Lake Huron and all of the area south (approximately 33,000 square kilometres). This is the deciduous forest region which lies within the agriculturally dominated region of southwestern Ontario.

Landsat Thematic Mapper (TM) imagery was obtained to cover the study area. All or part of three scenes were required. An attempt was made to acquire images from 1985 to correspond with the bird species data to be used. However, due to cloud cover, 1985 imagery for the appropriate season could not be obtained to cover the whole study area. The images acquired were Track 18 Frame 30 for September 20 1985; Track 19 Frame 30 for August 2 1988; and Track 19 Frame 30+10 for August 29 1986.

The imagery was analyzed using the Dipix ARIES III image analysis system in the Land Use and Environmental Planning Department. The forestry data was then transferred to the department's CARSS (Computer Assistance for Route and Site Selection) GIS. CARSS is a system of integrated processing functions using a variety of hardware and software componerts. It has the capabilities to record, store, manipulate and output data in both raster and vector format. The Raster Processing System, developed in-house, is a tool designed for

transforming, modeling, and analyzing raster data. Vector data is manipulated through the use of ESRI Arc/Info functions. The ARIES system is interfaced with the GIS through the host VAX 11/785 minicomputer.

CLASSIFICATION

All three TM scenes were analyzed using similar steps. Each scene was geometrically corrected to the UTM map base using 1:50,000 NTS topographic mapsheets for ground control point generation. This is standard practise for the department as information derived from image analysis is often transferred to the GIS and must be registered to the same map base. A supervised classification procedure was then carried out. Through the use of visual interpretation of the imagery, aerial photographs, 1:50,000 topographic maps, and field checking, training areas were developed for forest cover classes as well as for some agricultural cover types. It was found that it was necessary to include these non-forest cover types to help reduce some confusion between the classes of field crops (especially corn) and deciduous tree cover. Signatures for the cover types were created and applied to a maximum likelihood classifier. In an attempt to minimize the confusion, many signatures were manipulated. Combinations of TM bands and ratios were tested. As well, purification of training areas and signatures was performed to reduce the spectral overlap.

It was found that the manipulations of the signatures either did not improve the accuracy of the classification results or improved it in one area of the scene while decreasing it in another area. The signatures used in the final classification included six TM bands (bands 1 to 5, and 7) at a 95% statistical rejection level.

The result of the classification procedure was a thematic map which included deciduous, coniferous, mixed wood, corn and soybean. The forested areas were correctly classified and, for the most part, completely classified. Farm fields which remained misclassified as forest cover were removed manually. A 5 pixel filter was then applied to the classification to remove extraneous pixels.

DATA TRANSFER AND MANIPULATION

The information on breeding bird species is stored in the form of 10 km by 10 km sample squares located throughout the study area. This data was derived from the computer records of the Ontario Breeding Bird Atlas (Cadman et al., 1987). The Atlas represents the product of a five year effort to document the breeding status and distribution of the birds of Ontario. This data base was collected, geo-referenced and stored based on the Universal Transverse Mercator (UTM) grid system.

A stratified random sampling technique was used to sample 100 squares from the study area to test selected hypotheses from the predictive model and integration procedure. Random sampling of squares were chosen from three categories of woodland coverage: greater than 50% woodland; between 20% and 50% woodland; and less than 20% woodland. For each square selected, the forest cover data was transferred from the image analysis system, in raster (cell) format, to the GIS database. It was then copied to Arc/Info where it was converted to vector format. At this point, editing can be done to smooth the polygons, if desired.

The parameters chosen for testing selected aspects of the model include:

- number of woodlots in each square
- area of each woodlot in each square
- total forest area for each square
- length of forest edge of each woodlot in each square
- edge to area ratio for each woodlot and average for total square
- a measure of woodlot connectivity and spatial organization.

These measurements can be derived easily using Arc/Info functions. The area and perimeter of each polygon (woodlot) is automatically calculated when the data is vectorized and/or edited. Calculations, such as the ratio of edge (perimeter) to area, can then be done for each polygon and added to the Info listing for the file.

The use of these parameters in the impact prediction model involved the comparison of these forest habitat parameters with the bird species information contained in the computer database of the Ontario Breeding Bird Atlas. This work was conducted under the concurrent project objectives.

CONCLUSIONS

Forest cover in southern Ontario was successfully mapped through the use of satellite digital imagery. It was found that the image date is very important when classifying deciduous forest cover in agricultural areas. Ideally, the image should be selected for a period within the growing season which would minimize confusion between deciduous and crop signatures.

Satellite imagery proved to be an efficient source of information for environmental impact modeling. The imagery provides coverage over an extensive area and image analysis provides a means of extracting information in a cost-effective manner without the detailed analysis required to interpret Forest Resource Inventory maps, topographic maps or aerial photography.

The transfer capabilities between the image analysis system and the GIS simplify the acquisition of data by eliminating the need to manually digitize. This study has demonstrated that the integration of remote sensing and GIS technology can be used in an operational environmental impact project. It has established an effective and efficient method that can be used for future modeling and transmission route and site selection projects.

REFERENCES

Cadman, M.D., P.F.J. Eagles and F.M. Helleiner 1987 "Atlas of the Breeding Birds of Ontario" University of Waterloo Press, Waterloo.

Freemark, K.E. and H.G. Merriam 1986 "Importance of area and habitat heterogeneity to bird assemblages in temperate forest fragments" Biol. Conserv. 36: 115-141.

Hounsell, S.W. 1988 "Methods for assessing the sensitivity of forest bireds and their habitats to transmission line disturbances" Land Use and Environmental Planning Department, Ontario Hydro, Toronto. Unpublished report. pp. 469

Temple, S.A. and B.A. Wilcox 1986 "Predicting effects of habitat patchiness and fragmentation" In "Wildlife 2000 - Modeling habitat relationships of terrestrial vertebrates" Ed. by J. Verner, M.C. Morrison and C.J. Ralph. pp. 261-262. University of Wisconsin Press, Madison.

Whitcomb, R.F., L.S. Robbins, J.F. Lynch, B.L. Whitcomb, M.K. Klimkiewicz and D. Bystrak 1981 "Effects of forest fragmentation avifauna of the eastern deciduous forest" In "Forest island dynamics in man-dominated landscapes" Ed. by R.L. Burgess and D.M. Sharpe. pp. 125-205. Springer, New York.

INTEGRATED GIS AND IMAGE ANALYSIS SOFTWARE. APPLICATION IN VEGETATION MONITORING

P. Landriau
Dipix Technologies Inc.
1050 Baxter Rd., Ottawa, K2C 3P1
tel: (613) 596-4942
FAX: (613) 596-4914

M. J. Manore
Canada Centre for Remote Sensing
2464 Sheffield Rd., Ottawa, K1A 0Y7
tel: (613) 952-9735
FAX: (613) 952-7353

The integration of remotely sensed imagery with map layers contained in Geographic Information Systems (GIS) is recognized as an important step in effectively using satellite imagery for resource management. In addition to a variety of technical issues that need to be confronted, the methods for combined image and map analysis for different applications must be identified This paper describes one software implementation for GIS/image analysis integration and how it was used for regional vegetation monitoring in western Canada.

The approach adopted for image analysis/GIS integration in this case was to implement a sub-set of common GIS functions within the image analysis system. To this end, software was written to import and export DLG format vector map files (and their associated user attributes) to and from Dipix ARIES image analysis systems and to provide database query, statistics extraction, and vector edit functions. The functionality of the software package is described in the paper

The Crop Information System, jointly established by the Canada Centre for Remote Sensing and the Manitoba Remote Sensing Centre, routinely processes NOAA-HRPT image data into mapped, near cloud free images of the grain producing regions of the Canadian prairie. Map layers imported to the image analysis system from a GIS are used for assessing the accuracy of geometrically corrected imagery and to overlay on images for visual and geographic reference. Administrative and physiographic polygons, also imported from the GIS, are used as training areas and the derived image statistics are added as new attributes to the imported attribute database. Further analysis on the extracted radiance statistics may be performed when the data are returned to the GIS where complementary databases reside (meteorological, historic production, etc.).

KEYWORDS attribute, data transfer, GIS, polygon, remote sensing, statistics, view

Introduction

Digital remotely sensed imagery data have provided valuable data on earth resources for many years, and the ability to extract 'information' for resource management from these data has been largely through spectral and textural processing routines available in Image Analysis Systems (IAS) With the rapidly increasing use of Geographic Information Systems (GIS) as data bases for resource information, there is a strong incentive to link IAS and GIS technologies so that the complementary information contained within each system may be used together. The creation of such a 'spatial' information system has encountered many technical issues, including the combination of the raster and vector data structures commonly used by the two systems. There are many solutions to the mechanics of IAS/GIS integration, one of which is described in this paper Because IAS and GIS vendors have historically developed their products seperately, current IAS/GIS integration often involves exchanges of data between two proprietary data formats and specialized display hardware This paper describes software which permits the integration and manipulation of ESRI ARC/INFO map coverages within the Dipix ARIES image processing system The application of this software to vegetation monitoring in Western Canada (the Crop Information System) is also discussed.

Technical description of the Aries-ARC/INFO software.

This new Aries software was designed to provide an intelligent access to ARC/INFO covers. The first step in this software was to provide a mechanism to combine the graphics data from ARC/INFO with the digital remotely sensed raster data of the Aries system. Since the Aries display is designed around a bulk memory raster mapped architecture the vector graphics data could not be used in their original form. The intention of this software package was not to duplicate the

efforts of ESRI, but rather to offer a subset of GIS functions for data manipulation within a raster based image display system. Note that it is very important to ensure that the polygon cover is "clean" before exporting it out of ARC/INFO. In particular undershoots and overshoots will distort the results of the statistics generation process and will create difficulties with the editing

The software comprises several modules in three tasks. These are.

DT database transfer

The graphics data is read in and converted to an Aries vector data format which allows quick access to the graphics database. Another advantage of the Aries vector format is that the accuracy of the original map coordinates is maintained, even though subsequent display and manipulation is performed in the less precise raster environment. This Aries graphic file is then rasterized This rasterization can be done for the whole database or for any subarea a user wants to specify Since the size of the pixels can also be specified the rasterized vector data can be made to fit any type and size of remote sensed imagery The result of the rasterization is a cover theme file with a single theme representing the vector data

DQ database query

The database query task can then be run This task allows the user a number of high level functions such as displaying the rasterized cover, listing the attribute dictionary, examining and editing the attribute data of the elements (an attribute record being a record which contains information about a particular element), data queries on the database and saving the results into views (a view being a set of selected elements) and lastly statistics generation on any selected elements

With the display function a user can overlay an rasterized ARC/INFO cover on top of remotely sensed imagery. Up to 3 features or 4 theme background files can be loaded up along with the rasterized cover, as 16 or 32 bit composites. Note that the amount of the data that can be displayed is not limited to the size of the display monitor but the size of the bulk memory. So if a user wished to only display the rasterized cover, then even though the display is only 1024 by 1024, a 1600 by 1600 cover could be displayed; all of it accessible and displayable through the roam function of the display monitor. Of course much less can be displayed if the user loads up the rasterized cover with 3 8 bit features.

The data dictionary option is used to control access to an APS (attribute processing system) database by defining the structure of the database. By describing the attributes, the data dictionary records the structure of the attribute records. By describing views, the data dictionary enables access to the results of significant logical queries or selections applied against the database, and which the user has chosen to name and save. Besides describing the structure, the data dictionary also acts as an index which links the attributes

The attribute option allows the user to view or modify attributes. Attribute displayed are keyed by -ID number, provided by the ARC/INFO database. For the most part the attribute option is driven by the current select list (that being the list of elements last selected by a query or the last view recalled by a user). By default if there is no current select list then all elements are considered selected. A number of subfunctions are provided such as which attributes of an element to display and editing of the attribute data. Also an output list can be generated to a file or printer

The ad hoc query option provides the user with means of selecting a subset of the elements in a cover which satisfy certain properties as determined by attributes in the APS database, or the results of previous logical queries. A query may consist of up to 15 statements, each a simple logical expression. These statements form a logical list of .AND.s and .OR.s with which elements will be selected. This selected set is then the current element set. It can be edited with a manual selection option, which will highlight the current data set on the display monitor and allow the user to add or delete elements by pointing to them. It can also be saved through the view option and given a user selected name.

The view option lets the user save a set of selected elements under a user given name. It also allows the user to retrieve a previously selected set (if that set was saved) and make it the current data set.

Particularly useful is the statistics generation function. This function allows the user to generate a set of statistics from a set of background features or theme files on a per polygon basis. The set of polygons for which the statistics will be generated must be contained in a named view, or the polygon can be selected directly from the display monitor.

The statistics produced for the feature files are the number of pixels, the mean, the mode, the median, the standard deviation, the skewness, the minimum and the maximum The mean, median and skewness are calculated through the use of a histogram and their accurracy will depend on the bin (sample) size. This is usually within 1% of the accepted value for a normal distribution of values. Statistics calculated from theme files can display the number of pixels per polygon or the number of pixels per theme per polygon or the percentage of total per theme per polygon The user can generate statistics for up to 32 themes or 32 feature files These attributes can then later be transported back to the ARC/INFO system.

When using feature files, the statistics generation algorithm will allow the use of a user defined arithmetic expression to calculate the statistics. This permits arithmetic combinations of several features to be calculated and saved in real numbers, without the need to store results in an 8-bit integer image file A small 8 lines by 70 characters editor is provided to enter the arithmetic expression

In converting the vector data to raster information, an infinitely thin polygon border obtains a thickness of 1 pixel, therefore the user will have the option to include or not include the border pixels in the statistics If the border pixels are included, then they may be included as a half pixel (which assumes that all borders pixels are shared by exactly 2 polygons) or by using a pixel count file which will report on how many polygons each border pixel belongs to. Of course this pixel count process will take more time.

Polygon editing.

Another essential part of the software package is its element editing capabilities Since the vector cover can be rasterized and displayed as an overlay on top of actual remote sensing data, a user can base any editing on visual interpretation The need to update a cover is very important so a short description of the editing capabilities is in order In this case the polygon editing is described.

In a polygon coverage we have 2 types of elements. Arcs and polygons. Each polygon is made up of a number of arcs and each arc belongs to 2 polygons A number of functions permit the user to edit the database.

Arc reshape: this allows a user to select an arc from the display and to reshape it. The restriction being that it's end points are not moved. This permits the user to do small corrections to an arc with the minimum amount of database modifications.

Arc delete: the user selects an arc to be delete by pointing to it on the display monitor This arc is then deleted from the database and all relevant polygons are immediately updated. That is, two polygons are merged and, if the required of the arcs of the newly formed polygon are also merged.

Arc adding: arc adding will let the user create new polygons. This function provides useful tools such as snap to arc and snap to node. Again the database is updated immediately. That is the affected polygon is cut up into two new polygons, for which areas, perimeters and label points are calculated. The user is prompted for the attributes of the new polygons.

All these editing functions are performed directly on the display using a responsive bit-pad. Because of the method of storing the database in Aries format any update to it will take the same amount of time regarless of the size of the database. When a user is finished editing there is not need to rebuild the coverage on the Aries side as it is always current. The edited coverage can then be transferred back to the ARC/INFO system.

Example Application.

The Crop Information System (CIS) is a joint program between the Canada Centre for Remote Sensing, the Manitoba Remote Sensing Centre, and several other agencies with interests in agricultural activity in Western Canada. The purpose of the program is to provide timely information on vegetation condition during the growing season which may assist in forecasting potential yield and production. At present, the principal source of data is daily AVHRR imagery from the NOAA series of meterological satellites. The following is a brief description of the standard CIS image product.

Once geometrically registered to a map projection using standard ARIES transform and resampling software, a seven-day series of AVHRR images are combined into a single image comprising only the 'best' data from the input images. The selection of these data is based on the highest Normalized Difference Vegetation Index (NDVI) and has the effect of creating a near cloud-free image in which atmospheric path radiance has been minimized. . These products, termed 'composite images', have been shown to correlate with biomass and may be used to

assist in assessing vegetation condition (Tucker and Sellers, 1986; Goward et al., 1985). Registered composite image products are produced weekly during the growing season for the agricultural regions of Manitoba, Saskatchewan, and Alberta.

The link between ARC/INFO and the ARIES is used in several ways in CIS processing. A conceptual overview of the flow of data between the IAS and GIS for this application is available in Manore and Brown (1989). Some of the uses of the ARIES-ARC/INFO link for the CIS are described below.

Initially, a map coverage of the large bodies of water in the study region was created within the GIS and then transformed into the same projection to which the imagery had been mapped. The coverage was then transferred to the ARIES and rasterized to provide a simple overlay which is used to verify the success of image registration prior to the compositing process. Map layers of provincial and administrative boundaries (i.e., Crop Reporting Districts, Rural Municipalities) were also transferred and rasterized. These are used as theme overlays on the composited images to provide geographic reference and to assist in image interpretation. District names and other attributes in the coverage data base (i.e., historic yields) may also be interactively browsed and displayed on the image using the attribute processing functions. The water body map was also used to enhance the visual quality of the images by masking out the cloudy appearance of lakes (an artifact of the compositing process). Solid theme masks were generated from the transferred coverage by creating training areas directly from the selected set of water polygons.

A routine analysis in the CIS is to extract image statistics within each administrative or physiographic map unit from each weekly image. Through the query software, these values may be extracted in batch mode from any selected set of map polygons and placed directly into the attribute data base. When returned to the GIS as attributes, they are be used to produce cartographic representations of the image data, to feed external statistical packages, and in the near future, will be used in yield forcasting models.

The CIS does not make use of all the functions available in the software package. Among those functions not used are the editing of map polygons for map update, and the vectorizing of image classifications.

Conclusion

This application shows the importance and usefulness of integrating remotely sensed imagery and GIS data in vegetation monitoring. Note that this type of data integration can be carried over to many other applications.

References

Goward, S.N., C.J. Tucker and D.G. Dye, 1985. "North American Vegetation Patterns Observed by NOAA-7 AVHRR", Vegeation, Vol. 64, No. 3.

Manore, M.J. and R.J. Brown, 1989. "GIS in the Analysis of Satellite Data for Vegetation Monitoring: The Crop Information System", Proceedings, National Conference on GIS, Ottawa, February 28- March 3, 1989.

Tucker, C.J. and P.J. Sellers, 1986. "Satellite Remote Sensing of Primary Production", International Journal of Remote Sensing, Vol. 7, No. 11.

THE USE OF REMOTELY SENSED DATA IN THE CREATION OF A GEO-DATABASE FOR ANTIGONISH, NOVA SCOTIA UTILIZING A MICRO-COMPUTER BASED GIS

Tim Webster

TYDAC Technologies Inc.
1600 Carling Avenue
Suite 310
Ottawa, Ontario K1Z 8R7

M. S. Akhavi

Nova Scotia College of
Geographic Sciences
Lawrencetown
Nova Scotia B0S 1M0

ABSTRACT

The purpose of this project was to integrate several geologic data sets utilizing the SPANS micro-based GIS. Presently a variety of geological data exists in several formats; satellite imagery, airborne sensors, hardcopy maps, geochemical surveys, etc. In order to better understand the relationships of these various data sets, a GIS can be used to analyze the integrated database.

The creation of the geo-database involved data transfer between mini and micro-computer based software. The remotely sensed data consisted of Landsat MSS (79 x 79m) and TM (30 x 30m) coverages, airborne vertical gradient magnetics (125 x 125m), and seven channels of airborne radiometrics (250 x 250m). Hardcopy maps were digitized including; surficial geology, lineaments, and known mineral occurrences. The bedrock geology was input from as a scanned raster image. Finally a regional stream sediment geochemistry survey ASCII file was also integrated.

With these data sets residing in the SPANS GIS in a common projection and format the GIS functionality can now be used to examine spatial relationships. Both reports and maps can be generated from various modelling techniques available in the system.

Introduction

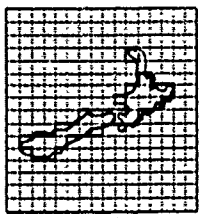
As with many disciplines, geology involves large volumes of data often of variable source and format. As a result of this inconsistency, data is often not used to it's fullest potential. In order to make the entire range of data accessible for analysis it must be integrated into a common reference and format. The TYDAC, SPANS Geographic Information System (GIS) was the tool used to integrate the various data sets and determine their inter-relationships.

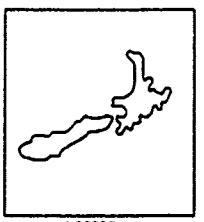
In 1987 the Geological Survey of Canada (GSC) contracted the CANMAP Research Institute at the Nova Scotia College of Geographic Sciences (COGS) to create a multi-layer geo-database for the Antigonish Highlands, Nova Scotia. This paper outlines the integration procedure used to establish the digital database. Using a GIS as the integrating tool also allows the use of analytical and modelling techniques on the database once it has been established. The GSC Mineral Resource Division are now using this approach and exploring new methodologies for data integration and mineral potential modelling.

The project involved data transfer between mini and micro-computer based image analysis and geographic information systems. The hardware platforms and software include: DIPIX, ARIES II (VAX 11/785), PCI, EASI/PACE (386 micro-computer), ESRI, ARC/INFO (PRIME 9650) and the TYDAC, SPANS (386 micro-computer). The data involved in the project as listed in Table 1 includes: Landsat MSS and TM imagery, airborne vertical gradient magnetics and radiometrics, digitized bedrock and surficial geology maps, lineaments, mineral occurrences, and stream sediment geochemistry.

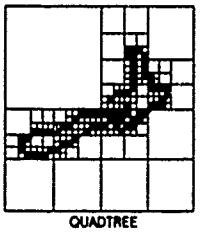
Data Structures

The systems employed in the project use three different data structures: raster, vector and quadtrees. The raster structure is used by the image analysis systems (DIPIX and PCI). Imagery is represented by a grid of equal size picture elements (pixels) of which each is assigned a value. The vector structure is used by the ESRI GIS. Thematic information is represented by a series of line segments, arcs and nodes, which define the polygon boundaries. The quadtree structure is used by the SPANS GIS. The information is stored in a raster like structure with a variable pixel size. In homogeneous areas the information is stored in large pixels, near the polygon boundary the pixel size is reduced to better define the boundary conditions. Quadtrees greatly reduce file size and hence reduce operating times. The efficiency of raster overlay procedures is maintained, and the accuracy of vector boundary placement is not compromised.





File Reduction Factor



Preparation of Raster Data

The satellite MSS and TM imagery was supplied by CCRS on tape and was loaded into the ARIES system.

Radiometric data were supplied by the GSC in PCI, UNIDISK format on floppy diskette. This was transferred to the VAX mini-computer, then converted to an ARIES file using locally developed ARIES task from COGS.

The magnetics were supplied by the GSC as standard VMS files. The data was stored as real values between -10 and +10 due to the gradient nature of the data. However, the scan lines were oriented vertically rather than the conventional horizontal style. The file was rotated through 90 degrees and converted to 8 bit (0-255) data by running a GSC supplied fortran program. The resulting VAX file was then converted to an ARIES file using the same COGS developed task as with radiometrics.

The satellite and geophysical imagery were then co-registered and resampled to a common 50m pixel size for a UTM projection. Enhancements and classifications were also carried out on the ARIES. The integrated raster data was then exported from the image analysis system by converting it to a VMS data file consisting solely of data, i.e. no header or trailer information. This data was then transferred from the VAX to the micro-computer

Vector Data

The vector data digitized in ARC/INFO consisted of: surficial geology, lineaments interpreted from satellite and magnetic imagery, and known mineral occurrences. This was exported using the U.S. Geological Survey Standard Digital Line Graph 3.0 (DLG) format and transferred from the PRIME to the micro-computer. The polygon attribute table (PAT) was also exported as a Standard Data File (SDF).

The Integrated Database: Spans Input

The raster imagery transferred from the VAX uses the same storage format as a SPANS raster file, left to right and top to bottom. An ASCII header file is used to spatially position the raster and define the x, y extent in terms of lines and pixels as well as meters. The raster files can be displayed and statistics generated for classification purposes. These raster files were then converted into the quadtree structure and maps built. The raster-quadtree conversion typically produces a file size reduction on the order of 5 times, while vector conversions typically reduces the file size by a factor of about 10 times.

The raster scanned bedrock geology was converted to a quadtree map directly, spatial positioning and projection information were supplied in the file.

The vector DLG files were read into the system through the SPANS vector interface. The surficial geology vectors were converted into a quadtree map and reclassified based on the polygon attribute table file. The lineaments and known mineral occurrences were imported as standard SPANS vectors and points, respectively. The UTM referenced stream sediment geochemistry file ASCII was directly imported into the system and stored as a SPANS point file.

The rather complex procedure used in this data integration methodology has since been simplified. Three suites of programs are required:

- SPANS (all modules)
- ARIES (or alternative Image Analysis System, mini or micro-computer based); and
- GSC data conversion programs

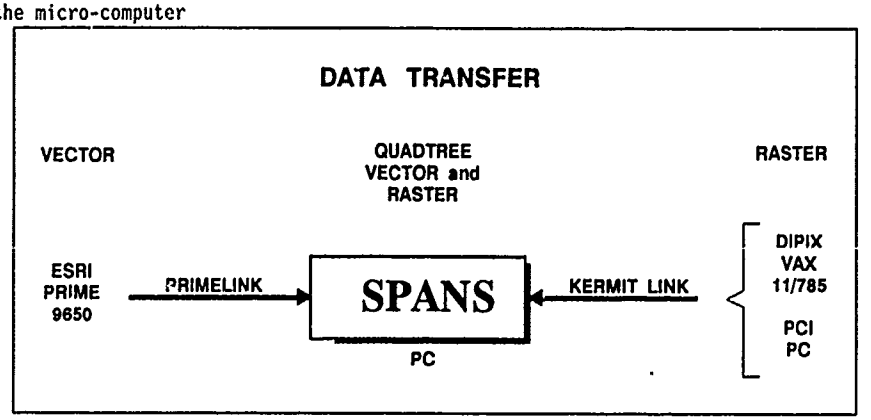


Table 1 ORIGINAL DATA SETS

DATA TYPE	STORAGE	RESOLUTION
Landsat MSS Multi Spectral Scanner	6 Bit Resampled to 8 0 - 255 value range	79 x 79 m Pixel Bands 4,5,6,7
Landsat TM Thematic Mapper	8 Bit 0 - 255 value range	30 by 30 m Pixel Bands 1,2,3,4,5,7
Airborne Vertical Gradient Magnetics	4 Byte Real -10 to +10 value range	125 by 125 m Pixel One Band
Airborne Radiometrics	4 Byte Real	250 by 250 m Pixel Total counts eU, eTH, %K and U/Th, U/K, Th/K
Surficial Geology	Vector	Digitized in ARC/INFO (arcs, nodes)
Bedrock Geology	Raster Scanned	Obtained from the Canada Land Data System
Stream Geochemistry	ASCII	Point data UTM reference element assay values

Advice on how best to proceed is available from the GSC, COGS, or the private companies mentioned in the text.

GIS FUNCTIONALITY

With the final geo-database residing on SPANS the analytical power of the system could be used. Along with the existing quadtree maps; surficial geology, bedrock geology, satellite and geophysical maps, new maps were created from the lineament vectors and point data. Proximity mapping was accomplished for the lineaments by generating corridors out from the vectors. Similarly cocentric circular zones were created around the known mineral occurrences. The stream geochemistry survey data was converted into a thematic trend surface for each element.

With the input maps completely established, analysis and modelling of the data can be used to determine relationships between the data. Various area analysis techniques may be used, on single maps as well as for two map correlations. Several overlay operations are available, for example, up to 14 maps can be superimposed to produce a unique output map. The maps and attributes of each can be weighted to produce the desired areas of interest. The system also employs its own modelling language which features preprogrammed mathematical functions (e.g. trigonometric, logarithmic, etc.) and the ability to express numerous complex relationships. The user can also query several maps in the database simultaneously, as well as large point files.

REVIEW OF SATELLITE APPLICATIONS TO FISHERIES

R. Michael Laurs

National Oceanic & Atmospheric Administration
National Marine Fisheries Service
P.O. Box 271
La Jolla, CA 92038 USA

ABSTRACT

Satellite oceanic remote sensing applications to fisheries research and fish harvesting in the U.S. are reviewed. Specific case studies are discussed to show how satellite data have been used in fisheries that are marketed as different types of fishery products. These include albacore tuna which is marketed primarily as a canned product, northern anchovy which are reduced for meal and industrial products, Pacific salmon species marketed mostly as fresh and fresh-frozen table fish, and Gulf of Mexico shrimp which are marketed mostly fresh.

INTRODUCTION

Satellite oceanic remote sensing can be a very powerful tool when used in fisheries research and in harvesting marine resources. The successes gained in experiences using satellite data in fisheries investigations and in harvesting thus far, indicate there is immense potential for expanding its utilization. Satellite remote sensing can supply environmental information, which can lead to improvements needed to ensure the effective management of marine fishery resources. It also, can provide information to detect potentially productive fishing grounds and to make improved weather and sea condition forecasts for the safety of fishermen and their vessels and equipment.

Variations in ocean conditions play crucial roles in causing natural fluctuations of fisheries resources and their vulnerability to catch. Comprehensive information on the changing ocean, rather than on average ocean conditions, is needed to effectively manage many stocks of living marine resources and to efficiently harvest them.

The strength of satellite remote sensing lies in the ability of spacecraft sensors to monitor ocean variability with the combined advantages of large scale synopticity, high spatial resolution and frequent repeatability. Each of these characteristics is prerequisite to understanding the effects of ocean variability on fishery resources.

Despite its strengths, satellite remote sensing has weaknesses, chief of which is the capability of sensors to measure only the surface skin or uppermost layer of the ocean. In addition, most present sensors provide useful data only in cloud- free areas. The application of satellite remote sensing to fisheries has also been hampered by difficulties in obtaining satellite data, the lack of affordable systems to process it, and by a shortage of trained people needed to conduct research and development and education of fisheries users. However, recent advances in the capabilities of data-processing systems and improved training opportunities in satellite oceanography, offer promise for the future.

In this manuscript, after a brief over-view of satellite data used in fisheries, experiences are summarized from marine fisheries investigations which have utilized satellite data. Specific examples have been selected to show how satellite data have been used in fisheries that are marketed as different types of fishery products including: albacore tuna which are marketed primarily canned, chinook and coho salmon which are sold fresh, northern anchovy which are reduced for meal and industrial products, and Gulf of Mexico shrimp which have several types of markets.

SATELLITE DATA USED IN FISHERIES

Satellite measurements that have been used in fisheries applications are: ocean temperature and color and ocean wind stress. Most fisheries research studies and operational uses of satellite sea surface temperature (SST) measurements have utilized data from thermal infrared sensors onboard the TIROS polar-orbiting meteorological satellites. Some have also used infrared data from geostationary satellites. However, only a very limited number of attempts have been made to apply ocean temperature measurements made by microwave instruments. The Coastal Zone Color Scanner (CZCS) on the Nimbus-7 satellite provided color measurements from space from early 1979 to about mid-1985. The excellent ocean color data from the CZCS have probably been more valuable for fisheries uses than any other ocean measurements made from space. There has been very limited success in using Landsat color imagery in fisheries studies. The scatterometer (SASS) aboard the Seasat-A satellite furnished data which demonstrated the importance to fisheries of high resolution

surface wind stress measurements from space. Satellite wind stress measurements can be used to calculate ocean surface layer transport, which controls the distribution of larval stages and the subsequent recruitment and harvests of many marine fishes and shrimps. Satellite measurements of ocean winds can also be extremely valuable in the detection and forecast of weather and sea conditions hazardous to safety at sea.

EXAMPLES OF SATELLITE APPLICATIONS TO FISHERIES

Albacore Tuna - A Canned Product Fishery

Albacore are highly mobile tuna which are widely distributed throughout the world's oceans. In the North Pacific Ocean, they migrate seasonally into waters off the coast of North America during July through October, where they support important U.S. commercial and recreational fisheries. Most of the albacore caught by commercial fisheries is consumed as a canned product. The distribution, availability and vulnerability of albacore off the west coast of the U.S. have been found to be related to oceanic fronts seen in AVHRR infrared and CZCS imagery (Laurs et al., 1984; Jurick, 1985; and Svejkovsky, 1987).

An investigation employing AVHRR infrared temperature and CZCS color imagery and concurrent daily commercial fishing catch records clearly showed that satellite temperature and color data can be used to define the environmental limits of the spatial distribution of fishable aggregations of albacore, and can do so more effectively than ship or aircraft data (Laurs et al., 1984). This study convincingly demonstrated that the distribution and availability of albacore off the west coast of California are related to oceanic fronts associated with coastal upwelling. Commercially fishable aggregations of albacore are found in warm, blue oceanic waters near temperature and color fronts that mark the seaward edge of waters which had been upwelled near the coast. Relatively intense fronts are favored and shoreward intrusions of warm, clear oceanic water are particularly favorable sites for albacore concentration. This study also found that during late summer, commercial concentrations of albacore several hundreds of miles offshore, were associated with oceanic boundaries noted by color fronts detectable from satellite, but without temperature gradients. The color boundaries probably distinguished the North Pacific Subtropical Front. In another study, the distribution of albacore in winter was linked to SST fronts observed in AVHRR imagery, which probably to mark the outer boundary of the California Current. (Laurs et al., 1981).

Possible mechanisms responsible for the aggregations of albacore in the vicinity of upwelling boundaries is being investigated using data from acoustic tracking of free-swimming albacore and concurrently collected oceanographic data from ships and AVHRR and CZCS data from satellites. The results suggest that albacore aggregate on the warm, clear oceanic side of coastal upwelling boundaries and avoid areas of higher productivity and forage density in upwelled waters, because they are not able to detect prey in the turbid, upwelled water. (Laurs, in prep.).

U.S. albacore fishermen make extensive use of satellite-derived fishery advisory products for determining favorable locations to fish and for safety. SST analyses based on satellite data have been routinely available for nearshore coastal waters off the U.S. west coast since 1975 (Breaker and Jurick, 1975). The form, degree of detail, and methods of dissemination of these analyses, which are in chart form, have varied over the years (Breaker, 1981). Presently, they consist of large-scale thermal boundaries which are derived from AVHRR imagery, and isotherms which are based mostly on in situ observations from ships and buoys. The charts are prepared twice a week and are distributed to fishermen by radio facsimile and by mail.

An AVHRR infrared sea surface temperature analysis for west coast waters, with considerable more detail is available to fishermen on a subscription basis from Ocean Imaging Company in San Diego, California. The satellite derived thermal analysis charts are distributed to fishermen at sea by radio facsimile, or photographs with false-color are disseminated by express mail or hand delivery. An increasing number of albacore and other fishermen subscribe to these ocean products because they are specially tailored to meet their needs, including high spatial resolution of surface boundary features.

To obtain satellite imagery when operating in distant water locations, some albacore fishermen have purchased low-cost systems designed for direct reception, processing and display of satellite imagery on board ship. These systems receive visual and infrared imagery with a 4 km pixel size by Automatic Picture Transmission (APT) signals directly from polar-orbiting satellites. The visual imagery is used to help avoid areas of hazardous weather, and the infrared imagery is being used to interpret SST patterns for more efficient track and fishing ground selection, in cloud-free areas.

During the early 1980's and continuing until the CZCS sensor failed, ocean color boundary charts based on CZCS imagery were utilized by U.S. west coast albacore to locate potentially favorable locations to fish (Montgomery et al., 1986). These charts were prepared and disseminated via radio facsimile in a program funded by the Jet Propulsion Laboratory and involving Scripps Institution of Oceanography and the U.S. National Marine Fisheries Service. According to surveys of fishermen, these satellite-derived ocean products were remarkably useful in locating favorable areas to fish for albacore, and could save them as much as 50% in search time.

Oregon Coho and Chinook Salmons - A Fresh Market Fishery

Chinook and coho salmon support lucrative markets as fresh fish products. An investigation is underway to evaluate the use of satellite imagery to determine the optimal time of release of salmon smolts from hatcheries on the Columbia River to correspond with ocean conditions favorable for their survival. Increasing the survival and subsequent contribution of salmon released from hatcheries to commercial and recreational fisheries can have significant economic benefit.

Up to 90% of the salmon caught in the waters off the Columbia River are released from hatcheries on the Columbia River. About 60% of the salmon caught in other areas off the Pacific Northwest are from salmon released from hatcheries on other coastal rivers and streams. While hatchery produced salmon contribute most of the fish which are harvested, the percentage of fish released that are caught is low, e.g. only about 2% for Columbia River hatcheries. About 98% of the salmon that are released suffer mortality. A major part of the mortality is believed to occur in the ocean soon after the smolts arrive there, subsequent to their release from the hatcheries (Pearcy, 1984).

The research is directed toward testing the hypothesis that the survival of young salmon released from Columbia River hatcheries is related to variations in characteristics of the Columbia River plume and its interactions with other oceanic processes, notably coastal upwelling. The goal of the research is to ascertain if satellite imagery can be used to determine when ocean conditions are favorable for young salmon, so that the release of the smolts from hatcheries may be timed for optimal survival. Even modest increases in survival could result in substantial increases in salmon available for harvest, and have significant economic value.

Findings thus far have demonstrated that ocean temperature and phytoplankton pigment images derived from AVHRR and CZCS data, respectively, show that the Columbia River plume has remarkable variability in coastal waters off the Pacific Northwest (Fiedler and Laurs, in review). The orientation, shape, intensity and relative temperature of the plume vary in response to coastal winds and wind-driven surface currents.

The use of satellite-derived SST distribution and frontal boundary analyses by commercial fishermen to locate potentially fruitful locations to catch fish is a further application of satellite remote sensing to Pacific Northwest salmon resources.

Northern Anchovy - A Fishery Where Catches are Reduced For Oil and Industrial Uses

Northern anchovy, *Engraulis mordax*, is a California Current fish targeted for an oil and meal reduction fishery. Like most fish species harvested for their oil and meal, the anchovy is relatively short-lived and spends its early life history stages, part of which are planktonic, in the upper mixed layer. Its spawning, early survival and recruitment to the fishery are believed to be largely controlled by ocean conditions (Lasker, 1978).

Satellite imagery has been used to investigate ocean processes in relation to northern anchovy spawning. In these studies, which were conducted for several years during anchovy spawning periods, satellite data were collected on a daily basis coincident with fine-grid oceanographic ship observations. Shipboard observations included sampling of anchovy eggs, larvae and adults as well as physical oceanographic measurements. The objectives of the investigations were to relate variations in mesoscale SST distributions (Lasker et al., 1981) and phytoplankton pigment and SST distributions (Fiedler, 1983) with anchovy spawning, and to identify and delineate ocean processes that might be important to the survival of fish eggs and larvae. Based on satellite imagery and confirmed by shipboard observations during the studies, which were conducted in the Southern California Bight, there were were distinct temperature and pigment regimes in the general geographic region where anchovy spawning normally takes place.

Lasker et al. (1981) found that anchovy avoid recently upwelled water and that the areal extent of upwelled water may be mapped using infrared satellite imagery. They concluded that the anchovy habitat for spawning in the northern part of the Southern California Bight could be defined using AVHRR satellite data. Fiedler (1983) corroborated Lasker et al. (1981) conclusion and also found that the limits of the spawning habitat in the southern part of the Bight could be differentiated using CZCS imagery. The spatial distribution of northern anchovy spawning can thus be defined by mesoscale patterns in satellite temperature and phytoplankton pigment images (Lasker et al, 1981 and Fiedler, 1983). While neither parameter alone is sufficient, both together may define the spatial distributions nearly completely. In general, the northern extent of spawning in the Southern California Bight and the offshore extent north of Santa Catalina Island are limited by cold, upwelled water advected south of Point Conception. Spawning to the south is limited by low phytoplankton pigment levels in oceanic water 20-100 km offshore rather than by temperature. However, these factors do not directly determine spawning success, but rather are indicators of oceanic conditions and processes that determine the spawning habitat.

Fiedler (1984) also demonstrated the effectiveness of using satellite AVHRR and CZCS imagery to monitor shifts in anchovy spawning habitat off California associated with the 1982-1983 El Niño warm-water conditions.

Shrimp In The Gulf Of Mexico - Mostly a Fresh Product Fishery

Surface layer transport processes play an critical role in the life cycle of shrimp. Dispersal mechanisms control the distribution of early life stages of shrimp and the transport of developmental stages from offshore waters to coastal estuarine areas in the Gulf of Mexico. Thus the recruitment and future harvest of shrimp resources are determined, in large, part by ocean surface layer transport.

Brucks et al., in prep., also see Laurs and Brucks (1985), conducted a case study to investigate whether high-resolution measurements of wind stress by the Seasat-A scatterometer (SASS) could improve estimates of surface layer and larval shrimp transport based on traditional geophysical models. The SASS- measured wind-stress field was used to calculate surface currents by the standard homogeneous, steady-state Ekman solution. The SASS approach reduces the number of assumptions involved in

the calculation of the wind-stress vector and presumably provides more accurate estimates of surface circulation. Striking spatial and temporal variability in wind-drift pattern and regions of convergence and divergence were apparent in the Gulf of Mexico using the SASS wind-stress data to calculate surface-layer transport processes. The investigators used surface-layer transport estimates calculated from the SASS data to define potential offshore areas where shrimp spawn material and very early-life stages could be rich in the plankton and coastal areas where larval and post-larval planktonic stages could accumulate prior to settling. They found that processes conducive for accumulation in the offshore regime mainly in the west central Gulf and some projections to the south, west and north. The predicted coastal areas were in close proximity to shrimp nursery grounds in Louisiana, Texas and Florida.

Brucks et al. concluded that synoptic and repetitive direct measurements of wind stress by the SASS provided a significantly enhanced capability to determine variability in wind-driven events that strongly influence shrimp recruitment processes in the Gulf of Mexico.

LITERATURE CITED

1. Brucks, J.T., T.D. Leming, and S.B. Buckett, Jr., "A model investigation using high resolution SASS wind stress measurements to derive wind surface layer transport properties in the Gulf of Mexico", In prep.

2. Breaker, L.C., "The application of satellite remote sensing to west coast fisheries", J. Mar. Tech. Soc. Vol.15, pp32-49, 1981.

3. Breaker, L.C. and F.A. Jurick., "Providing near real-time sea surface temperatures to the California north coast fishing fleet", In: Ocean 75 (Combined meeting of 1975 IEEE Conference on Engineering in the Ocean Environment and 11th Annual Meeting of the Marine Technology Society), San Diego, Calif., N.Y. IEEE, 952pp, 1975.

4. Fiedler, P.C., "Satellite remote sensing of the habitat of spawning anchovy in the Southern California Bight", CalCOFI Rep. No. 24, pp202-209, 1983.

5. Fiedler, P.C., "Some effects of El Niño 1983 on the northern anchovy", CalCOFI Rep. No.25, pp53-58, 1984.

6. Fiedler, P.C. and R.M. Laurs, "Variability of the Columbia River plume observed in visible and infrared satellite imagery", Remote Sensing Environ., In rev.

7. Jurick, F.A., "A comparison between North Pacific commercial albacore *(Thunnus alalunga)* catch and the satellite derived West Coast thermal boundary analysis", M.S. Thesis, Humboldt State University, Arcata, Calif., 79pp, 1985.

8. Lasker, R., "The relation between oceanographic conditions and larval anchovy food in the California Current: identification of factors contributing to recruitment failure",Rapp. P.-V. des Reun. Cons. int. Explor. Mer. Vol.173, pp212-230, 1978.

9. Lasker, R., J. Pelaez and R.M. Laurs, "The use of satellite infrared imagery for describing ocean processes in relation to spawning of northern anchovy *(Engraulis mordax)*", Remote Sensing Environ., Vol.11, pp439-453, 1981.

10. Laurs, R.M., "The role of water clarity in albacore aggregations in upwelling frontal areas", In prep.

11. Laurs, R.M. and J. Brucks, "Living marine resouces applications", In Satellite Oceanic Remote Sensing, Ed. B. Saltzman, Adv. Geophys., Vol.27, pp419-452, 1985.

12. Laurs, R.M., R.J. Lynn, R. Nishimoto, and R. Dotson, "Albacore trolling and longline exploration in eastern North Pacific water during mid-winter 1981", NOAA Tech. Memor., NMFS-SWFC-10, 52pp, 1981.

13. Laurs, R.M., P.C. Fiedler, and D. Montgomery,"Albacore tuna catch distributions relative to environmental features observed from satellites", Deep Sea Res., Vol.31, pp1085-1099, 1984.

14. Montgomery, D.R., R.E. Wittenberg-Fay, and R.W. Austin, "The applications of satellite-derived ocean color products to commercial fishing operations", MTS Jour. Vol.20, No.2, 72-86, 1986.

15. Pearcy, W.G., "The influence of ocean conditions on the production of salmonids in the North Pacific, a workshop", Oregon State Univ., Sea Grant College Program, ORESU-W-83-001, Corvallis, Oregon, 327pp, 1984.

16. Svejkovsky, J., "Remotely Sensed Ocean Features and Their Relation to Fish Distributions", In: Organisms As Indicators, Eds. G. Kleppel and D.F. Soule, Springer-Verlag Publishers, N.Y., pp179-197, 1987.

REMOTE SENSING AND MODERN FISHING TECHNOLOGY - ITS IMPACT ON CONSERVATION AND PROTECTION OF THE FISHERIES RESOURCE

J.C. Davis, Regional Director of Science, Department of Fisheries and Oceans, Institute of Ocean Sciences, P.O. 6000, Sidney, B.C. V8L 4B2 (FAX 604 356-6479)
TELEX 0636700764

World population is increasing steadily while the food supply becomes taxed and mankind turns increasingly to the oceans for food and livelihood. The fishery has changed dramatically in the past few decades with powerful instrumentation now available to fishery fleets which readily allow location, identification and efficient harvesting of the resource. In addition, better knowledge of oceanic features and the distribution of resource species in relationship to those features increases harvesting effectiveness. This paper focuses on the impact of sophisticated equipment and remote sensing technology on aquatic resources and their harvesting, management and conservation. Particular emphasis is placed on the British Columbia salmon fishery which suffers particular problems related to excessive harvesting capacity.

COMPARISON OF SHIP AND AERIAL SURVEYS OF SARDINE SCHOOL

Ichiro Hara

Seikai Regional Fisheries Research Laboratory
Kokubu-machi Nagasaki, 850 Japan
phone 0958-22-8158 FAX 0958-21-4494

Field tests were made to compare estimates of sardine schools resulting from simultaneous ship and aerial observations. The aerial/ship observations were performed around the areas where big catches of sardine have been taken. In the surveyed area, the sardine schools can be sighted and counted easily, making the direct counting very advantageous. The schools are only distinguished from their dark color against lighter background of ocean.

A comparison of density estimated with the two types of surveys, namely ship- and air-borne observation are discussed here as a sea truth. In case of air-borne observation schools counted were only 16.5% of ship-borne observation. This may be caused by the difference of the visibility between two ways. I think the direct detection of schools and counting may be used widely and expect to be able to count schools from space in future.

KEY WORDS: aerial survey, sardine, line transect, sonar.

INTRODUCTION

Development of rapid and direct method for assessing stock abundance is one of the most important subjects of fish population studies. The author has developed a method to count fish schools directly and assess fish stocks on the basis of aerial/ship surveys (Hara,1983,1986). The day-time aerial/ship observations were performed around the areas where big catches of sardine have been taken (Fig.1). The sardine catch in the area off southeastern coast of Hokkaido, which is of concern in this paper, exceeded one million tons in during the July-October fishery. In the surveyed area, the sardine schools distributed vertically in shallow areas (Hara,1984), and the schools are sighted and counted easily from the air. The schools are only distinguished from their dark color against the lighter background of the ocean. The shapes of the schools are usually elongate and narrow, but there are other shapes as well. The shapes of the schools are classified into three types, such as oval,

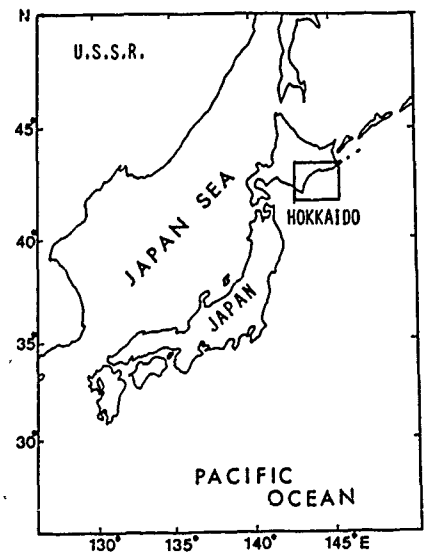


Fig.1 Locations of the area surveyed.

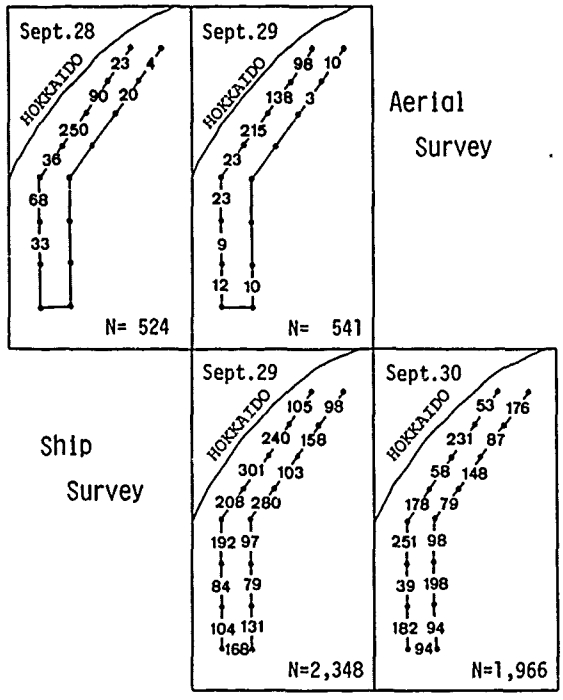


Fig.2 Transect lines of the aerial and ship surveys totaling 142 n.m. The upper part shows the result of aerial surveys and the lower part shows the result of scanning sonar surveys. Numbers show the number of sardine schools observed.

crescent and elongate (Hara,1985). The typical horizontal dimensions of the large elongate or crescent-like schools are 10m in width and 100-200m in length. Oval-like schools has 20-30m in maximum axis. In this survey area, the maximum sardine school dimensions ranged from 20-200m. Moreover, we can easily estimate a moving direction of schools by observing the changes in shape and color (Hara,1985).

MATERIALS and METHODS

Sea truth data were obtained to compare estimates of sardine school density resulting from simultaneous ship and aerial observations on Sept. 29, 1987. Furthermore duplicating surveys were conducted on the same transect line by aircraft or ship before or after sea truthing (Fig.2). For all fish schools sighted within a given distance from the observer, school density is estimated (Eberhardt,1978 and Gates,1979), as follows,

$$D = n/2LW,$$

Table 1. Comparison of densities (schools/n.m.^{2}) of sardine schools derived from sonar and aerial counts over the same transects (142n.m.).

Time	No. of schools	Distance (miles)	Density
Aircraft			
Sept.28	524	142	4.1
29	541	142	4.2
Sonar			
Sept.29	2,348	142	61.2
30	1,966	142	51.3

where D=number of schools per square nautical mile; n=number of schools sighted; L=length of the transect line; and W=the average distance under observation to the right or left of the transect line (half of the swath). Aerial counts were made from a Cessna U-206G flying at an altitude of 500m with a velocity of 110 Kn. The estimated width of the aerial counting strip was 839m (839 x 2m total width). During the visual observation from the air, schools which were in deeper layers could not be sighted. However, the schools could be counted from the ship using a scanning sonar or a vertical echo sounder. Counts on the sonar image were taken at a range of 250m (250 x 2m total width) and an angle of depression between 0 and 5 degrees.

RESULTS and DISCUSSION

As a result of the field test, sonar observers reported considerably higher mean densities than did aerial observers (Fig.2, Table 1). In order to examine in detail, the transect line (Fig.2) was separated into 4 transect segments by north-south line and a inshore-offshore line. In case of the air-borne observations, schools counted were only 16.5% of ship-borne observations. Northern part of inshore segment contained the highest number of counts. The counts of other three transect segments ranged from 0.6 to 3.5%, in comparison with northern part of inshore these values were extremely low. This may be caused by the difference in visibility in two ways.

Table 2 shows the vertical distribution of sardine schools measured by vertical echo sounders. The number of schools found in a duplicating survey carried out on the following day was low. That is to say, the number of schools observed was 388 schools on

Table 2. The vertical distribution of sardine schools measured by Mishio-maru No.5 using the echo sounder. The right column (Sept.29*) shows the distribution at the northern part of inshore segment where many sardine schools were observed.

Depth	Sept.29	Sept.30	Sept.29*
0- 2m			
2- 4m	34	4	9
4- 6m	139	81	41
6- 8m	58	37	14
8-10m	44	37	9
10-12m	28	24	7
12-14m	17	19	4
14-16m	16	23	3
16-18m	15	16	7
18-20m	9	9	1
20-22m	8	9	4
22-24m	5	7	4
26-28m	1	3	1
28-30m	1	2	0
30-	5	9	4
Total	388	287	113

September 29th and 287 on September 30th. A similar trend was found in sonar observations. The modes of vertical distribution are seen in the class 4-6m. Counting vertically downward from the surface of the sea, the 16.5% of the schools correctly determined belonged to the class 4-6m. It is assumed that all fish schools which were distributed vertically were detected by sonar. Calculations made with this assumption show that 16.5% of the underwater schools could be observed from the air. In other words, aerial observers detected fish schools within a depth of 4 to 6m from surface.

With aerial surveys, important factors to be considered are: depth of school, clarity of water, and brightness of daylight. The clarity of water decreased appreciably when moved towards inshore from offshore region. At the northern part of inshore segment described above, the optical transparency of water in this region is 4-5m usually (Hara,1986). The results of the fish detection scheme carried out consecutively for five years, the current year results show a close correlation with those of the previous year.

As shown in Table 2, the sardine schools are not vertically distributed in deep layers. It is not clearly determined whether it was the same group of schools observed by ship and the aircraft since there was a time lag between the ship and aircraft surveys. The influence of time lag between two levels of observations and the levels of water clarity, that might have on the results of these surveys is yet to be determined.

A sonar survey may be efficient in counting fish schools however, an aerial survey provides additional information such as shape, size of individual schools, the patterns of their movement and distribution. The main disadvantage of making a survey by ship is its slow speed, it takes a very much longer time to complete a survey, where as the same area can be covered by aircraft within a fraction of that time. At this stage, it can be said that the combination of acoustic and visual observations are important in direct stock assessments. Author believes that this method of direct detection of fish schools may be used widely in future with the employment of multiple data gathering platforms such as ships, aircraft and spacecraft.

REFERENCES

1. Eberhardt,L.L., "Transect methods for population studies", J. Wildl. Manage., Vol.42, No.1, pp1-31, 1978.
2. Gates,C.E., "Line transect and related issues", In : Sampling Biological Populations. Statistical Ecology Vol.5. Int. Co-op. Publ. House, pp71-123, 1979.
3. Hara,I.,"Estimation of fish density using the line intercept method", Nippon Suisan Gakkaishi, Vol.49, No.11,pp1619-1625, 1983.
4. Hara,I.,"Distribution and school size of Japanese sardine in the waters off the southeastern coast of Hokkaido on the basis of echo sounder surveys", Tokai Reg. Fish. Res. Lab. No. 113, pp63-74, 1984.
5. Hara,I.,"Shape and size of Japanese sardine school in the waters off the southeastern Hokkaido on the basis of acoustic and aerial surveys", Nippon Suisan Gakkaishi, Vol.51, No.1, pp41-46, 1985.
6. Hara,I.,"Moving direction of Japanese sardine school on the basis of aerial surveys", Nippon Suisan Gakkaishi, Vol.51, No.12, pp1939-1945, 1985.
7. Hara,I.,"Stock assessment of Japanese sardine in the waters off the southeast coast of Hokkaido using line transect method", Nippon Suisan Gakkaishi, Vol.52, No.1, pp69-73, 1986.

REMOTE SENSING OF FISH SCHOOLS: EARLY RESULTS FROM A DIGITAL IMAGING SPECTROMETER

Brian S. Nakashima[1], Gary A. Borstad[2], David A. Hill[2], and Randy C. Kerr[2]

[1]Department of Fisheries & Oceans, Science Branch
P.O. Box 5667, St. John's, Newfoundland, A1C 5X1
Tel. (709) 772-4925, Fax. (709) 772-2156

[2]G.A. Borstad Associates Ltd., Suite 100 Marine Technology Centre
9865 West Saanich Road, Sidney, British Columbia, V8L 3S1
Tel. (604) 656-5633, Fax. (604) 656-3646

Abstract

To examine annual fluctuations in capelin (Mallotus villosus) stocks on the east coast of Newfoundland, aerial surveys have been conducted since 1982. This method is time-consuming, labour-intensive, and restricted to light conditions suitable for colour photography over water. To overcome these disadvantages we are devising a digital imaging technique which relies on a newly developed imaging system, the Compact Airborne Spectrographic Imager (CASI). The advantages of the CASI over aerial photography are: 1) data can be collected during light conditions unfavorable for colour photography; 2) the digital data can be analyzed following each flight; 3) data are stored digitally in real time; and 4) stored data can be easily incorporated into a geographic information system. In this paper, we describe the CASI and its operation, show how it will be used to develop spectral signatures to differentiate fish schools from other features, and discuss the future operational role of CASI in monitoring fish stocks.

Keywords: digital imagery, spectrometer, spatial, spectral, CASI, capelin schools

Introduction

Capelin (Mallotus villosus), a small silvery pelagic fish species, inhabits Arctic and sub-Arctic waters in both the Atlantic and Pacific Oceans. Five stocks of capelin are commonly found on Canada's east coast, four of which migrate from offshore areas to coastal beaches to spawn during a four to six week period in June and July. Mature capelin segregate into schools according to sex. Males tend to school close to beach environments and females congregate in schools in deeper water. When females are ready to spawn, they move towards the beach through the male schools. These ripe females are the target of an intensive fishery employing purse seines, traps, and beach seines to supply a lucrative Japanese capelin roe market.

Estimation of spawning biomass is based on hydrographic surveys conducted prior to the actual spawning period. To verify the annual fluctuations in biomass predicted from acoustic estimates, a catch rate index and a school surface index are estimated on a routine basis (Nakashima, 1988). The latter is based on annual aerial photographic surveys of capelin schools conducted during the spawning period. Photographs are taken at an altitude of 457 m utilizing an aerial photo-mapping camera and colour negative film (Kodak Aerocolour Negative 2445) for limited water penetration. Experience has shown that photography must be restricted to sunny days when the sun angle is 20 to 50 degrees, the lower limit being too dark and the upper limit producing sun glint from the water surface. Following the surveys, prints are examined to count and measure the area of each capelin school (Nakashima, 1983, 1988). However, the aerial photographic method is time-consuming, labour-intensive, and limited to light conditions required for successful colour photography. In addition, the school surface area index is not available until several months after the survey and they cannot be converted readily to abundance estimates, a necessary step in the assessment process (Nakashima, 1983).

To overcome limitations associated with photographic techniques and to expedite analysis of data we consider herein the application of digital remote sensing with the Compact Airborne Spectrographic Imager (CASI) as a viable alternative. The CASI is a pushbroom sensor which operates in a spectral range of 423-946 nm. The instrument is compact and easily installed aboard small fixed-wing airplanes and helicopters. Further details on CASI are provided by (Borstad & Hill, Apr. 89, Borstad et al., July 89).

In this manuscript, we present spectral and spatial data of capelin schools collected by the CASI during its first developmental mission in July 1988. The results are the first step toward the development of a fully operational digital aerial survey of capelin schools.

Methods

In July 1988, survey tracks were flown in a Piper Aztec over the experimental area, Coley's Point in Conception Bay, Newfoundland. The aircraft was equipped with a Zeiss RMK photo-mapping camera with a 153 mm lens and with the first prototype of the CASI. Colour negative photographs and digital imagery were collected simultaneously at the test site at an altitude of 610 m.

The CASI is an imaging spectrometer which utilizes a two-dimensional (612 by 288) charge couple device (CCD) array to acquire both spatial and spectral data (Borstad & Hill, Apr. 89). A spatial image consists of a series of swaths across the direction of travel formed on 577 elements of the 612 on the CCD array. The remaining elements are used to obtain dark and electronic offset values. A diffraction grating is utilized to obtain spectral information across 288 elements in the along track dimension of the array. The spectral resolution of each element is 1.8 nm. During this experiment data was collected in two imaging modes; Spatial and Spectral.

In Spatial mode, or Imaging mode, the CASI is similar to other pushbroom imagers, however the spectral band widths, positions, and number are programmable during flight. These spectral bands can be as narrow as 1.8 nm or several hundred nanometers wide. For this flight the CASI used a 388 by 288 CCD array, the present version now uses the 612 by 288 array (Borstad et al., July 89). A total of 8 spectral bands were used during this project, with a spectral range from 545 to 757 nm.

In Spectral mode, or Multispectrometer mode, the CASI can be thought of as many independent spectrometers with each spectrometer having a resolution of 1.8 nm. The number and positioning of these spectrometers is programmable during flight. CASI also obtains a co-registered monochromatic high spatial resolution reference image, 'track recovery image', so that each spectrometer row can be accurately located (Fig. 1).

Results

Capelin schools can be identified from aerial photographs as dark patches with fluid edges. Identification and delineation of the schools is subjective and requires experienced interpretators. For present assessment purposes the area of the school, on the aerial photograph, is measured with a planimeter. Digital data on the other hand ,albeit more complex to interpret, should yield consistent objective results.

Early results show that the CASI has the spectral and radiometric sensitivity to detect small changes caused by capelin and can display these changes as images. At this time the array was uncalibrated so this has precluded analysis of the spectral data. Later data acquired with a calibrated array show that the track recovery image is an aid to the researcher in identifying the location of a given spectral signature (Borstad &

Hill, April 1989). By e ng these spectral signatures spatial ba: .. ctrum can be selected to optimally r erentiate between targets such as fish schools and other ocean features.

Discussion

The results of the July experiment demonstrate that the CASI can be employed to collect data from capelin schools. Radiance spectra from capelin schools compared to other features such as kelp, rock, water, etc. are different and can be targeted inflight for a monitoring survey such as the one proposed to replace the capelin aerial survey. In March 1989 the CASI was used successfully to image Pacific herring (Clupea pallasi) schools on the west coast of Canada (Borstad and Hill, 1989). The capability to detect fish schools of two different species has exciting potential for managing fish stocks which are amenable to this methodology. Besides performing a monitoring function, digital imagery of fish schools can be analyzed to examine the distribution of fish with respect to kelp beds, water features such as upwelling areas, and bottom bathymetry in a geographic information system. Airborne sensors such as the CASI can complement research from survey vessels and can be used to provide synoptic coverage of large areas in a time frame much shorter than could be achieved from research vessels.

In 1989 a study will be conducted to compare the results from a capelin monitoring programme using colour aerial photography and digital imagery from the CASI. The results from the 1988 experiment will permit data acquisition from CASI to focus on a few pertinent spectral bands to reduce the amount of data written to tape. In addition to real time examination of the digital data, a procedure to enumerate schools and calculate school areas following each flight will be implemented. This will significantly enhance the utility of the method for resource managers over aerial photography which now requires up to 2 months to obtain prints and a further 2 months for analysis of fish schools. Preliminary studies have also demonstrated that the CASI can detect schools under cloud and other light conditions lower than required for colour photography and when the sun angle is greater than 50 degrees. These features will make more effective use of survey time during a period when school behavior is dynamic and weather conditions can be quite variable.

Future development of the CASI will allow for the recording and correction of incident light levels along the flight path. This development will add to the CASI's versatility as a synoptic research instrument. The direct detection of fish schools by airborne sensors such as CASI will enable researchers to monitor school behavior with respect to oceanographic phenomena. The ability to assign spectral signatures to fish schools may lead to classification of schools using spectral data which raises the possibility that colour variation among schools may be associated with differences in density. An observed relationship between these two measurements would lead to conversion of school area into abundance estimates, a major goal in fisheries resource surveys.

Acknowledgements

The CASI was developed by Itres Research Ltd., Calgary, Alberta. The study was funded by the Department of Supply and Services (Unsolicited Proposal Fund) and the Department of Fisheries and Oceans.

References

Borstad, G.A. and D.A. Hill, "Using visible range imaging spectrometers to map ocean phenomena", Conference on Advanced Optical Instrumentation for Remote Sensing of the Earth's Surface from Space, International Congress on Optical Science and Engineering, Paris, France, April 24-29, p. 7, 1989.

Borstad, G.A., D.A. Hill and R.C. Kerr, "The Compact Airborne Spectrographic Imager (CASI): Flight and Laboratory Examples", IGARRS'89/12th. Canadian Symposium on Remote Sensing, Vancouver, B.C., July 9-14, p. 4, 1989.

Nakashima, B.S., "Aerial photography of capelin (Mallotus villosus) schools in the coastal waters of Newfoundland", Proc. RNRF Symp. Application of Remote Sensing to Resource Management, Seattle, Washington, U.S.A., May 22-27, pp. 655-660, 1983.

Nakashima, B.S., "Capelin schools surface area index for NAFO Div. 3L, 1982-87", Northwest Atlantic Fisheries Organization Research Document 88/11, p. 11, 1988.

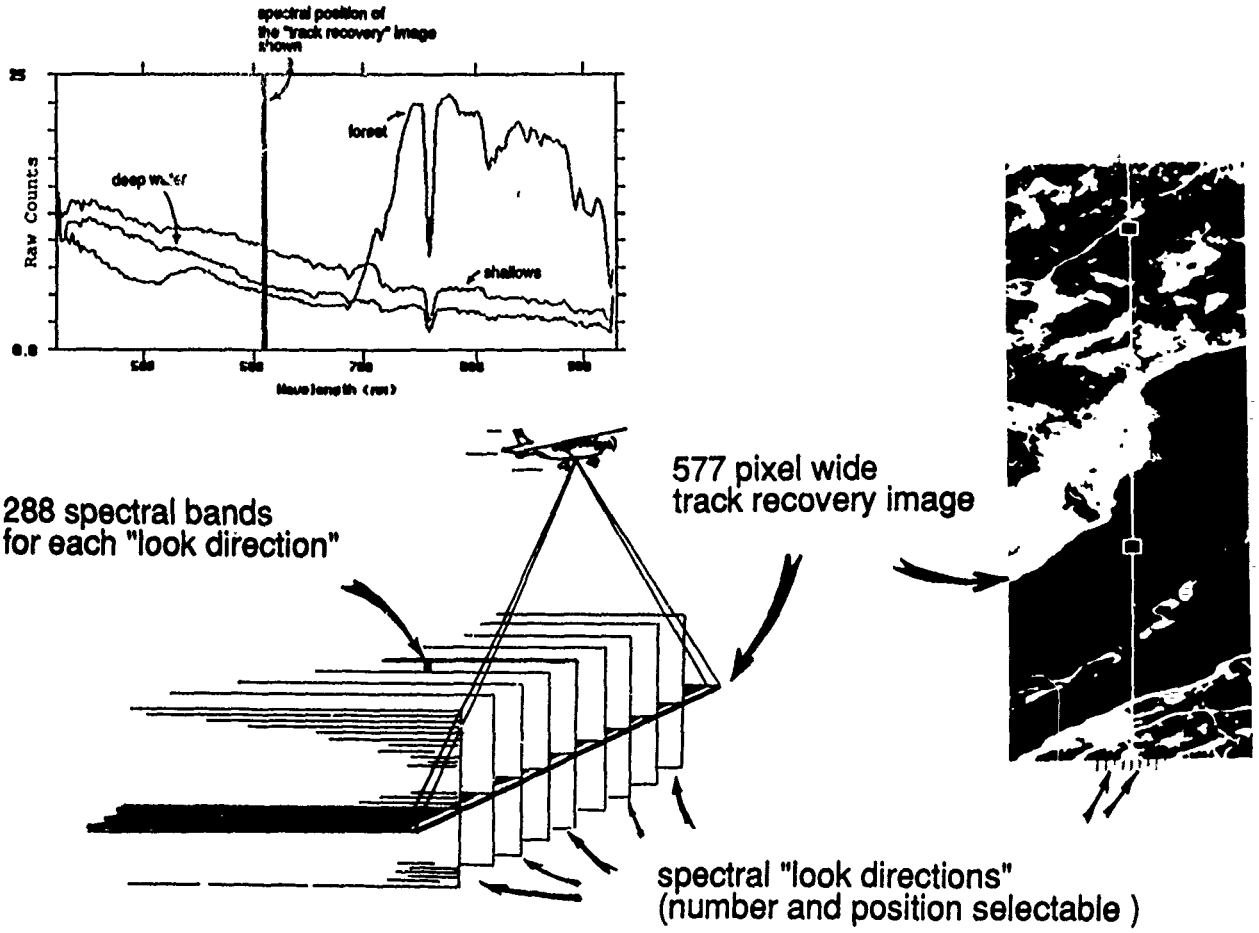


Figure 1: In CASI Spectral Mode, data is read from one or more spectral columns to provide time varying spectral data along the track of the aircraft, as well as from one row of the array to form a monochromatic high spatial resolution 'Track Recovery' image. The raw count spectra shown are derived from single pixels at the positions indicated in the Track Recovery image (Borstad & Hill, 1989).

THE TOXIC *CHRYSOCHROMULINA POLYLEPIS* BLOOM IN SCANDINAVIAN WATERS - MAY/JUNE 1988

Ola M. Johannessen, Lasse H. Pettersson, Johnny A. Johannessen, Peter M. Haugan, Tor I. Olaussen, and Kjell Kloster

Nansen Remote Sensing Center
Edvard Griegsvei 3a
N-5037 Solheimsvik, Norway
Telephone: +47 -5-297288

ABSTRACT

During the late spring of 1988 an extensive bloom of the toxic algae, *Chrysochromulina polylepis*, occurred in the Skagerrak region, which influenced most life in the upper 30 meter of the ocean. The algal front was transported northward with the Norwegian Coastal Current along the coast of southern Norway, where it became a severe threat to the major part of the Norwegian seafarming industry. Based on the need of guidance to the seafarming community and for public information about the movement and effect of this algal front, an expert team of biologists, oceanographers and remote sensing specialists were established among scientists from the Institute of Marine Research, the University of Bergen and the Nansen Remote Sensing Center in Bergen. Data were obtained from research vessels, aircrafts, satellite and research ocean model were made available to be used in the forecast of the growth and movement of the algae front. Remote sensing data of sea surface temperature from the operational US NOAA satellites demonstrated its capability, in combination with the *in situ* biological, chemical and physical observations, in following the movement of the algal front, which in this case appeared consistent with the temperature front, prior to the culmination of the bloom.

Key words: PLANKTON BLOOM, AQUACULTURE, REMOTE SENSING, OCEAN MODELLING

INTRODUCTION

During spring, increasing light intensity and available inorganic nutrients create favorable conditions for intensive growth of algae (blooms), which can last as long as nutrient supplies are replenished. Diatom spring blooms are a normal and necessary part of the productivity of the ocean. Discoloration of the sea due to the extreme abundance of planktonic algae is a phenomenon frequently observed.

Since 1966 a number of toxic blooms of dinoflagellate algae have been reported along the southern coast of Norway. Most were caused by the species *Gyrodinium aureolum*, producing toxins which can cause fish kills, and by *Prorocentrum minimum* producing toxins which may accumulate in filter feeding animals (e.g. mussels) without causing harm to these animals but rendering them toxic to consumers (Tangen, 1983).

Recently the recreational interest in coastal areas has increased markedly, causing a heightened public awareness of unusual blooms. With the advent of aquaculture, toxic blooms, causing shellfish poisoning or loss of valuable cultured fish, achieve the status of media events.

In 1988 the alga *Chrysochromulina polylepis* had a massive and unpredicted bloom during May–June in the Skagerrak–Kattegat area, where it occasionally outgrew all other algae, (Dundas et al., 1989).

Four weeks after its outbreak in eastern Skagerrak the bloom covered major parts of the Kattegat and the Skagerrak causing kills of both wild and caged fish. The bloom spread northwards with the Norwegian Coastal Current (NCC) along the western coast of Norway to about 60°N.

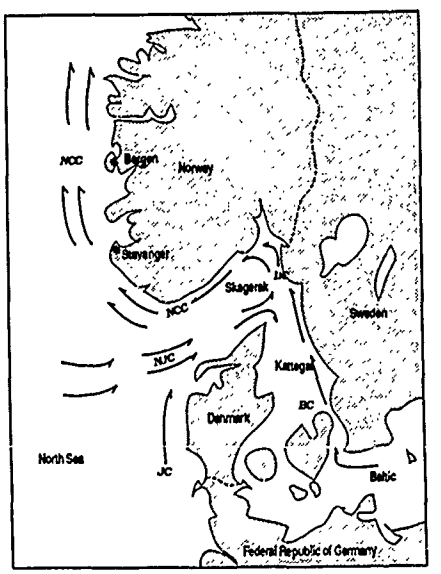


Fig. 1: Surface circulation in the North Sea, Skagerrak and Kattegat area. The Baltic Current: BC. The North Jutland Current: NJC. The Jutland Current: JC. The Norwegian Coastal Current: NCC.

In spite of the high concentrations of cells (up to 70–80 million cells per liter) the bloom was not very striking visually, partly because maximal algal populations often were found at some depth, with less dense populations at the surface, causing weaker discoloration. The bloom was accordingly mainly noticed by its lethal effect on caged fish in fish farms along the coast. About 500 tons of caged fish with a market value of approximately US$ 5 million were lost along the southern coast of Norway before precautions were taken.

On the west coast of Norway, some 200 sea farms containing fish with a value of approximately US$ 200 million were evacuated during the bloom. Seafarm cages were relocated by towing them from the bloom–exposed coastal region, to relative safety in the brackish waters of the inner parts of the fjords.

The surface water of the Kattegat is influenced by the outflowing brackish water from The Baltic Sea. In the eastern Skagerrak, this water mixes with water from the central and southern North Sea entering via the North Jutland Current (NJC), and flows along the entire Norwegian coastline from the eastern Skagerrak into the Barents Sea, forming the Norwegian Coastal Current (NCC), (Fig. 1). Remote sensing studies (e.g. Johannessen et al., 1983) have concluded that the NCC is characterized by a complex pattern of shifting meanders and eddy structures (Johannessen et al., 1989b).

The surface water of the central Skagerrak originates mainly from North Sea water overriding Atlantic water brought in via the Norwegian Trench. The anticlockwise circulation in the area favours an upwelling of deeper, nutrient rich water in the central part of Skagerrak (Svanson, 1975), which is of of great importance for the fertilization of the euphotic layer and consequently the productivity of this area. In addition advective transport of nutrients to this system adds to the natural fertilization of the coastal waters and is a potential source for an increasing number of plankton blooms.

The described circulation pattern makes the NCC the only outflow for the North Sea and Baltic waters. The waters of the NCC accordingly reflect any pollution or nutrient enrichment, caused by human activities, of North Sea and Baltic waters.

In the Skagerrak region about 90 species of fish are recorded, of which some 30 are of commercial importance. Lobster, crab, shrimp and mussels are also harvested commercially in the area. The total annual commercial catch of all species amounts to 400–500 tons (Hognestad, 1984). In addition, the sheltered areas of the Swedish fjords and fjords along the entire Norwegian coast offer excellent sites for a rapidly growing fish farming industry with an annual production value of approximately US$ 750 million.

THE 1988 *CHRYSOCHROMULINA* BLOOM

The first documented indication of this toxic bloom was a report from the west coast of Sweden on May 9th. Dead wild fish were first observed May 13th. The bloom extended around the Skagerrak basin and were advected northward along the west coast of Norway, were culminated in early June. During the bloom, algal concentrations in the surface water (upper 20 m) were measured up to a maximum of 100 million cells per liter. The highest concentrations could be found in a narrow layer between the mixed, nutrient poor, upper layer and the nutrient rich deeper water. Here the algal density could be high enough to be recognized by sonar as an echo scattering layer (Horstman and Jochem, 1988).

In response to the algal bloom, an ad-hoc expert group with members from the University of Bergen, the Marine Research Institute and the Nansen Remote Sensing Center (NRSC) were established, with the responsibility of coordinating the monitoring of the algal distribution and relevant environmental conditions (Berge and Føyn, 1988). Daily forecasts of the algae front position (Fig. 2) were distributed by the Director of Fisheries to the radio, television and newspapers. Very little was previously known about the alga concerned. Investigations of its physiology and its effect on other organisms were carried out concomitantly with the monitoring efforts. The monitoring and the forecasts were used by the seafarming industry and their insurance companies for formulating advice on precautionary evacuation of caged fish. The Seafarming Sales Association in Trondheim established its own monitoring group to assist their members.

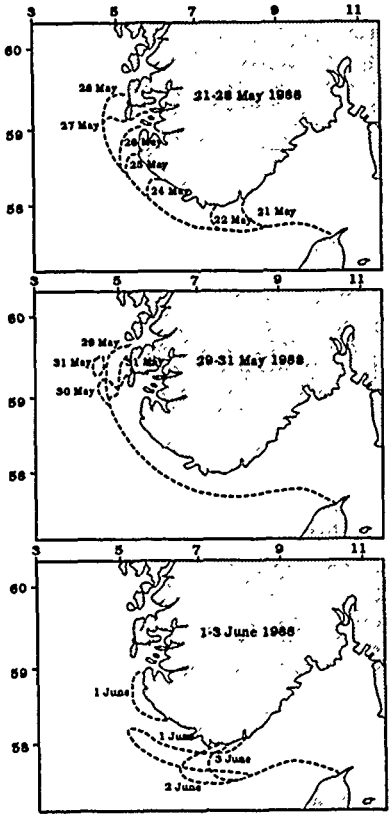


Fig. 2: Summary of the observed algal front including both advance and retreat in the Norwegian Coastal Current from 21 May to 3 June defined by 0.5-1.0 million algal cells pr. liter.

Several research vessels from the Marine Research and University of Bergen were used in addition to vessels from the Norwegian Coast Guard to make *in situ* observation of the biological and physical conditions in the bloom area. After the species had been identified as *C. polylepis*, its concentration was determined with a ship based flow-cytometer and by direct counting of unpreserved water samples in light microscopes (Berge and Føyn, 1988). The movement of the algae front and advection of the warm water from the Skagerrak and up along the west coast turned out to be consistent. The early observations indicated a close correlation between the algal front and the satellite-derived surface warm water front. The spreading and advection of the algae was thus indirectly monitored by satellite infrared surface temperature (IR) data, during cloud free periods (Johannessen et al., 1989a). Further utilization of aircraft patrols from the State Pollution Authority, Royal Norwegian Aircraft and the Swedish Coast Guard were performed. None of these aircrafts were instrumented to do spectral ocean color measurements, but information from infrared (IR), ultraviolet (UV), Side Looking Airborne Radar (SLAR) and Airborne Expandable Bathy-Thermographs (AXBT's) sensors turned out to give some additional information of the algal front zone. Highly dependent on the observation condition (sea state) and depth of major concentrations visual observation of the front signature were possible. Free drifters tracked by the ARGOS satellite system were also deployed in the algae front zone, and were tracked by the ARGOS satellite. As the algae and coastal water advected northward along the coast of west Norway.

The propagation of the algal front, defined as a concentration between 0.5 – 1.0 million cells per liter, indicates the advance and retreat of the bloom along the southern and southwestern coast of Norway from May 21st to June 3rd. (Fig. 2).

Between May 15 and May 21, the algal front moved southwestward at an average speed of 5 km per day, while between May 21 and May 22 the westward advection of the front rapidly increased to about 30 km per day. On May 30 the IR image showed that a narrow front, close to the coast, had reached Stavanger, indicating a mean advection speed of about 25 km per day. *In situ* mapping of the sea surface temperature field, current measurements, and also the algal concentrations registered off Stavanger, verified these satellite observations.

Synoptic satellite and *in situ* observations, data from the monitoring program, and historical data from a previous research program on the NCC were used in the production of daily forecasts of the algal front movement (Fig. 2). In addition, numerical model simulation was done in support of the forecasts (Fig. 3), using a two layer quasi-geostrophic model (Ikeda, et al., 1989) with a grid scale of 5 km. Wind effects and fjord and coastal exchange processes were not included in the numerical model. The model was initiated with the IR-derived surface temperature field from May 22 and later compared to IR images, current measurements, and algal observations, until June 1 (Johannessen et al., 1989a). Neglecting local growth or

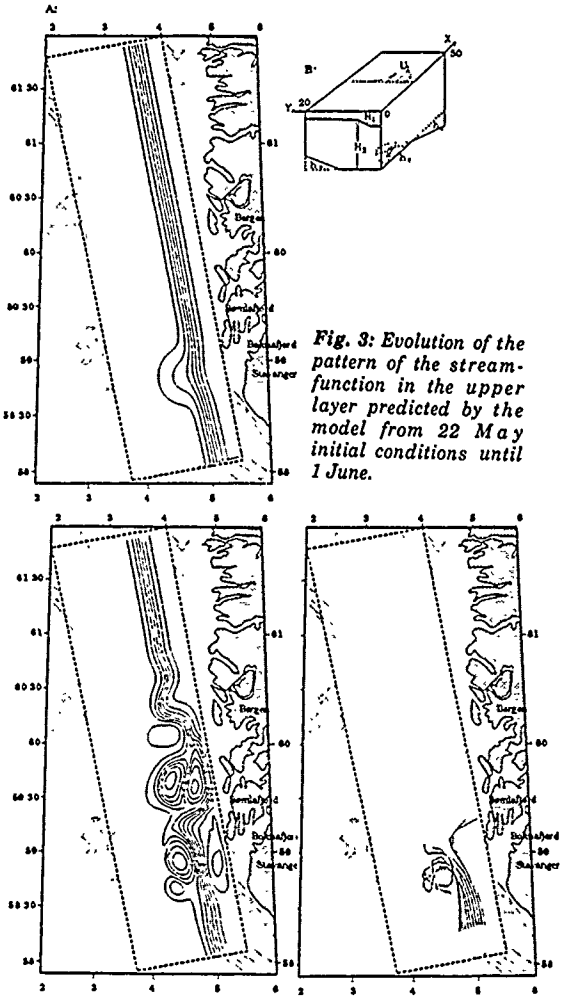


Fig. 3: Evolution of the pattern of the stream-function in the upper layer predicted by the model from 22 May initial conditions until 1 June.

death, the algae were assumed to drift passively with the NCC. Streamlines of the flow pattern for the upper layer in the model on May 22 to June 1 are shown in Fig. 3. During this period the wind was modest, supporting the omission of wind effects from the model.

In the model calculation, the initial meander from the May 22nd IR image developed into two vortex pairs on June 1, one pair off Stavanger and the second pair farther north. Assuming that the algae drift passively, the advection of a simulated algal front was included in the model from May 26 to June 1. The downstream evolution of the algae trajectories showed a meander-like propagation pattern in agreement with the vortex pair off Stavanger June 1. The central part of the front propagated downstream at an average speed of 20-25 km per day. Model particles simulating the algal front, placed near shore, initially followed a clockwise path, towards the coast, whereas particles placed on the offshore edge of the jet-like current followed cyclonically curved trajectories in accordance with the formation of the cyclonic eddy off Stavanger. The impression from the model is that particles remained trapped in the cyclonic feature.

The positions and configuration of the algal front on May 30, and the detached algal plume registered on May 31 (Fig. 2), are in good agreement with the results of the model tracer simulations, which demonstrates the impact that meander and eddy features may have on the dispersion of biological material. The algal front advected northwards with the NCC until the culmination of the bloom around May 30th. Thereafter, the algal front retreated into the Skagerrak during the first three days of June (Fig. 2). Cessation of growth and possibly also algal mortality apparently caused the retreat of the algal front, and dominated over the effect of advection by the northward flow of the NCC.

CONCLUDING REMARKS

Although the forecasting team was established on a short notice, the resulting observations and forecasts have turned out to be of high quality. The benefit from both research ocean models and remote sensing technology have clearly stated it's operational applications. Further this experience have emphasized the need for remote sensing tool to spectral observation of ocean color.

On initiative from NRSC a Norwegian program; "Norwegian Remote Sensing Spectrometry for Mapping and Monitoring of Algal blooms and Pollution - NORSMAP", have been initiated. The program will utilize airborne spectrometers in conjunction with biological and physical *in situ* observation during the seasonal bloom periods.

REFERENCES

Berge, G. and L. Føyn, 1988: Report on the Chrysochromulina polylepis bloom during May-June 1988. Monitoring, forecasting and follow-up. Institute of Marine Research, Bergen, Norway, 16 pp.

Dundas, I., O.M. Johannessen, G. Berge and B. Heimdal, 1989: Toxic Algal Bloom in Scandinavian Waters, - May/June 1988". In press "Oceanography".

Horstmann, U. and F. Jochem, 1988: Report on the Activities and First Results of the Investigations on the Chrysochromulina Bloom in the F.R.G. Institut fur Meereskunde an der Universitat Kiel, pp. 17.

Hognestad, P.T., 1984: Ed. Assessment of the environmental conditions in the Skagerrak and Kattegat. ICES Coop. Res. Rep. No. 149, 1-45.

Ikeda, M., J.A. Johannesen, K.Lygre and S. Sandven, 1989: A process study of mesoscale meanders and eddies in the Norwegian Coastal Current. Journ. Phys. Oceanogr., 19(2), pp. 20-32.

Johannesen, J.A., O.M. Johannesen and P.M. Haugan, 1989a: Remote sensing and model simulation studies of the Norwegian Coastal Current during the algal bloom in May 1988. Int. J. Rem. Sens, (in press).

———, E.A. Svendsen, O.M. Johannesen and K. Lygre, 1989b: Three-dimensional structure of mesoscale eddies in the Norwegian Coastal Current. Journ. Phys.Oceanogr., 19(2), pp. 3-19.

Johannesen, O.M., J.A. Johannesen and B.A. Farrelly, 1983: Application of remote sensing for studies, mapping and forecasting of eddies on the Norwegian Continental Shelf. Proceedings of EARSEL/ESA Symposium on Remote Sensing for Environmental Studies, ESA SP-188, July 1983, 179-187.

Svanson, A., 1975: Physical and chemical oceanography of the Skagerrak and Kattegat. Fishery board of Sweden. Institute for Marine Research. Report No.1, 88 pp.

Tangen, K., 1983: Shellfish poisoning and the occurrence of potentially toxic Dino-flagellates in Norwegian waters. Sarsia 68, 1-7.

Developing Operational Fisheries Oceanography in Newfoundland.

by Scott A. Akenhead

Department of Fisheries and Oceans
Box 5667, St. John's Nfld., A1C 5X1

Abstract

Identifying data products with fisheries applications is just the first of many developments required for operational fisheries oceanography (OFO). The utility of OFO in the Newfoundland region is outlined, and recent Newfoundland initiatives are described. Proving that fishing profits can be increased by OFO is an experiment currently underway aboard trawlers off Newfoundland and Labrador. An inter-ship computer network to exchange data and maps at ranges of 1000 km was created using an intelligent modem for HF packet radio. This low baud rate link appears sufficient for collecting proprietary data from trawl monitors on commercial vessels, and distributing data products genuinely useful to trawler skippers. A need for sea bottom temperature maps (SBT) from sparse data led to a bathymetrically guided contouring. Maps of SBT residuals, after removing the mean temperature at depth, show spatial autocorrelation scales of about 100 km along isobath, and about 100 m in depth (across isobaths).

Acknowledgements

The OFO developments reported are a cooperative effort by DFO, DOC, Seaconsult, Ultimateast, and Fishery Products International, supported by ACOA and NRC. We particularly respect the contributions of Capt. Mike Hogan of FPI, and Ron Wilcox and Sherman Chow of DOC.

Introduction

The annual catch from the Newfoundland and Labrador fisheries is about 650,000 tons and has an economic value to Canada of order one billion dollars. About half of the wholesale cost of the fish products involved is the cost of trawler operations (G. Etchegary, pers. com.). As explored below, fisheries oceanography services promise to increase profit in the Newfoundland fisheries. Except for ice maps, oceanographic data is presently not utilized to find fish in this region.

The oceanographic regime around Newfoundland involves strong currents, ice, and year-round cold water (Petrie et al. 1988). End of winter temperatures for the top 150 to 200 m approach freezing for seawater and are often below the freezing point of fish blood. Cold temperatures persist below the spring and summer thermocline as the Cold Intermediate Layer that extends to the ocean bottom in many areas. Cold water forces many fishes off the shallow banks to find deep 3°C waters on the continental slope. Cod concentrate near the intersection of the cold layer and the bottom (Lilly 1982). For American plaice and yellowtail flounder, large scale mortality is suspected to be due to cold water in the recent cold periods 1972-3 and 1983-4 (R. Bowering, DFO, pers. com.). Fig. 1 shows the SBT regime on the Grand Banks for 1978 and 1984, years of above and below average temperature (from Wells et al 1988). There is important interannual variation in the SBT pattern, and attendent changes in the distribution of fishes.

SBT and water column temperatures influence fish aggregations directly, and also indirectly through prey (Lilly 1982). The basis for OFO in Newfoundland is not just that fish prefer water of optimal temperature or prefer water that contains prey, although these factors are just as important in Newfoundland as anywhere else The added factor that makes OFO uniquely important in Newfoundland is the polar water that floods the fishing banks from time to time and may be cold enough to actually kill the boreal fishes A caveat is required here; American plaice and Atlantic cod are often found in sub-zero water, and it is not known how the cold in 1972 and 1983 might have killed a large fraction of the plaice.

Methods

SBT capture. There are about 4 research vessels offshore in Atlantic Canada on any day, and about 150 offshore trawlers If these silent trawlers were to send environmental data to shore, it would create an new capacity to observe regional ocean events and

to provide ocean services. Every trawler could become a research vessel, if we could capture the data that now crosses the trawler's bridge and is then discarded. For instance, some trawlers use doppler current profilers for midwater trawling operations but despite the value of accumulating this data they keep no records whatsoever. Modern trawlers utilize trawl monitoring devices capable of recording SBT, reporting from the trawl by cables or hydroacoustics. For the Newfoundland experiment we concentrated on the Norwegian SCANMAR units aboard FPI trawlers, and built a computer package to capture, enhance, and display the hydroacoustic trawl signals, including a sonar record of fish entering the trawl.

Satellite SST and ocean colour are not simply applicable to determining submarine fronts and other SBT features, despite the 50–100 m depth of the Grand Bank. The fishing banks in this region have two very distinct layers in summer. The transition from say 10°C, 32 PSU in the surface slab to say 2°C, 34 PSU at 50 m is abrupt, and acts to isolate horizontal temperature patterns in the two layers. On the Southeast Shoal of the Grand Bank, a circular 'hot spot' in SBT is, until the autumn deepening of the surface mixed layer, largely independent of the pattern in SST. Obviously a Gulf Stream ring bumping into the southwest slope of the Grand Bank is an alert to changes in SBT patterns; the matter is in active research (J. Loder, BIO, pers. com.) Overriding these considerations is the 80% mean cloud and fog cover during spring and summer on the Grand Bank. Theoretically this does not prevent SST maps from being accumulated piecewise through holes in the clouds and fog, but operations will complex and still unreliable. Less frequent orbiters that require clear skies, eg Landsat, will be only an opportunistic aspect of OFO in the Newfoundland region. Partly because of the inadequacy of other satellite sensors, we expect Radarsat to become a staple of OFO in Newfoundland.

DFO fish surveys include XBT profiles at each station, but for the Newfoundland region these are never available within a week of collection These trawling surveys do provide valuable climatological statistics, although there is poor seasonal coverage. The NOAA SEAS program (one-way GOES transmissions) has enabled a smattering of real-time SBT from the fisheries surveillence vessels.

Communications. From the Tail of the Bank (42°N) flatfish fishery to the Davis Strait (67°N) shrimp and turbot fisheries, two-way data communication in this project must span 2800 km. We found using voice for data transfer by HF radio (even when possible) impossibly tedious, and satellite capital and operations costs prohibitive. OFO traffic must be secure, since the operations of fishing boats are not only company secrets but the secrets of competing fishing skippers. A new 'intelligent modem' for surface wave HF packet radio, invented by DOC, met our requirements. This error-correcting modem varies packet size according to noise conditions, and scans HF frequencies to find a relatively quiet channel. Surface wave radio has inherent limitations such that our OFO system must cope with a minimum data exchange rate of 100 baud.

Statistics. Mapping the SBT on the Grand Bank in a statistically defensible way is not trivial, even with abundant data. Yet the problem at hand is to produce SBT maps from the sparse data available in some operational time frame. Problems of synopticity, feature resolution, error recognition, and map accuracy must still be addressed. A few XBTs are not sufficient to define the *a priori* temperature field, but if the spatial scales of changes in temperature are large enough, it may be possible to map these changes with relatively few data. Borrowing from the techniques of meteorology, our method is to make an intelligent first guess, currently by using the long term mean for each month. At the large spatial scales we are dealing with, water flows are constrained to follow bathymetry. The result is both mean and residual SBT patterns have relatively short scales across isobaths, and long scales along isobaths. This anisotropy prevents the application of commercial mapping packages. Objective interpolation (OI) (Bretherton et al 1976) also known as kriging (Hardy 1984), depends upon making a model of the autocorrelation structure of the data to control interpolation. We examined two dimensional autocorrelation models in OI, assuming orthogonality of the component axes. The concept is simple. a gridpoint that is separated from a SBT datapoint by 1 kilometer horizontally (h) but 50 meters in depth (z) might be just as far away in correlation space as another SBT datapoint that was 100 kilometers away but at the same depth as the gridpoint.

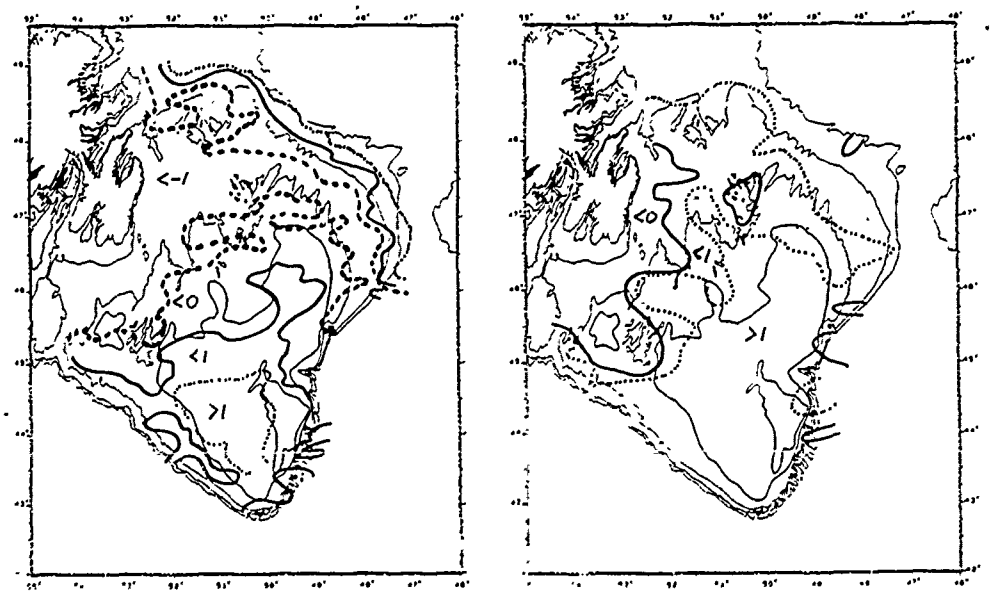


Figure 1. May-June Grand Bank SBT from a cold year (1984, left) and a warm year (1978, right).

Results

Figure 2 contrasts maps of SST and SBT on the Southeast Shoal of the Grand Bank June 15-27 1978. The existence of such a two layer system limits our utilization of satellite SST data to identifying features such as the Gulf Stream ring intrusions which occur on the southwest slopes.

A DOC license to Ultimateast led to production of the described modem (trade name Datahail) and allowed a computer network to be set up between ships thousands of miles apart at sea, of course including data processing stations on shore. Communications charges are small. Trials aboard trawlers produced an unexpectedly high volume of traffic that included engineering orders and stock market quotes as well as XBT profiles and weather reports. We found unexpectedly distant reception: a network node being demonstrated for sales in the Solomon Islands exchanged packets with a ship in Davis Strait. The OFO products developed in this project are coded in NAPLP for international marketing and to ensure compatibility with marine GIS developments.

Figure 3 is a contour map of the correlation structure for SBT residuals on the top of the Grand Bank, using a second degree markov model for autocorrelation (Thiébaux and Peder 1987). The equation fitted by nonlinear least squares was $r = r_0 r_h r_z$ where r_0 accounts for within station error, $r_h = e^{-ch}[cos(ah)+(c/a)sin(ah)]$, and r_z is similar. Because we extract simultaneous vertical and horizontal components from the data, we claim to have extracted the along and across isobath components. In practise, we extract a three dimensional autocorrelation model by including temporal separation. Although we don't at present use the temporal autocorrelation directly in OI, it indicates the elapsed time after which SBT residuals are no longer synoptic. This multivariate correlation structure OI should not be confused with two dimensional OI.

Discussion

Products. Behind the bridge display of accumulated catch, catch rate, and SBT history that develops during each trawling, the network receives a report on catch rate and SBT every 5 minutes. This provides unprecedented detail on catch statistics and temperature, dealing with those infrequent and ephemeral concentrations that are above the threshold of commercial interest. Unless the collections are from working trawlers, the events that actually comprise the fishery are not likely to be observed. This presents a 'Catch 22' situation: the research observations required to design OFO products and services can only be collected from a system designed for commercial operations. The investment in OFO in Newfoundland is speculative, but the potential payoff is quite handsome in three ways: 1) the fishing fleet can become more efficient. 2) scientists get a flood of environmental data from the trawler fleet, enough to reveal events like the suspected temperature related mortality of American plaice on the Grand Bank. 3) stock assessment biologists get new detail in catch rate data to help them disentangle changes in stock biomass from changes in fishing technology and the environment.

In an era of fibre optics and gigabaud transmissions, our low cost 100 baud network may seem anachronistic, but it is important to ask, "How much data really needs to be sent?" Consider that the ship's microcomputer has the software to draw a display with axes and bathymetry, given just the lower left and upper right coordinates. A few more coordinates and spline coefficients produce a smooth isotherm on this map. Modern data compression algorithms reduce the volume of transmissions dramatically. This project benefitted enormously by a contract from DOC to develop new data compression algorithms for NAPLP. In theory our network could transmit 'image' products in a reasonably short time, but is a one megabyte false colour SST image actually useful to a fishing skipper? Or is a more highly processed 'cartoon' of ocean features with various fisheries annotations more useful? Our conclusion was that we ought to present a relatively small amount of highly processed and pertinent information in as spare and clear a fashion as possible: the quintessence of all available data. We went for the cartoons.

Acceptance. We tried to avoid a reactionary stance from

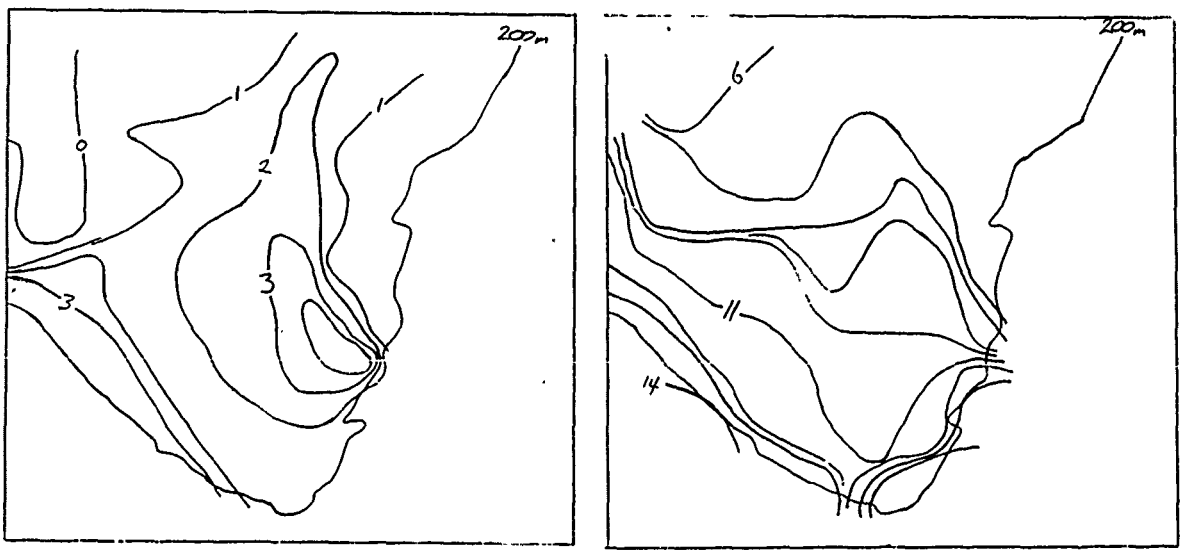


Figure 2. SBT (left) and SST (right) on the Southeast Shoal of the Grand Bank, June 15-27, 1978.

trawler crews faced with new technology by ensuring each trawler received immediate benefit from each new tool. For instance, the low-cost data link allowed every crew member to exchange electronic mail with his family ashore, and was very popular. Computerizing the SCANMAR improved its immediate utility and was seen as genuinely useful Simplifying the weather reports from voice to computer mail was gratefully received These improvements led to acceptance of computers on the trawler bridge, and we are convinced that access to OFO products, including past and current catch data, will lead to increased demand for operational products.

The Newfoundland project is to establish an prototype OFO program which will develop through positive feedback. Better data will allow better analyses and more effective products that increase fishing profits. Identifying these new profits and encouraging their reinvestment for subscriptions to remote sensing data products is the key to developing OFO. Proving that fishing profits can be increased by taking advantage of environmental data is an experiment currently underway on trawlers off Newfoundland and Labrador. Ships with all the OFO information we can muster are to be compared with ships working without this advantage. Only hard data comparing searching time and trip length will demonstrate that OFO is worth buying.

Future. For the purposes of OFO, there are two scales of temporal variation in the ocean: climatic (weeks and years) and weather (days). Both are important, but involve different data capture, products, and budgets. The first regime is the subject of this paper, and can be addressed statistically because the flow of data is potentially sufficient to maintain the required maps. The faster regime requires dynamic physical oceanography models (eg. Robinson et al. 1987) to utilize incoming data. In the present experiment we can do no better than ignore high frequency components in the absence of resources to maintain a dynamic model. Our existing product might immediately be improved in synopticity by relocating past SBT observations according to static flow fields, eg./ from Greenberg and Petrie (1988).

References

Bretherton F P, R. E Davis and C B Fandry 1976 A technique for objective analysis and design of oceanic experiments. Deep-Sea Res. 23:559-582

Petrie, B., S. Akenhead, J. Loder and J. Lazier. 1988 The cold intermediate layer on the Labrador and Northeast Newfoundland continental shelves, 1978-86 NAFO Sci. Coun. Studies 12:57-69.

Greenberg, D. A., and B. D. Petrie 1988 The mean barotropic circulation on the Newfoundland shelf and slope J. Geophys. Res. 93:15541-15550

Hardy, R. L. 1984. Univeral kriging, least squares collocation, and optimum interpolation (what's the difference?). 16th European Meeting of Statisticians (EMS) University of Marburg, Germany, 3-7 September 1984, revised October 1984 13 p.

Lilly, G. 1982 Influence of the Labrador Current on predation by cod on capelin and sand lance off eastern Newfoundland. NAFO Sci. Coun. Studies 3:77-82.

Robinson, A. R., J. C. J. Nihoul and B. M. Jamart. 1987 Predicting open ocean currents, fronts and eddies. Three dimensional models of marine and estuarine dynamics. Elsevier Oceanogr Ser 45:89-111

Thiébaux, H. J. and M. A. Pedder 1987 Spatial objective analysis with applications in atmospheric science Academic Press. 299 p.

Wells, R., W. B. Brodie, C. A. Bishop and J. W. Baird. 1988 Distribution and abundance of three fish species on the Grand Bank in relation to depth and temperature of the water. NAFO SCR Doc 88/94. Ser No N1546 173 p.

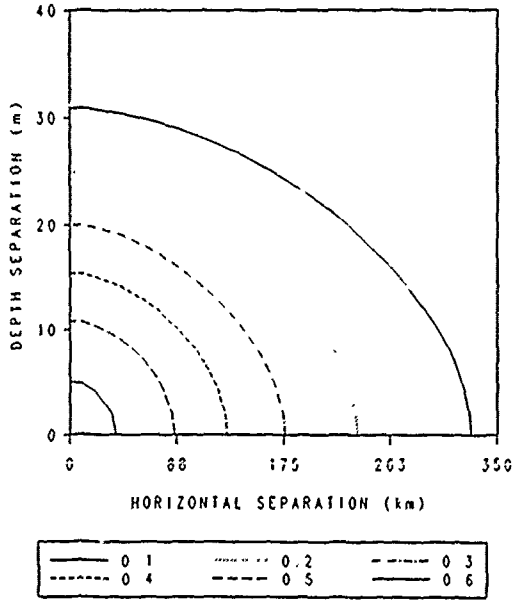


Figure 3. Autocorrelation structure with horizontal and depth separation, of Grand Bank SBT.

Assessing Fishing Effort by Remote Sensing in the Scotia Fundy Region of Fisheries and Oceans Canada

G. Sharp, J. Pringle and R. Duggan

Benthic Fisheries and Aquaculture Division, Fisheries and Oceans Canada, P.O. Box 550, Halifax, Nova Scotia, Canada, B3J 2S7

ABSTRACT

Data on the distribution and level of fishing effort is critical in the development and management of fisheries. Near shore small boat fisheries do not lend themselves to conventional methods of documenting effort including log books and vessel observers. Colour aerial photography at 1:6000 scale has been used to delimit lobster fishing effort to a resolution of several meters at an error level of 11.3 % on interpretation. Contouring of digitized data enabled identification of areas of peak fishing effort which could be correlated to substrate,community type, depth, temperature and water movement. Baseline data can be used to evaluate changes in fishing effort with catch levels, and environmental impacts. Gear conflicts between types of fixed gear and between traditional fisheries and aquaculture can be prevented or moderated by a detailed knowledge of gear deployment.

Key Words: fishing effort, colour photography

INTRODUCTION:

The nearshore area (water depths (<40 m) of the Scotia-Fundy region are the grounds for a wide variety of mobile and fixed fishing techniques. Fixed gear includes gillnets for pelagic and benthic fish, trap nets, wiers, baited hooks on moored trawl or longlines. However the predominant fixed gear is the trap fishery for the american lobster (Homarus americanus). Single or multiple traps are placed on the bottom and marked at the surface by a float. This fishery is limited by the number of fishing licenses, number of traps per fisherman and a closed season. Within these limitations the level of fishing effort varies both during the season, annually and geographically.

Fisheries and Oceans Canada as a part of its resource management mandate monitors fishing effort. Vessels over 25.5 gross tons must provide logbooks detailing catch, fishing time gear type and amount, location etc. However, the nearshore is dominated by smaller vessels and effort data must be obtained indirectly. The amount of fixed gear (traps, gillnets etc anchored to the bottom) is limited per vessel by fishing district. Landings, from these vessels are reported at point of sale and the fishing grounds cannot be determined accurately. Interview surveys of lobster fishermen provide detailed data on vessel characteristics; gear type and amount; maximum fishing distances; fishing depth; fishing effort (trap hauls) (Duggan, 1985). The lobster fishing grounds were quantified from interviews on maximum depth and distances fished and the intensity (traps km^{-2}) of fishing estimated for lobster fishing areas (Pringle and Duggan, 1985).

A wide range of fishing intensities have evolved due to climate, changing socio-economic trends and management initiatives along the Atlantic coast of Nova Scotia. However the level of spatial resolution of the above effort assessment methods is in the range of tens of kilometers.

Fisheries and Oceans has turned to remote sensing methods to provide high resolution spatial and quantitative data on the distribution of fixed fishing effort. This paper outlines the development, types of applications of this method and examines its advantages and problems.

METHODS:

In the development phase of this technique lobster buoys of two types and four color schemes were placed at known locations in the near shore. Photographs were taken using Kodak no. 2424 black and white infrared and Kodak no. 2448 color positive in a Wild RC-8 camera at

altitudes of 405m, 458 m, 763m and 1,220 m with 60% endlap and 30% sidelap (Pringle and Duggan, 1983). Loran C and two dimensional triangulation with known location points were used to position photographs. Photos were interpreted on a light table with an eight power lens magnifier. The buoys identified were digitized and plotted on a Geobased mapping system ([1]K. Speight pers comm). Buoys were identifiable at 1:5000 scale on color photographs at an error level of 11.3%. Subsequent surveys used color film at 1:6000 scale in a Zeiss RMK with 15 cm lens during calm weather when possible at a sun angle less than 30°. Buoys were counted within 2.25 ha grids to allow contouring of buoy densities at intervals of 50 buoys km^{-2}. DISPLA[2] generated a linear interpolation of adjacent surface grid points in both the x and y direction. The mean distance between buoys within a variety of grid sizes and buoy densities was calculated on the raw data file by a nearest neighbour program (Cottam and Curtis, 1956). Three sites were surveyed one to provide base line data on lobster fishing effort to determine impact of coal waste rock dumping in the nearshore of Cape Breton, one to compare seasonal changes in fishing effort distribution in Pubnico N.S. and a third to help select a dredge spoil dump site near Cape Sable Island N.S.

RESULTS/APPLICATIONS:

Level and distribution of effort.
The resolution of lobster fishing effort using aerial photography was 1 buoy per $50m^2$ and provided a range of means for 1 km^2 sub sections of the coast. (Table 1). Data obtained for the entire Lobster fishing areas by interview techniques and extrapolation by fishing grounds and trap limits were dependant on definition of the fishing ground by a maximum depth limit (Pringle and Duggan, 1985).

Table 1. Comparison of lobster trap buoy densities from aerial surveys and indirect methods for selected areas in Nova Scotia, Canada.

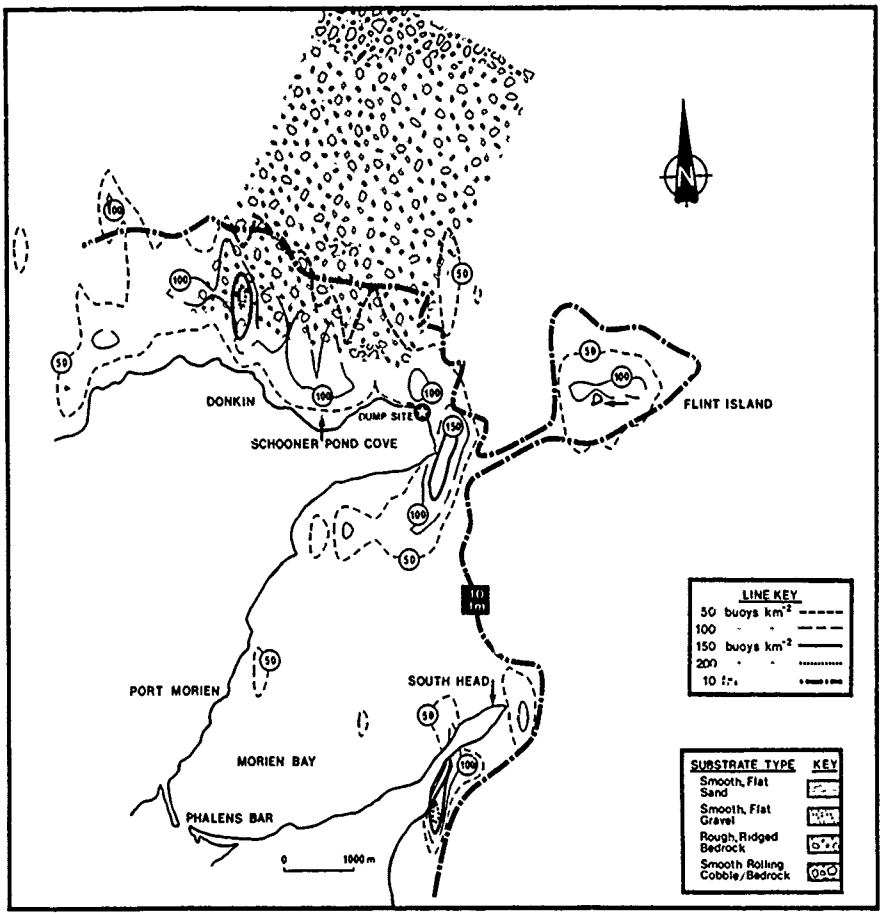
	Location			
	Cape Sable	Pubnico	Sydney	Test Site
	Lobster trap buoys km^{-2}			
Aerial Photo	44-954	44-1021	46-109	70-50
Interviews	-	-	-114	74
Trap limit x licenses	424	424	145	142

Contours of lobster trap distribution in the Sydney area define areas of intense fishing effort >200 km^{-2} (Fig. 1).

[1] Maritime Resource Management Service, P.O. Box 310, Amherst, N.S., B4H 3Z5

[2] Integrated Software System and Plotting Language. Integrated Software Systems Corp., 4186 Sorrento Valley Blvd., San Diego, California 921121, U.S.A.

Figure 1 The distribution of lobster trap buoys in 50 km^{-2} contour intervals with depth and substrate type at a coastal site near Sydney, Nova Scotia.



Only a small portion of the fishing grounds (10%) support this level of fishing effort and the majority of the area (64%) has less than 50 traps km^{-2}. Nearest neighbour distances between buoys reflect the density contour of the sample but have a wide range of overlap (Table 2).

Table 2 : Nearest neighbour distances m between trap buoys in the Cape Breton survey in a range of buoy density contours. sd (standard deviation)

Density contour buoys km^{-2}	>200	100-150	50-100
Mean distance	34 sd 17	72 sd 30	105 sd 37
Range	5-88	25-159	51-174

Successive measures of effort provide the resource manager wtih a data base defining both the level and variation in fishing effort allowing evaluation management initiatives such as license buy back, enforcement of trap limits and expansion of the fishing ground.

Biological data
Fishing effort within the parameters of a minimum economic return reflects the relative distribution of adult lobster stocks. Although there are long distance movements of large adults (>100 mm carapace length) fishing effort in the nearshore follows local movement of most animals (<20 km). Over the fishing season traps sample or remove between 60% to 80% of the legal size animals (>81mm carapace length) (Paloheimo, 1963). In the Sydney area 90% of traps were located within the 18 m depth contour (Fig 1). Sidescan sonar defined a homogeneous substrate suitable for lobsters well beyond this depth. Temperature was 2 to 3°C lower at the 20 m level than at the surface during the fishing season. It was concluded animals were optimizing their environment for growth and egg development within the fishing ground. Records of catch per trap haul number of days fished and cumulative catch can be assigned to levels of effort density to determine the productivity of a fishing ground. In the Cape Breton area cumulative catch per trap was 29.5 kg in 1985 and 30.7 kg in 1987. This catch was transformed to a yield of 1.5 g m^{-2} within the 50 trap km^{-2} contour and 6.0 g m^{-2} in the 200 trap km^{-2} contour. In Pubnico fishing grounds the total level of effort between the December and following May surveys were not greatly different at 9,595 and 10,207 buoys respectively. However, the successive surveys of fishing effort detected small changes in distribution resulting in significant changes in depth and habitat (Fig 2). Adult animals have moved from a shallow kelp covered bottom to a deeper rock and gravel habitat.

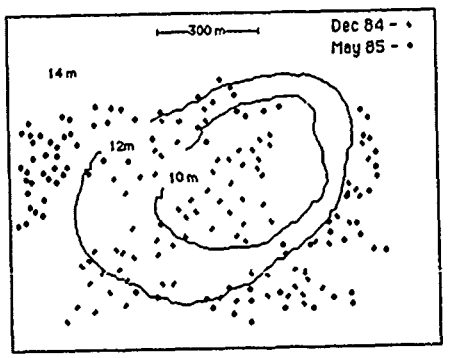


Figure 2. Seasonal Distribution of lobster trap buoys on and near a shoal near Pubnico, N.S.

Environmental impact assessment.
Documentation of a baseline fishing pattern is needed to monitor changes in habitat or water column characteristics due to point sources of pollution or more general impacts of oil spills, thermal effluents, sedimentation and alteration in water circulation patterns. Identification of areas devoid of fishing activity can be utilized to select dredge spoil ocean dumping sites. A deep channel (>12 m) without fishing effort was identified near Cape Sable Island for a dredge spoil disposal site.

Gear conflict resolution
This detailed fishing effort information can be used to prevent gear conflict between types of fixed gear. The rapid development of aquaculture with new types of fixed gear needs an evaluation of space utilization of the nearshore (Sharp and Lamson, 1989). A knowledge of trap buoy distribution allows positioning of aquaculture equipment with minimal interference with traditional fishing gear. The use of new mobile gear may also be restricted to certain areas dependant on trap buoy distribution.

DISCUSSION:

Assumptions required to extrapolate buoy densities to adult lobster distribution have some weak points. The first is the assumption one buoy represents one lobster trap. In the nearshore (<10 km) a single trap per buoy is normal with some exceptions. However in deeper water (>20 m) it is more efficient to increase the traps per buoy from 10 to a maximum of 50 (Sharp and Duggan, 1985). Fishermen interviews and direct observation of fishing activity can be used to correct these figures.
Aerial photography provides only one look at the level and distribution of effort. Trap design, bait, soak time are some of the factors which affect catch rates (Krouse, 1988). However traps are rarely moved on mass but are finely adjusted on a daily basis. Resurveys on a biweekly basis would provide a sufficiently detailed seasonal pattern and abundance. Mapping of effort distribution is a more accurate depiction of adult animals than animal tagging. Tagging returns rarely have precise location data and small scale movements are not detectable. Storm conditions can destroy up to 25% of traps particularly in exposed areas. These events can be documented and considered in the analysis of data. Distribution of traps provides data only on the adult or legal size animals although traps can capture animals to approximately one half legal size they are not retained by the fishermen.

Prediction of future landings will continue to depend on ground research work.
The survey and interpretation costs totaling 0.16 man-hours and $ 109.10 (Canadian) for the aerial photography method are 10 times the cost of aerial visual scanning and 30 times the cost of systematic fishermen interviews (Pringle et al 1986). To make this tool more cost effective alternative flight platforms such as tethered balloons or remote controlled vehicles may offer savings. Greater magnification of the image and/or computer aided image analysis has yet to be tested.

References

1. Cottam, G. and J.T. Curtis, "The use of distance measures in Phytosociological sampling", Ecology, vol 37, pp451-460, 1956.

2. Duggan, R.E., "Characteristics of the lobster fishery on the oceanic coast of Nova Scotia", Can. MS Rep. Fish. Aquat. Sci., no.1883, 37pp, 1985.

3. Krouse, J.S., "Performance and selectivity of trap fisheries for crustaceans", In: Marine invertebrate fisheries: their assessment and management (ed.J.F.Caddy), pp307-325, 1989.

4. Paloheimo, J.E.,"Estimation of catchability and population sizes of lobsters" J. Fish. Res. Board Can., no.20, pp59-88, 1963.

5. Pringle, J.D. and R.E.Duggan, "An estimate of yields for oceanic Nova Scotia grounds", CAFSAC res. doc 85/102, 16pp., 1985.

6. Pringle, J.D., R.E. Duggan and G.J.Sharp, "A cost evaluation of techniques designed to assess Canadian lobster fishing effort", Fish. Res., no.4, pp283-295, 1986.

7. Pringle, J.D. and R.E. Duggan, "A remote sensing technique for quantifiying lobster fishing effort", Can. Tech. Rep. Fish. Aquat. Sci., no.1217, 16pp, 1983.

8. Sharp, G.J. and R.E. Duggan, "An aerial survey of near-shore and mid-shore lobster fishing effort in southwestern Nova Scotia, spring and fall 1983-1984", Can. MS Rep. Fish. Aquat. Sci., no. 1847, 37pp, 1985.

9. Sharp, G.J. and C. Lamson, "Approaches to reducing conflict between traditional fisheries and aquaculture", Bulletin of the aquaculture association of Canada, no.88-4, pp150-152, 1988.

LA TÉLÉDÉTECTION: UN OUTIL ADDITIONNEL POUR LES ÉTUDES QUI PORTENT SUR LE RECRUTEMENT LARVAIRE DE CERTAINES ESPECES MARINES

André LAVOIE [1-3], Christiane HUDON [2], Jean-Marie DUBOIS [3] et Gérald VIGEANT [4]

(1) Dépt. de géographie, Uni. de Montréal, QC (2) Pêches et Océans Can., Ste-Anne de Bellevue, QC
(3) CARTEL, Univ. de Sherbrooke, QC (4) S.E.A., Env. Canada, Saint-Laurent, QC

RÉSUMÉ

Le recrutement résulte d'une sélection naturelle des larves qui survivent à maturité. La température de l'eau, entre autres, est un élément qui conditionne les stades de croissance larvaire des individus. Le mouvement des masses d'eau est également un élément important puisque les larves sont soumises à ses déplacements. Le but du travail est de vérifier d'une manière synoptique les conditions de température de surface de l'eau et, si possible, les mouvements des masses d'eau. À cette fin, une série temporelle d'images thermiques, prises par les satellites NOAA6 et 7, a été retenue pour les étés 1981 et 1984. On a formé des classes de température de la surface du golfe du Saint-Laurent, et les patrons de distribution représentés ont été reliés à des données additionnelles sur les conditions météorologiques. Il a été possible, lors de l'interprétation des images, de mettre en évidence: un gradient de température de surface croissant des côtes du Québec vers celles de Terre-Neuve, des résurgences d'eau froide le long des côtes du Québec et des apports en eau froide, issus de la mer du Labrador et entrant par le détroit de Belle-Isle. Ces conditions semblent défavorables à la croissance de larves le long de la Côte-Nord du Québec. Une vérification de cet état a été faite par le biais de statistiques de débarquements du homard américain (*Homarus americanus*).

Mots clés: Imagerie thermique, NOAA, Pêcheries, Golfe du Saint-Laurent

INTRODUCTION

En 1985, la Direction de la recherche sur les pêches du ministère des Pêches et Océans du Canada désirait vérifier le potentiel d'un outil supplémentaire pour aider à la compréhension du recrutement larvaire de certaines espèces marines tels les mollusques et les crustacés. L'outil envisagé est l'image infrarouge thermique.

Le but général poursuivi par la Direction est d'obtenir des informations additionnelles permettant à ses chercheurs de mieux définir, spatialement et temporellement, les zones de productivité biologique et les trajectoires empruntées par les larves de poissons ou de crustacés. À plus long terme, les chercheurs visent à pouvoir prédire l'abondance des stocks et stabiliser les débarquements dans le but d'une meilleure gestion des pêches.

Le premier objectif spécifique de l'étude est de fournir des cartes montrant l'évolution des températures et la circulation des eaux à la surface. Si possible, on doit également porter une attention particulière aux panaches de rivières ainsi qu'aux remontées d'eau froide le long des côtes. La température, les courants et les remontées d'eau froide font partie d'un système air-mer qui les associe à des facteurs météorologiques comme les vents. Ainsi, un deuxième objectif de l'étude consiste à vérifier les diverses conditions de surface de l'eau par le biais de données météorologiques qui correspondent au moment d'acquisition de l'image. On confirme d'une certaine façon l'interprétation faite sur les images.

Par ailleurs, nous nous sommes intéressés à l'impact possible des paramètres physiques et météorologiques sur l'abondance de ressources marines. À cette fin, nous avons utilisé des statistiques de débarquements de poissons, et plus spécifiquement de homards, le long des côtes du golfe du Saint-Laurent. Il s'agit d'un intérêt tout à fait extérieur au cadre fixé à l'origine par cette étude, mais ces statistiques se sont révélées particulièrement intéressantes.

LE NORD-EST DU GOLFE DU SAINT-LAURENT

Le secteur faisant l'objet de l'étude se situe le long de la Basse-Côte-Nord du Québec, dans la partie nord-est du golfe du Saint-Laurent, entre les détroits de Jacques-Cartier (Havre-Saint-Pierre) et de Belle-Isle (Blanc-Sablon) (figure 1). On s'intéresse particulièrement à ce qui se passe à la surface d'une bande d'eau côtière de 40 km de largeur environ. Cette région a été choisie parce que l'on désire connaître les conditions de courant en surface de manière à pouvoir déduire le transport possible de larves, et aussi parce que c'est une région moins bien connue au point de vue océanographique et biologique. Ce sont les larves du homard américain (*Homarus americanus*) qui sont surtout visées ici. On désire savoir si les faibles débarquements de cette région, par rapport au reste du Québec, sont liés à des facteurs limitants de types environnementaux. Même si dans cette étude on s'intéresse aux larves de homard, les résultats obtenus peuvent néanmoins s'appliquer à toutes les espèces de poissons, mollusques ou crustacés qui ont un stade de développement larvaire en suspension à la surface de l'eau. Leur répartition spatiale est également fonction des conditions de courants en surface.

LES OUTILS ET LES MÉTHODES

Compte tenu de l'envergure du territoire à couvrir, l'outil privilégié pour rencontrer nos objectifs est l'image infrarouge thermique acquise par le capteur AVHRR du satellite NOAA. La période d'intérêt s'échelonne entre les mois de mai et d'octobre. L'année 1981 a été retenue à cause de la disponibilité de données de température de l'eau acquises lors d'une campagne océanographique.

Pour diverses raisons nous n'avons pu obtenir de suivi temporel régulier. On a néanmoins retenu 17 images réparties sur 2 années soit 1981 et 1984, de manière à couvrir la période d'intérêt fixée (tableau 1). De plus, le Service de l'Environnement Atmosphérique d'Environnement Canada à Montréal nous a permis d'accéder aux données météorologiques correspondantes aux images et pour lesquelques jours qui les précèdent (tableau 1). Des mesures de température de surface de l'eau ont aussi été obtenues du même Service. L'ensemble de celles-ci provient de stations

terrestres (météo.) et d'observations (T°) faites par les bateaux circulant dans cette partie du golfe.

Les images satellites ont été corrigées pour les effets géométriques mais elles ne l'ont pas été pour les effets atmosphériques. Les valeurs numériques des images ont ensuite été transformées en températures suivant une fonction de transfert connue dans la littérature (Mouchot et Lambert, 1986; Malila and Anderson, 1986). On obtient alors des températures apparentes de l'eau. La corrélation entre ces températures et celles prises par les bateaux a été vérifiée et donne 99%. De plus, on a noté un écart moyen de 2,2° C entre les deux (écart-type 0,31°C). Les températures relevées sur les images ont donc été réajustées et les nouvelles limites de température apparente de l'eau ont été établies (tableau 1). Le résultat est une image des classes de température de l'eau interprétable en fonction des gradients que l'on y observe et de la répartition des températures à la surface. Cette répartition peut être indicatrice de la direction générale du courant de surface de l'eau.

INTERPRÉTATION DES IMAGES

Les images satellites ont été utilisées dans cette étude parce qu'elles étaient susceptibles d'apporter des éléments supplémentaires d'information tels les mouvements inhabituels de masses d'eau non révélés par des moyennes statistiques à long terme, ou des variations importantes de températures de l'eau à des moments particuliers, etc. Elles permettent de vérifier ou de confirmer des situations connues dans la littérature, mais surtout d'avoir une idée plus précise de l'ampleur, au niveau spatial, de certaines conditions de surface de l'eau.

Une interprétation détaillée de chacune d'elles est fournie par Lavoie *et al* (1986). Cependant, nous ne donnerons ici que les faits saillants qui font état de quatre éléments d'importance dans ce secteur du golfe. Chacun de ces éléments, pris séparément, a d'ailleurs fait l'objet de recherches et de publications. Par le biais des images, nous pouvons les y regrouper dans la mesure où ils sont apparents au même moment. Ces quatre éléments sont:

- le gradient thermique de surface (température);
- les remontées d'eau froide;
- la contribution en eau froide du courant du Labrador;
- la circulation tourbillonnaire dans le chenal d'Esquiman.

Le tableau 2 montre chacun de ces éléments ainsi que les images sur lesquelles on peut les observer. Certains ouvrages de référence y sont également mentionnés, lesquels font état de ces éléments par le biais des méthodes traditionnelles.

Reprenons chacun des éléments un peu plus en détails. Le premier élément est la variation des températures de l'eau. En effet, sur plusieurs images, il est possible de bien voir le gradient thermique sur le secteur. L'augmentation progressive de la température de surface se fait dans le sens ouest-est, soit depuis la côte du Québec jusqu'à Terre-Neuve. La variation moyenne est de l'ordre de 3,6°C. Les références citées au tableau 2 font état des mêmes constatations à la différence que ces auteurs ont utilisé des moyennes de données ponctuelles sur des périodes pouvant aller jusqu'à 30 ans. Les températures de l'eau en surface semblent donc moins favorables à la croissance de larves, sur les côtes québécoises.

Le second élément observé sur les images, soit les remontées d'eau, permet d'expliquer en partie les températures froides de l'eau le long de la côte du Québec. Ces remontées sont générées par des vents de l'ouest ou du nord-ouest. Nous avons pu le vérifier dans 80% des cas en comparant les directions du vent et la présence sur les images d'une nappe d'eau froide le long de la côte. Pour le 20% qui reste, un changement rapide de direction du vent quelques heures avant la prise de l'image est à l'origine de la différence. De plus, il semble y avoir un facteur de proportionnalité entre la force du vent et l'étendue de la zone froide en surface. Enfin, de façon générale, l'eau froide qui remonte à la surface va s'étendre en direction sud et sud-est sur des distances variant entre 4 et 25 km d'après nos images. On a cependant observé une étendue de 55 km les 29 et 30 septembre 1984. Le vent soufflait alors à des vitesses entre 25 et 50 km/h.

Le troisième élément du tableau 2 a également un impact sur les températures de l'eau du secteur d'étude. Il s'agit d'une contribution en eau froide qui provient du courant du Labrador. Suivant la littérature, celui-ci entre par le côté nord du détroit de Belle-Isle et suit la côte du Québec jusqu'au détroit de Jacques-Cartier à l'occasion. Cette eau étant extrêmement froide contribue aussi à abaisser la température de l'eau du côté du Québec. Il s'agit là d'une condition courantométrique reconnue par plusieurs auteurs. Cependant, comme il ne s'agit pas de quelque chose de stable et de régulier dans le temps, nous avons tenté d'établir la fréquence à laquelle ces entrées d'eau froide surviennent, à partir des hypothèses fournies par les auteurs cités au tableau 2 (sous ce thème). Ils associent ces entrées à une différence de pression atmosphérique entre le golfe du Saint-Laurent et la mer du Labrador. Après examen des images montrant les entrées d'eau froide, on s'aperçoit que les étendues d'eau froide ne se retrouvent que sur les images du printemps. Nous avons ensuite vérifié les conditions atmosphériques correspondantes. Pour chaque date, on note la présence d'une crête de haute pression qui passe sur le golfe. On ne possède cependant pas de confirmation de la présence d'une basse pression sur la mer du Labrador. Ainsi, en condition de haute pression sur le golfe, le courant résiduel entre par le détroit de Belle-Isle et suit généralement la Côte-Nord du Québec. La situation inverse est difficilement observable puisqu'en condition de basse pression sur le golfe, celui-ci est généralement couvert de nuages.

Le dernier élément du tableau concerne la présence de tourbillons dans le secteur du chenal d'Esquiman. Ceux-ci (mentionnés dans la littérature) pourraient avoir un impact certain sur le transport de larves dans le secteur d'étude. L'examen des images nous permet difficilement de percevoir leur présence. On observe certaines formes qui ressemblent à la structure d'un tourbillon mais rien d'aussi évident que ce qui a déjà été observé dans l'estuaire et la partie nord-ouest du golfe du Saint-Laurent (Lacroix, 1986; Lavoie *et al*, 1985; Tang,1980).

En somme, pour les masses d'eau qui longent la côte du Québec, il semble que l'un des principaux traits à retenir soit les températures froides de l'eau à la surface. Ensuite, on retient l'impact des vents (généralement de l'ouest) sur le transport de masses d'eau et des larves en surface. Finalement, le dernier trait est le gradient thermique ouest-est. On voit tout de suite par chacun de ces traits que la Côte-Nord du Québec n'est pas favorisée. D'abord, les températures de l'eau étant plus froides, les larves auront plus de difficultés à survivre ou prendront plus de temps à atteindre la maturité, contrairement à Terre-Neuve où l'eau atteint plus rapidement une température favorable pour les larves. Au point de vue du transport de ces larves, comme les vents sont généralement de l'ouest, celles-ci risquent d'être exportées et de se déposer ailleurs que sur la Côte-Nord, si cette dernière est le lieu de ponte.

LA PECHE SUR LA COTE-NORD

La partie nord-est du golfe Saint-Laurent est considérée comme un secteur moins riche au niveau biologique (Steven, 1971). Ceci se réflète-t-il au niveau de la pêche maritime? À cette fin, les statistiques sur les débarquements totaux au Québec sont un bon indice de la situation. Une attention particulière sera apportée à la situation du homard.

Durant la période entre 1953 et 1984, les débarquements totaux (poissons, mollusques et crustacés) variaient entre 40 000 et 130 000 tonnes au Québec. Plus particulièrement entre 1973 et 1984, les débarquements de la Côte-Nord et d'Anticosti, bien que décroissant, constituaient en moyenne 60% des débarquements. En ce qui concerne le homard américain, les débarquements totaux variaient entre 1 000 et 2 000 tonnes pour la période de 1953 à 1984. La Côte-Nord ne contribue en moyenne que pour 1% de ce total. On doit cependant considérer que l'effort de pêche dans ce cas est beaucoup moins grand sur la Côte-Nord qu'ailleurs au Québec. Par exemple, d'après Therriault (1988), 70% des bateaux des Iles-de-la-Madeleine pratiquent la pêche au homard contre 8% sur la Côte-Nord et ce malgré une flotte de bateaux 2,5 fois supérieure à celle des Iles. La pêche au homard n'est donc pas

l'activité première de cette région.

À l'interprétation des images, on a supposé que la température de l'eau et les courants étaient plus favorables à la survie et à la croissance de larves le long de la côte de Terre-Neuve que de celle du Québec. Les statistiques de débarquements l'indiquent aussi. Prenons la division des zones de pêches de l'Organisation des Pêches de l'Atlantique Nord-Ouest (OPANO) et comparons les débarquements totaux (poissons et invertébrés) des secteurs 4R (côte ouest de Terre-Neuve) et 4S (Côte-Nord et Anticosti, au Québec) (figure 1). On voit que les débarquements totaux du secteur 4R (1973 à 1984) sont en moyenne le double de ceux du secteur 4S et ce malgré la plus grande superficie de ce dernier (figure 2). Plus spécifiquement, si l'on compare les débarquements de homards de la Côte-Nord et ceux de la côte ouest de Terre-Neuve, pour la période entre 1953 et 1984, on s'aperçoit que la Côte-Nord enregistre un peu moins de 2% des débarquements (figure 3). La zone de pêche la plus près du détroit de Belle-Isle, à Terre-Neuve, enregistre à elle seule des débarquements qui sont environ 10 fois supérieurs à ceux de la Côte-Nord en entier. Les tonnages débarqués rendent compte d'un effort de pêche sûrement plus important à Terre-Neuve qu'au Québec, mais nous n'avons cependant pas les chiffres nécessaires, à l'heure actuelle, pour le confirmer. Toutefois, cette différence est suffisamment importante pour palier à l'imprécision des statistiques de débarquements et indique que la ressource est nettement plus importante à Terre-Neuve.

RÉSULTATS ET CONCLUSIONS

On peut donc constater, par le biais des débarquements, que la pêche au homard est meilleure à Terre-Neuve que sur la Côte-Nord du Québec. La température de l'eau serait plus favorable à la survie et à la croissance des larves le long des côtes de Terre-Neuve, compte tenu qu'elle est de 3 à 4°C supérieure à celle des côtes du Québec. Cette constatation est fondée à la fois sur l'analyse de statistiques à long terme et sur l'observation directe au moyen d'images satellites. Le régime de courant semble également contribuer à ce système puisque la Côte-Nord semble être souvent sous l'influence de l'eau froide qui entre par le détroit de Belle-Isle. Par ailleurs, Terre-Neuve est favorisée par une entrée d'eau plus chaude depuis le détroit de Cabot pour ensuite longer la côte en direction nord-est. Les vents sont également importants puisqu'en plus de favoriser l'arrivée d'eau froide en surface, ils forcent un déplacement de cette couche de surface vers la côte de Terre-Neuve et le centre du golfe.

Un examen plus détaillé de ces éléments (température de l'eau, vitesse et direction des vents, conditions météorologiques générales) serait nécessaire sur une plus longue échelle temporelle que seulement celle qui correspond aux images satellites analysées dans cette étude. La corrélation entre les divers éléments et les débarquements pourrait être également vérifiée. Diverses tentatives ont déjà été faites, avec plus ou moins de succès, pour prédire les débarquements de homards dans le sud du golfe. On corrélait différents paramètres environnementaux comme les débits d'eau douce (Sutcliffe, 1973), les vents dominants (Caddy, 1979) et la température de surface (Boudreault *et al*, 1977; Dow, 1977) avec les débarquements obtenus quelques années plus tard.

Il est donc intéressant de continuer sur cette voie en examinant avec plus de soins les paramètres physiques qui ont une influence réelle sur la survie, la croissance et le transport de larves, ainsi que leurs variations à long terme. Suivant l'expérience acquise par d'autres auteurs dans le sud du golfe, nous tenterons un exercice similaire pour la partie nord-est du golfe cette fois, avec un ensemble plus complet de données. Il sera peut-être possible alors d'élaborer un modèle de prédiction des bonnes et mauvaises années de débarquements en fonction de plusieurs conditions environnementales. Une publication sur le sujet suivra bientôt dans le Journal Canadien des Sciences Aquatiques et Halieutiques.

REMERCIEMENTS

Nos remerciements s'adressent à la Direction de la recherche sur les pêches, de Pêches et Océans Canada, pour le financement par contrat (OSD 85-00249), ainsi qu'au Service de l'Environnement Atmosphérique pour l'accès aux données météorologiques et de températures par bateaux.

RÉFÉRENCES

Anonyme, "L'industrie des pêches maritimes au Québec - Description statistique", Dir. des services économiques et de développement, Min. des Pêches et Océans Canada. 184 p., 1985.

Bailey, W.B., "On the dominant flow in the strait of Belle-Isle", Jour. Fish. Res. Bd. of Can., vol. 15, no. 6, pp. 1163-1174, 1958.

Boudreault, F.R., Dupont, J.N. et C. Sylvain, "Modèles linéaires de prédiction des débarquements de homard aux Iles-de-la-Madeleine (Golfe du Saint-Laurent)," J. Fish. Res. Bd. Can., 34, pp. 379-383, 1977.

Bureau de la statistique du Québec, "Pêches commerciales", Industrie et Commerce, Québec, bulletin annuel, 1953-1983.

Caddy, J.F., "The influence of variations in the seasonal temperature regime on survival of larval stages of the american lobster (*Homarus americanus*) in the southern Gulf of St.Lawrence", Rapp. P.V. Réun. Cons. Int. Explor. Mer, 175, pp. 204-216, 1979.

Dow, R.L., Relationship of sea surface temperature to american and european lobster landings", J. Con. int. Explor. Mer, 37, pp. 186-191, 1977.

El-Sabh, M.I., "Surface circulation pattern in the Gulf of St.Lawrence", J. of Fish. Res. Bd. Can., 33, no.1, pp. 124-138, 1976.

Garrett, C. and B. Petrie, "Dynamical aspects of the flow through the strait of Belle-Isle", J. of Physical Oceano., 11, pp. 376-393, 1981.

Garrett, C. and B. Toulany, "Variability of the flow through the strait of Belle-Isle", J. of Mar. Res., 39(1), pp. 163-189, 1981.

Hunstman, A.G., Bailey, W.B. and H.B. Hachey, "The general oceanography of the strait of Belle-Isle", J. Fish. Res. Bd. Can., vol. 11, no. 3, pp. 198-260, 1954.

Lacroix, J., "Étude descriptive de la variabilité spatio-temporelle des phénomènes physiques de surface de l'estuaire maritime et de la partie ouest du golfe du Saint-Laurent à l'aide d'images thermiques du satellite NOAA-7", Mémoire de maîtrise, Dépt. d'Océanographie, Uni. du Québec à Rimouski, 185 p. + annexes, 1986.

Lavoie, A. Bonn, F., Dubois, J.M.M. et M.I. El-Sabh, "Structure thermique et variabilité du courant de surface de l'estuaire maritime du Saint-Laurent à l'aide d'image du satellite HCMM", J. Can. de Télédétection, vol. 11, no. 1, pp. 70-84, 1985.

Lavoie, A., Dubois, J.M.M., Lacroix, J., Royer, A. et M. Carignan,"Télédétection de la circulation des eaux de surface dans le nord du golfe du Saint-Laurent: application à la dérive de larves de poissons et de crustacés", CARTEL, Univ. de Sherbrooke, rapport 86-R-15, 97 p., 1986.

Lauzier, L., Trites, R.W. and H.B. Hachey, "Some features of the surface layer of the Gulf of St.Lawrence", Bull. Fish. Res. Bd. Can., no.111, pp. 195-212, 1957.

Malila, W.A. and D.M. Anderson, "Satellite data availability and calibration documentation for land surface climatology studies - Final report", ERIM report 180300-1-F, 214 p., 1986.

Messieh, S.N. et M.I. El-Sabh, "Répartition des températures au fond et à la surface, et circulation dans le golfe du Saint-Laurent", Cahier d'information no. 4, Dépt. d'Océanographie, Univ. du Québec à Rimouski, 139 p., 1979.

Mouchot, M.C. et E. Lambert, "Exemple de restitution de la température de surface de l'océan à partir de données AVHRR de NOAA-9", Centre Canadien de Télédétection, manuel de l'usager 86-1, 24 p., 1986

Murty, T.S. and J.D. Taylor, "A numerical calculation of the wind driven circulation in the Gulf of St.Lawrence", J. Oceanogr. Soc. Japan, 26, pp. 203-211, 1970.

Pêches et Océans Canada, "Les pêches canadiennes - Revue statistique annuelle", Direction de l'analyse économique et commerciale, Pêches et Océans Canada, bulletin annuel, 1968-1984.

Rose, G.A. and W.C. Leggett, "Atmosphere-Ocean coupling in the northern Gulf of St.Lawrence. Frequency-dependent wind-forced

variations in nearshore sea temperatures and currents", Can. J. Fish. Aqua. Sci., vol. 45, pp. 1222-1233, 1988.

Steven, D.M., "International biological programme study of the Gulf of St.Lawrence", Dans " 2nd Gulf of St.Lawrence Workshop", E.M. Hassan (coord.), Nov. 30 - Dec.3, 1970, B.I.O. Dartmouth, N.S., pp. 146-159, 1971.

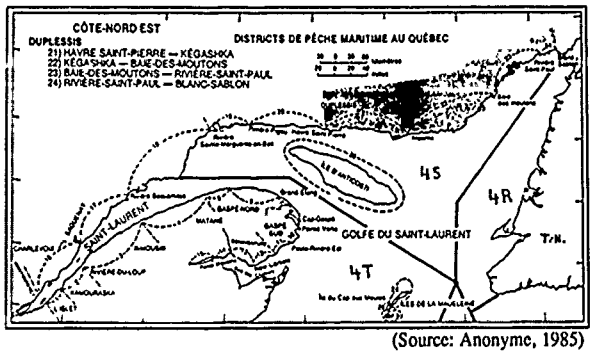
Sutcliffe, W.R. Jr., "Correlations between seasonal river discharge and local landings of American lobster (*Homarus americanus*) and Atlantic halibut (*Hyppoglossus hyppoglossus*) in the gulf of St.Lawrence", J. Fish. Res. Bd. Can., 30, pp. 856-859,1973.

Tang, C.L., "Observation of a wavelike motion of the Gaspé current", Jour. of Physical Oceano., vol. 10, pp. 853-860, 1980.

Therriault, M., "Analyse économique et commerciale du homard au Québec", Div. services économiques, Pêches et Océans Canada, 36 p., 1988.

Vigeant, G., "Température mensuelle de l'eau en surface dans l'estuaire et le golfe du Saint-Laurent", Div. des services scientifiques, Serv. de l'Environnement Atmosphérique, Environnement Canada, 34 p., 1987.

Weiler, J.D.M. and J.R. Keeley, "Monthly sea surface temperature for the Gulf of St.Lawrence", Marine Environmental Data Service, Fish. and Oceans Canada, technical report no. 7, 43 p., 1980.



(Source: Anonyme, 1985)

Figure 1 Les zones de pêche de la Côte-Nord (secteur ombragé). Limites provinciales en trait pointillé et limites selon l'Organisation des Pêches de l'Atlantique Nord-Ouest (O.P.A.N.O.) en trait foncé.

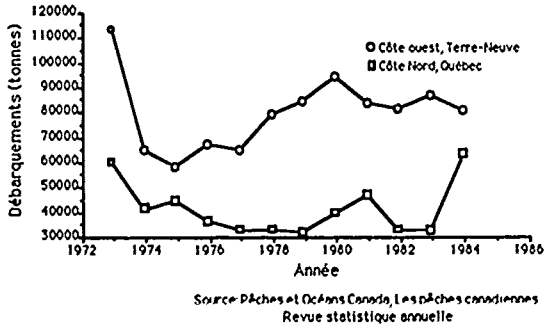
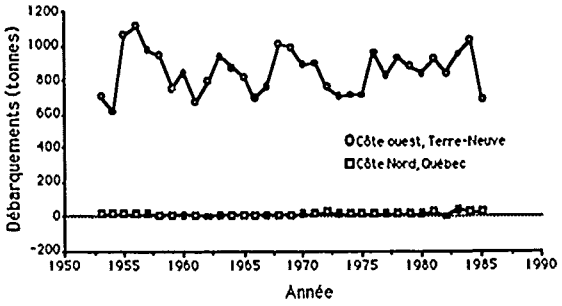


Figure 2 Débarquements totaux sur la côte ouest de Terre-Neuve (4R) et sur la Côte Nord et Anticosti au Québec (4S)



Source. - B S Q, Pêche commerciale
- Pêches et Océans, Les pêches canadiennes
Revue statistique annuelle

Figure 3 Débarquements de homard américain sur la Côte Nord au Québec et sur la côte ouest de Terre-Neuve

Tableau I
Liste des images NOAA retenues, température apparente de l'eau et conditions météorologiques générales pour la partie nord-est du golfe du Saint-Laurent

Image	Date jr/mo /an	Heure GMT	Température app. de l'eau Min	Max	Vent Vitesse km/h	Direction N-S-E-O	Condition barométrique sur le secteur
A	21-05-81	12 04	2,73	6,97	10-20	O	basse pression
B	22-05-81	11 41	2,18	6,97	15	N-E	haute pression
C	12-09-81	07 17	14,47	19,14	10	N-O	basse pression
D	17-09-81	11.56	10,59	17,05	0	-	basse pression
E	18-09-81	17 40	11,40	17,89	15-20	S-O	haute pression
F	22-09-81	18 35	9,59	16,19	20-40	O	haute pression
G	24-09-81	18 12	10,04	16,62	20-40	N-E	?
H	26-09-81a	17 50	10,04	16,62	30-65	N-O	basse pression
I	26-09-81b	23 16	6,97	13,65	20-50	N-O	basse pression
J	27-09-81	07:46	9,13	15,77	10-20	N-O	?
K	08-05-84	19.14	-0,13	8,16	10-25	O	haute pression
L	27-05-84	16 41	2,46	11,05	30-40	O-SO	?
M	24-06-84	19.38	5,49	12,98	15-25	S-SO	haute pression
N	25-06-84	19 25	5,49	12,99	15-25	S-O	haute pression
O	08-08-84	18 45	12,57	19,68	10-20	O	?
P	29-09-84	12:33	8,91	17,63	25-50	O	?
Q	30-09-84	19 34	3,52	11,98	15-30	S-O	haute pression

Tableau 2
Liste des éléments observés sur les images du secteur Natashquan à Saint-Augustin, accompagnés des références pertinentes

Thèmes à observer	Images	Ouvrages de référence
Gradient thermique (O - E)	E, F, I, J, K, N, O, Q	Vigeant, 1987, Weiler and Keeley, 1980, Lauzier et al, 1957, Messieh et El-Sabh, 1979
Remontées d'eau froide	F, H, L, P, Q	Hunstman et al, 1954, Bailey, 1956; Lauzier et al, 1957, Rose and Leggett, 1988
Contribution en eau froide de la mer du Labrador	B, K, M, N	Garrett and Petrie, 1981, Garrett and Toulany, 1981
Tourbillons Chenal d'Esquiman	J, K, L	El-Sabh, 1976, Murty and Taylor, 1970
Anticosti - Terre-Neuve	P ?	

Automated Spectral Recognition of Alteration Mineralogy in Imaging Data Using a Knowledge-Based Expert System

Tod D. Rubin

Department of Applied Earth Sciences
Stanford University
Stanford, CA 94305-2225
USA

Key Words: imaging spectrometry, expert systems, spectral recognition, alteration mapping

Abstract

Analysis of imaging spectrometer datasets is difficult because of the tremendous volume of data and the need for expertise in recognizing moderate- to high-resolution spectral signatures. In the knowledge-based expert system presented here, data are pre-processed with log residuals to remove topographic and atmospheric effects. Pixels from training areas are used to obtain spectral signatures of representative areas known to be dominated by certain minerals or alteration types. The spectra are then feature-coded using predetermined spectral windows based on a knowledge of the spectral features of the minerals or alteration zones. The result is a knowledgebase of the feature codes of representative minerals. Each pixel spectrum in the image is then matched against the knowledgebase of codes, and the pixel is assigned a color (represented by a single-digit number) corresponding to its alteration type, or black if there is no match. The resulting image-map of numeric codes corresponding to minerals is much smaller than the original multichannel dataset. The system was tested in the Yerington District of western Nevada, USA, and the image-map has excellent correlation with known geology.

Introduction

The purpose of this work was to develop techniques to automatically recognize and classify spectral signatures in imaging spectrometer data. There are two significant problems associated with such automated classification. The first is the tremendous volume of data contained in image data with high spectral and spatial resolution. The second is the specialized knowledge of mineral spectroscopy necessary to interpret moderate- to high-resolution spectra.

The approach taken was to utilize spectral coding in order to compress the data and to facilitate automatic interpretation. The coding scheme was designed based on a knowledge of the absorption features of the minerals of interest. A database, or knowledgebase, of the codes was assembled based on the minerals known to occur in the study area. After appropriate pre-processing, all the pixels in an image are codified. The expert system then searches the knowledgebase to match the pixel codes to the known codes. If a match occurs, the pixel is reassigned a one-digit value corresponding to a particular mineral and a color for display. Pixels which don't match any known code are assigned a digit corresponding to black. The colors in the resulting image correspond directly to mineral spectra in the knowledgebase.

Image data were obtained from the Geophysical Environmental Research Imaging Spectrometer (GERIS). GERIS measures 63 channels at wavelengths from 0.5 to 2.5 μm. There are 512 pixels per line; for this study, 800 lines were analyzed. The spatial resolution of the system as used at Yerington was 20 meters. Approximately 51 megabytes of data were required for the 10 km by 16 km study area.

The Yerington hydrothermal system, located 80 km southeast of Reno, Nevada, USA, is a Mesozoic porphyry system. Miocene extensional basin-and-range faulting rotated the deposit nearly 90°, resulting in a horizontal exposure of the originally vertical hydrothermal system. It was heavily mined for copper from the 1950's through the 1970's. Parts of the system have been mapped in great detail (Dilles 1983; Proffett and Dilles 1984). The excellent horizontal exposures of batholith-scale alteration zoning make this an ideal test area for imaging spectrometry.

Methods

I. Processing Sequence:

1. Pixel Averaging

Four pixels are averaged into one in order to increase the signal-to-noise ratio and to reduce the size of the image. While the image data are of very adequate quality for displays of individual channels, the success of the automatic spectral recognition algorithm described here was significantly improved by pixel averaging. An additional advantage is that the compressed display allows the entire 16 km length of the exposures at Yerington to be easily viewed on a single 400 line display.

2. Log Residuals

The data were transformed by log residuals (Green and Craig 1985) in order to remove atmospheric and topographic effects from the spectral data. In raw form, the spectra for all pixels appear very similar. The spectral shape is dominated by the absorption of solar radiation by atmospheric gases, primarily H_2O and CO_2, with major absorptions at 1.4, 1.9, and 2.7 μm. Overall spectral radiance of different pixels varies due to shading. The transformation consists of converting the data to logarithms, and subtracting from each log pixel spectrum: 1. the average log spectrum for the entire image, which accounts for the effects of atmospheric absorption and systematic instrument variations; and 2. the average brightness of that pixel in all channels (i.e. a pixel albedo), which accounts for differences in illumination between pixels. The differences are usually the result of shading by clouds or steep topography. The residuals which result from these transformations are then rescaled to a convenient range of values.

3. Feature Extraction

One of the most important, and most difficult steps in the processing sequence is the extraction of significant spectral features from the data. The features consist of absorptions, or areas of low log residual brightness, at wavelengths characteristic of certain target minerals. The human expert can usually distinguish anomalous data values (noise) from significant spectral information in an individual spectrum, but automating the process for large numbers of spectra is more difficult. Two approaches to automated feature extraction were tested.

The first approach relied on curve shape to detect absorption features. To be labeled as a minimum, a point in the spectrum had to be a local minimum (determined by the immediately adjacent points on the right and left), and the neighboring five points on either side also had to meet criteria for an appropriate range of slopes down toward the minimum point.

The second approach relied on deviation from average pixel brightness to detect absorption features. The average albedo of the pixel was computed, i.e. the average value for all channels of a pixel spectrum. Any channel (which was also a local minimum) falling more than a given threshold below the average was determined to be an absorption. Figure 1 shows the criteria for an absorption feature based on deviation form the spectrum average of a typical log residual spectrum in an area of sericitic alteration.

The results from both techniques were a list of the channels in each pixel spectrum which satisfied the criteria for being absorption features. The first technique was more appealing in that it mimicked the style of the human expert's interpretation of spectral curve shape. However, it was more computationally intensive than the second method, and was less successful at extracting significant spectral features. The second method, utilizing deviation of local minima from average pixel albedo, produced better results.

4. Feature Coding

Feature coding has been successfully applied to individual laboratory spectra (Yamaguchi and Lyon 1986; Goetting and Lyon 1986) The work here uses similar techniques applied to the thousands of spectra in an imaging spectrometer dataset. Each target mineral has a small number (usually from one to three) of significant absorption features measurable remotely by an imaging spectrometer. Depending on the number of target minerals and the spectral resolution of the measurements, the few features in any one mineral spectrum can occur in a limited number of wavelengths. These wavelengths are determined by the molecular structure of the minerals, and can be obtained from laboratory measurements or published spectra (e.g. Hunt and Salisbury 1970). Absorption features in imaging spectrometer data can vary over a range of wavelengths depending on mineral composition, mineral mixtures within the ground area covered by a pixel, instrument noise, etc. Thus the pre-defined wavelengths are actually wavelength bands, and can include more than one channel of imaging spectrometer data. For GERIS imagery at Yerington, only five wavelength bands were sufficient to characterize the significant absorptions. Figure 2 shows an expanded version of two pixel spectra, one from the sericitic zone and one from the advanced argillic zone, along with the five wavelength bands.

Each of the five locations was assigned a number from one to five. The code for a spectrum consists of a list of the numbers corresponding to those wavelength bands where absorption features occur. Absorptions which occur outside of the pre-defined wavelengths are not recorded in the code. The numbers are then arranged in order of decreasing absorption depth. The code for the upper spectrum shown in Figure 2 would be 4, while that for the lower would be 35. For comparison, ground spectra (with artificial illumination) are shown for the same areas in Figure 3. Artificial illumination allows the ground instrument to detect absorption features near 1.9μm.

5. Code Matching

The code for each pixel was then matched against a list of known codes in a database. The codes in the database were obtained from training areas as described below. Each code in the database was associated with a specific mineral or alteration type. If the code for a pixel matched one of the codes in the database, that pixel was then relabeled with a color code corresponding to that alteration type. If there was no match for the pixel, the pixel was assigned the color black. The resulting image consisted of pixels color-coded for specific alteration zones, and black elsewhere. (Each color is assigned an integer, e.g. sericite=red=1, alunite=blue=2, etc.) The color-codified pixels are a significant reduction in data size from the original pixel spectra. The original data consisted of 63 channels of 12-bit data for each pixel; the alteration mineral color code consists of a single-digit integer for each pixel.

II. Formation of the Knowledge Base

Parts of the Yerington District have been mapped and studied in detail (Dilles 1983, Proffett and Dilles 1984, Lyon 1987, Lyon and Rubin 1988). Certain areas known to be dominated by a single mineral or alteration zone were used as training areas to obtain representative mineral codes. The procedure consisted of visually scanning a list of the codes for all the pixels within the training area, and noting which codes were most common. In contrast to the usual use of training areas to obtain an "average" signature for the training area, no averaging was performed. Each training area, which consisted of between 10 and 100 pixels, could usually be characterized by just a few codes. In this way, the few codes formed an exhaustive list of all the different but significant spectral signatures which should be assigned the same color in the final output classified image.

The database of codes was stored in a separate file which could be easily edited. Editing allowed the user to change the assignment of codes to alteration types in order to obtain the image which best matched the known distribution of minerals and alteration zones. The database structure also allows separate commentary for each mineral code. The longer codes had the most specific or rigorous description of the original spectrum, and were therefore used for matching in preference to shorter codes. For example, a code of 3412 contains more information than the code 34, and a pixel matching 3421 should be assigned to that alteration type rather than the one associated with the shorter code 34.

Results and Conclusions

Imaging spectrometer datasets are very large and require expertise in spectroscopic analysis for interpretation. The expert system described here utilizes the knowledge of both mineral absorption features and the distribution of those minerals in field training areas to create a knowledgebase. The knowledgebase is then used to automatically interpret and classify the thousands of spectra in an imaging spectrometer dataset. The technique was tested in the Yerington District, western Nevada, USA, where there are excellent exposures of a hydrothermal alteration system which has been well studied with traditional and remote sensing/spectroscopic techniques. The classified images from the expert system agree very well with the known distribution of minerals and alteration types.

References

Dilles JH, The Petrology and Geochemistry of the Yerington Batholith and the Ann-Mason Porphyry Copper Deposit, Western Nevada, unpublished Stanford University PhD Thesis, Stanford, CA, 389p, 1983

Goetting H and RJP Lyon, A Knowledge-Based Software Environment for The Analysis of Spectroradiometer Data, Proc IEEE-IGARSS86, Zurich, Switzerland, pp 235-238, 1986

Green AA and MD Craig, Analysis of Aircraft Spectrometer Data with Logarithmic Residuals, Proc Airborne Imaging Spectrometer Data Analysis Workshop (ed by G Vane and AFH Goetz), JPL Publication 85-41, Jet Propulsion Laboratory, Pasadena, California, pp111-119, 1985

Hunt GR and JW Salisbury, Visible and Near-Infrared Spectra of Minerals and Rocks. I. Silicate Minerals. Modern Geology v1 pp283-300, 1970

Lyon RJP, Comparison of the 1984 and 1985 AIS Data over the Singatse Range (Yerington) Nevada. Proc Second Airborne Imaging Spectrometer Data Analysis Workshop, (ed by G Vane and AFH Goetz), JPL Publication 86-35, Jet Propulsion Laboratory, Pasadena, California, pp86-95, 1986

Lyon RJP, and T Rubin, Evaluation of Four Airborne Scanner Systems over the Ann-Mason-Buckskin Range Porphyry Copper-Geothermal System, Yerington,Nevada. Proc Sixth Thematic Conference for Exploration Geology, Environmental Research Institute of Michigan, Ann Arbor, Michigan, p237, 1988

Proffett JM and JH Dilles, Geologic Map of the Yerington District, Nevada, Map 77, Nevada Bureau of Mines and Geology, Reno, Nevada, 1984

Yamaguchi Y and RJP Lyon, Identification of Clay Minerals by Feature Coding of Near-Infrared Spectra, Proc Fifth Thematic Conference for Exploration Geology, Environmental Research Institute of Michigan, Ann Arbor, Michigan, pp 627-636, 1986

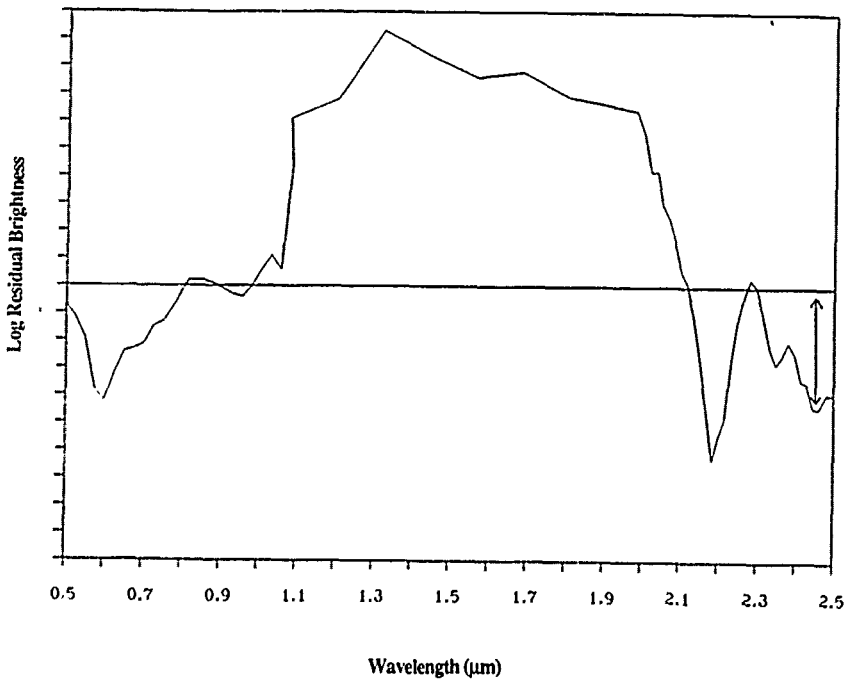


Figure 1. The horizontal line is the average of the spectrum, the arrow at 2.45μm shows the deviation below the average.

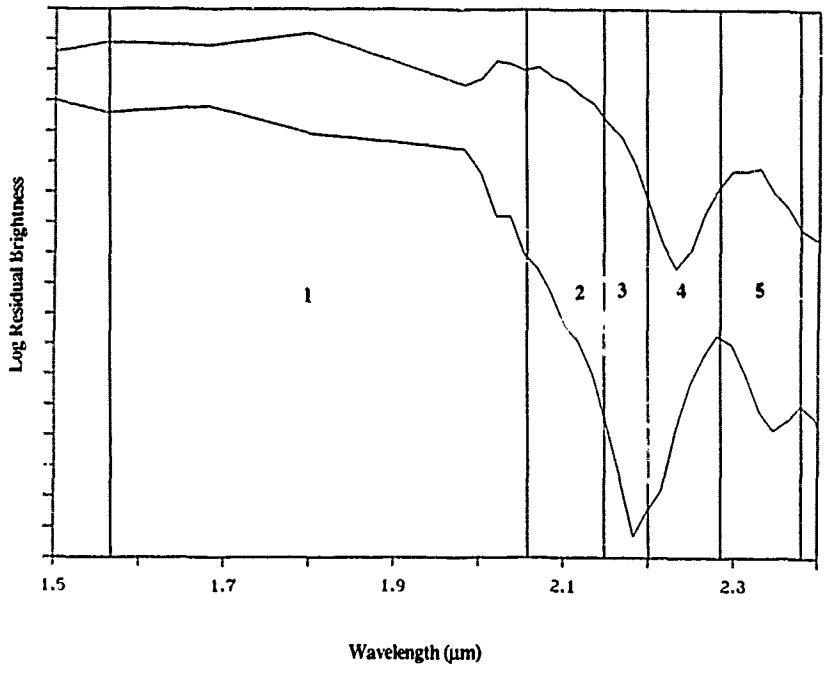


Figure 2. Expanded plots from a pixel in the sericitic zone and the advanced argillic zone. Numbers refer to the wavelength bands used for codification.

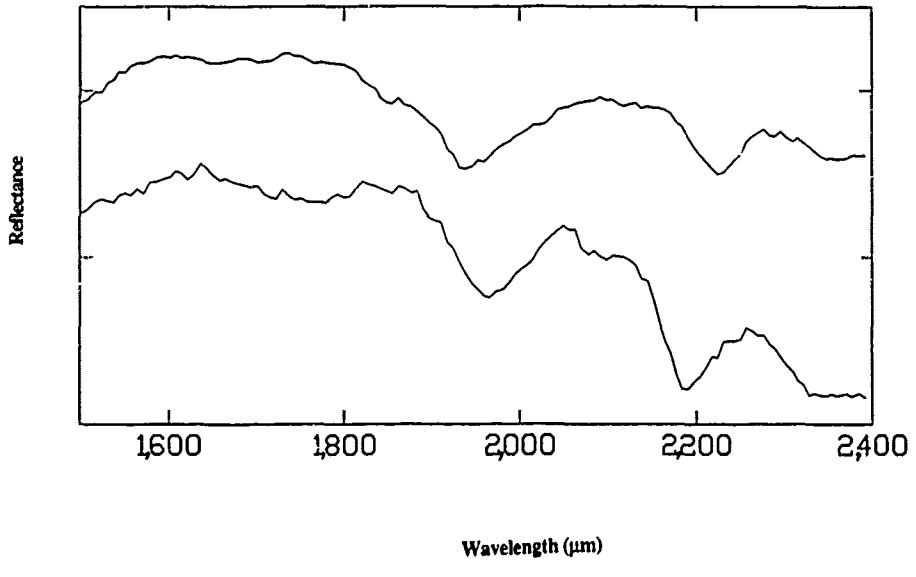


Figure 3. Ground-based reflectance measurements from the same areas as the pixels in Figure 2.

INVERSION OF IMAGING SPECTROMETRY DATA USING SINGULAR VALUE DECOMPOSITION

Joe W. Boardman

Center for the Study of Earth form Space (CSES)
CIRES, University of Colorado, Boulder, CO, USA 80309
Phone (303) 492-5051, FAX (303) 492-5070

ABSTRACT

The advent of imaging spectrometers, which acquire data that are both spectrally contiguous images and spatially contiguous spectra, promises to provide new capabilities for quantitative remote sensing of the earth. Such data sets cannot be analyzed fully using either existing spectroscopic or image techniques. New methods must be developed to take full advantage of the unique combination of spatial and spectral data. Singular Value Decomposition (SVD) is a powerful matrix analysis technique that has applications in spectral unmixing and determination of the spatial scales of mixing.

KEY WORDS: imaging spectrometry, spectral unmixing, singular value decomposition

INTRODUCTION

The developing technology of imaging spectrometry is providing new and promising data that can be used for geologic investigation of the Earth (Goetz and others, 1985). The data sets are at once both a spatially contiguous set of radiance or reflectance spectra and a spectrally contiguous set of spatial images. Such data sets are best envisioned as three dimensional cubes of values. Two dimensions are spatial, corresponding to the sample and line numbers of the pixel. The third dimension is spectral, indicating the wavelength or band number of the observed value. These unique characteristics of imaging spectrometry data are both the sources of its promise for new capabilities and the factors that necessitate the development of new techniques for analysis and interpretation of such data sets. Singular Value Decomposition (SVD) (Golub and Van Loan, 1983; Press and others, 1986) is a powerful tool used in matrix calculations and has several applications in imaging spectrometry analysis. Two such applications, spectral unmixing and determination of the spatial scale of spectral variance are described here.

Analysis techniques for imaging spectrometry data cannot be directly adapted from the existing methods used for either spectroscopic data or multispectral image data. The sheer number of individual spectra requires an automated and efficient method for extraction and full use of the spectral information. Similarly, the large number of spectral bands precludes a full utilization of the data in standard image processing environments. New analysis and interpretation techniques must be developed that take full advantage of this unique coupling of spatial and spectral information. The high spectral resolution common in such data sets encourages a shift from statistical and empirical interpretation techniques to more deterministic and quantitative ones. Individual mineral may be detected and identified based upon their diagnostic spectral responses (Kruse and others, 1989). However, the finite spatial resolution, typically tens of meters,of the current instruments results in each observed spectrum being a mixed spectrum. These mixed spectra are combinations of the spectra characteristic of the spectrally homogeneous components at the surface within the pixel. Quantitative interpretation of these mixed spectra requires an understanding of the modes and spatial scales of mixing of the spectrally homogeneous components in the area of interest. Additionally, one requires a method of inverting this mixing process to find the spatial abundance distribution patterns of the mixing components. The scales of spatial mixing can be examined and an unmixing technique developed using two different applications of SVD.

SPECTRAL UNMIXING

Spectral mixing results from the fact that the observed reflected electromagnetic radiation from a pixel has rarely interacted with a volume composed of a single homogeneous material. The spatial mixing of materials within the volume bounded by the pixel area and the depth of penetration is the cause of spectral mixing. Depending upon the geometry of this spatial mixing either a linear or nonlinear model may be used to describe the resulting spectral mixing. If a single incident photon is multiply scattered and encounters more than one material before escaping as reflected light then a nonlinear spectral mixing results. Pure linear mixing is the opposite case in which no single photon encounters more than one material. Models have been developed describing both the linear (Singer and McCord, 1979; Adams and Smith, 1986) and nonlinear (Hapke, 1981; Johnson and others, 1983) models for spectral mixing. Inversion of these

spectral mixing models has shown great promise in providing quantitative abundance estimates for the individual mixing components (Mustard and Pieters, 1987; Smith and others, 1985). In the nonlinear model the problem is linearized through a change of variables. So all of these techniques rest ultimately on a linear relationship between an observed, or linearized, spectrum and a library composed of spectra associated with the individual mixing components. Determination of the unknown abundance values is accomplished by inversion of the library matrix and multiplication by the observed spectrum. This is shown in a simplified form in Figure 1. SVD is a particularly useful method for inverting the library matrix as it allows insight into the resolving power and suitability of the library and control on the trade off between resolution and solution variance.

A is an M by N endmember library matrix

X is an N by 1 unknown abundance vector

B is an M by 1 observed data vector

M is the total number of bands
N is the number of mixing endmembers

$$A * X = B$$

$$X = A^{-1} * B$$

Figure 1. Forward and inverse linear spectral mixing models in matrix and vector symbols.

Singular Value Decomposition takes advantage of the fact that any matrix may be decomposed into the product of two column orthogonal matrices and a diagonal matrix. This decomposition is accomplished through bidiagonalization through by a series of Householder transformations and a subsequent diagonalization by QR shifts (Press and others, 1986). This is outlined in in Figure 2. Since the original matrix is now decomposed into the product of three orthogonal matrices, inversion of the library matrix is done by simply taking the transpose of the U and V matrices and using the reciprocal values in the W matrix. Thus an inverse of the original library is formed. A simple vector-matrix multiplication between this inverse library and the observed spectrum gives a least-squares estimate of the unknown endmember abundance vector. Singular Value Decomposition is a better method to use for inversion of the library matrix than Gaussian elimination since it can handle singular and near singular matrices. In addition it allows quantitative examination of the spectral separability of the library endmembers through an examination of the three matrices produces by the decomposition.

A is an M by N endmember library matrix

U is an M by N matrix of orthogonal columns

W is an N by N diagonal matrix of singular values

$\mathbf{V}^T$ is an N by N matrix of orthogonal columns

$$A = U * W * V^T$$

$$A^{-1} = V * 1/W * U^T$$

Figure 2. Symbolic representation of Singular Value Decomposition of the library matrix and the formation of the inverse library matrix.

Analysis of the values in the W matrix, the singular values, permits a determination of the spectral volume spanned by the mixing endmembers, measurement of their mutual spectral separability and analysis of the spectral distribution of information in the library. Through such studies estimates can be made of the unmixing accuracy possible for a given instrument and a given library (Goetz and Boardman, 1989). If the library of endmembers is completely spectrally separable, orthogonal endmembers, the normalized singular values would all be equal. For a wholly degenerate library all but one singular value would be zero, indicating that all of the endmembers are merely linearly scaled versions of each other. Real spectral libraries are intermediate to these two extreme cases. Small singular values come about as the result of spectral degeneracy in the library. In the inversion process the reciprocal values of these elements of the W matrix tend to have exceedingly high values. Including them in the inversion process contributes greatly to the solution variance but only adds slightly to resolution between the endmembers. When using SVD, an optimum balance may be struck between better resolution and less solution variance by damping some of the contributions of the smallest singular values. This may be done through either zeroing the large reciprocal values, to remove their contribution entirely, or multiplying them by a damping factor, to decrease their effect.

Application of the described spectral inversion techniques to real imaging spectrometry data sets involves choosing an unmixing library and the subsequent inversion of each spectrum. The library of mixing endmembers may be either spectra chosen from the data itself or laboratory and/or field spectra. This inversion process results in the entire data cube being mapped into an information cube whose spatial slices are endmember abundance distribution images. Figure 3 illustrates how the inverse mixing model is used to do this damping in an overdetermined case, where the number of endmembers is less than the number of bands used. In addition to the endmember abundance images other inversion products may be useful in improving the interpretation of the results. An error image may be produced by comparing the best fitting spectrum produced by the inversion to the original observed spectrum. This gives a pixel by pixel estimate for the accuracy of the inversion and the appropriateness of the chosen

endmembers. Also an average residual spectrum can be formed by summing the residuals for each pixel band by band and dividing by the number of pixels used. This spectrum can be used to help decide what additional endmembers could be added to the library to improve the inversion. Initial results using real data sets provide encouraging agreement between previously mapped and calculated distributions.

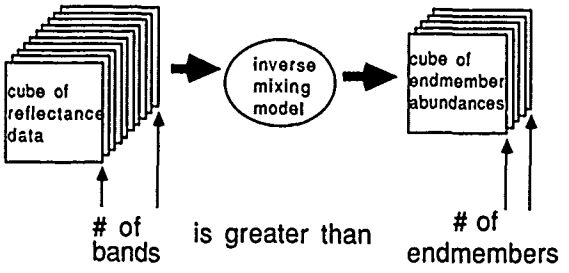


Figure 3. Processing technique for converting a cube of reflectance data into a cube of endmember abundance images through overdetermined spectral mixing inversion.

SPATIAL SCALES OF SPECTRAL VARIANCE

The scales of spatial mixing of spectral endmembers coupled with the spatial resolution of the instrument determine the degree of spatial oversampling or undersampling in an imaging spectrometry data set. If no spatial mixing occurs on scales less than 20 meters then a 20 meter spatial resolution will suffice for adequate sampling. The scale of spatially mixing of natural materials of geologic interest ranges from microns to many meters and is itself spatially variable. Through analysis of the data themselves a determination may be made of the spatial distribution of the scales of mixing for those scales greater than the pixel size. This is a new application of SVD to imaging spectrometry data analysis utilizing the unique combination of spectral and spatial information. It allows examination of the spatial distribution of spectral variance in both the spectral and spatial dimensions.

To map the spatial distribution of the spatial scales of spectral variance a suite of SVD operators is convolved with the data set. Square windows of dimensions 3 by 3, 5 by 5 and 7 by 7 pixels are used. For each of these windows the center spectrum and the surrounding spectra within the window are used to form a spectral library. SVD is applied to the library and the 9, 25 or 49 singular values are assigned to the center pixel as a singular value "spectrum". The sum of the squares of the singular values, representing the total spectral variance of the library, is also recorded. After this is done for all three window sizes the three total variance images are compared. Two different mean removal techniques are used to suppress variance associated with changes in overall albedo. Either each library spectrum is demeaned individually or the entire library formed by the spectra in a window is corrected by the mean of that library group. Figure 4 shows examples of the singular value "spectra" for the extreme maximum and minimum variances observed in a 63 channel imaging spectrometer data set using the 3 by 3 window. The maximum variance singular value "spectrum", the top curve, is from a region where significant spectral variability is found within the 3 by 3 pixel window, indicating spatial mixing on scales less than three times the pixel size. The minimum spectral variance singular value "spectrum", the bottom curve, is from a region of high spectral homogeneity on this scale.

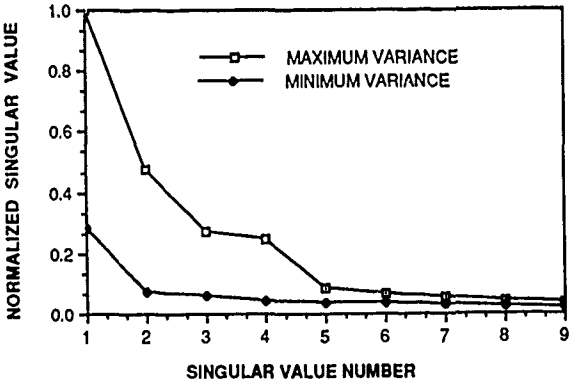


Figure 4. Normalized maximum and minimum variance singular value "spectra" for a 3 by 3 pixel window in a 63 channel imaging spectrometer data set of Cuprite, Nevada.

Analysis of the results of this procedure have shown it to be of use in outlining and identifying both regions of spectral homogeneity and regions of small scale spatial mixing. The dominant local scale of mixing, at superpixel scales, can be determined by observing a single pixel in the total variance images for the different window sizes. In the 7 by 7 window image it might appear with a high value , indicating it would be spatially undersampled if the pixel size was increased by a factor of seven. At the same time it may appear as a low value in the 3 by 3 window image, indicating that a pixel size three times larger than present would have sufficed to give adequate spatial sampling. Thus the local scale of spectral variance is between 3 and 7 times the pixel size. The entire processes described above can be done with spectral subsets instead of the entire spectral range. In this manner the spectral dependence of the degree of spatial undersampling can be determined.

CONCLUSIONS

Imaging spectrometry data represent a unique and valuable combination of spectral and spatial information. In geologic applications the most useful applications involve the quantitative mapping of lithologies and mineralogies. Since spatial mixing of geologically interesting components occurs on a wide range of scales, it is important to have accurate, flexible and informative techniques to invert the spectral mixing inherent in the data sets. Singular Value Decomposition provides a tool that can be used in both determining the scales of spatial mixing and the actual unmixing itself. When used to invert the mixing endmember library this technique allows for more insight into the library characteristics and allows more control of the inversion process than others

commonly used matrix inversion techniques. It can be used to show the mutual separability of the endmembers, to determine the spectral volume spanned by the endmembers and to provide a method for balancing the solution resolution and variance.

CITED REFERENCES

1. Goetz, A.F.H., Vane, G., Solomon, J.E. and Rock, B.N., " Imaging spectrometry for earth remote sensing", *Science*, v. 228, p. 1147-1153, 1985.
2. Golub, G.H. and Van Loan, C.F., Matrix Computations, Johns Hopkins University Press, Baltimore, 1983.
3. Press, W.H., Flannery, B.P., Teukolsky, S.A. and Vetterling, W.T., Numerical Recipes, Cambridge University Press, New York, 1986.
4. Kruse, F.A. , Kierein-Young, K.S. and Boardman, J.W., " Mineral mapping at Cuprite, Nevada with a 63 channel imaging spectrometer", Photogram. Eng. and Remote Sensing, v. 40, no. 4, *submitted*, 1989.
5. Singer, R. and McCord,T.,"Mars: large scale mixing of bright and dark surface materials and implications for analysis of spectral reflectance", Proc. 10th Lunar and Plan. Sci. Conf., p. 1835-148, 1979.
6. Adams, J.B. and Smith, M.O., "Spectral mixture modeling : a new approach to analysis of rock and soil types at the Viking lander 1 site", J. Geoph. Res., v. 91, p. 8098-8112, 1986.
7. Hapke, B.,"Bidirectional reflectance spectroscopy 1. theory," J. Geoph. Res., v. 89, p. 6329-6340, 1981.
8. Johnson, P., Smith, M., Taylor-George, S., Adams, J., "A semiempirical method for analysis of the reflectance of binary mineral mixtures:,J. Geoph. Res. v. 88, p. 3557-3561, 1983.
9. Mustard, J. and Pieters,C., "Quantitative abundance measurements from bidirectional reflectance measurements", Proc. 17th Lunar and Plan. Sci. Conf., J. Geoph. Res., v. 92, p. E617-E626, 1987.
10. Smith, M.O., Johnson, P.E. and Adams, J.B., "Quantitative determination of mineral types and abundances from reflectance spectra using principal components analysis", J. Geoph. Res., v.80,supp.,p. C797-C804,1985.
11. Goetz, A.FH. and Boardman, J.W.,"Quantitative determination of imaging spectrometer specifications based on spectral mixing models", IGARSS 1989, this publication.

ACKNOWLEDGEMENTS

This study was funded in part by NASA/JPL contract # 958039.

A HYBRID SYSTEM OF DATA ANALYSIS FOR IMAGING SPECTROMETRY

Don Jayasinghe and John Miller
Institute for Space and Terrestrial Science
4850 Keele Street,
North York, Ontario,
Canada. M3J 2K1
(416) 665 - 5406

Abstract

This paper describes an efficient method of image analysis suitable for imaging spectrometry. The method described is composed of most of the important image analysis techniques. It is divided into three stages through which data flows in analysis procedure. The stages taken in order are: statistical pattern recognition, structural pattern recognition and artificial intelligence. The efficiency of such an system, comprising of several stages lies in the fact that what one stage fails to accomplish is completed by another.

INTRODUCTION

Remote sensing technology has made great advances in recent years with the pioneering work of Goetz et.al. (1981), Collins et.al.(1983), Milton et.al. (1983) and Hollinger et.al. (1987), leading to the development of Airborne Imaging Spectrometer (AIS) in USA and Fluorescence Line Imager (FLI) in Canada. The necessity for such instruments became evident when it was discovered that utilization of high-spectral resolution radiometery offered the potential to directly identify surface mineral, detect vegetation stress induced by sub-surface minerals or observe in detail ocean processes. FLI was developed at Moniteq (Canada) Ltd. with the funding from The Department of Oceans and Fisheries, Ottawa, Ontario, Canada. This imaging spectrometer operates in two modes, spectral and spatial. In 'spectral mode' the device operates as 40 spectrometers, each with a resolution of 2.5 nm. In this mode it can also be thought of as a low spatial resolution, 40 element push broom scanner with 288 spectral bands. The location and with of the footprint of each spectrometer can be selected with the algorithm that forms spectral bands in spatial mode. In practice FLI's field of view and covers the ground, like the lines on a garden rake. This mode therefore sometimes refered to as a 'rake spect. ometer'. In 'spatial mode' the spatial resolution is high (0.7 mrad) providing 8 channel imagery in a manner similar to other airborne sensors but providing programmability in the selection of the bandwidth down to 2.5 nm and spectral position to anywhere in the spectral range 430 nm to 805 nm.

Since the development of AIS and FLI, several other improved imaging spectrometers have been designed such as Airborne Visible and Infrared Imaging Spectrometer (AVIRIS), which is an improved version of AIS, and accommodates the visible and infrared region of the spectrum. It has the spectral resolution of 224 chanels of each 20 nm wide and 550 pixels per each swath width. These imaging spectrometers (AIS, FLI,AVIRIS) produce a large volume of data per each flight line. The acquisition of this high volume of data poses a strain on conventional methods of image processing procedures.

Conventional image analysis methods such as maximum likelihood classifier or cluster analysis would take an enormous amount of computing time to process 1000 lines (i.e. 1000 x 40 x 288 x 2 bytes) of FLI data or about (1000 x 550 x 224 bytes) five times more of AVIRIS data. It would not be very efficient to use statistical methods of image classification using all the available spectral channels of data. Statistical classifications focus on the largest spectral differences between features as the basis of delineation. The spectral channels that will give the largest variations could be selected with a priori knowledge of the features in the scene. This method of selection of a few channels for statistical classification greatly reduces the computing time.

This paper outlines a methodology that would ease the burden on computer processing in classifying imaging spectrometer data. The main objective of this method is to employ all the available image processing techniques effectively and efficiently in extracting the maximum amount of information from imaging spectrometery. The final result one would expect from analysis of remotely sensed data is a correctly classified thematic map. This could be achieved by using the method discussed in the present work. This method is named Hybrid Image Analysis System, since it is a hybrid of all the available important image analysis methods. Fig.1, shows the data flow in hybrid system. The Hybrid Image Analysis System is composed of three major components:

1. statistical pattern recognition methods,
2. structural pattern recognition methods,
3. artificial intelligence techniques.

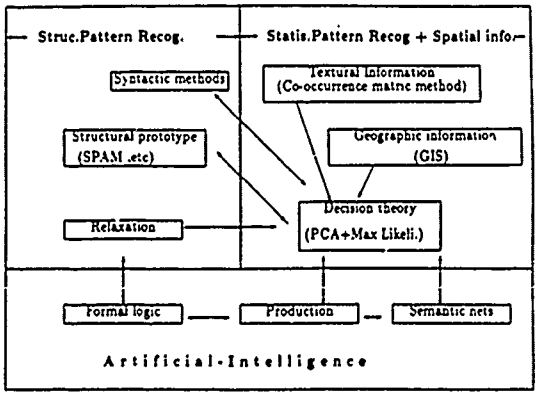


Figure 1: Data flow in a Hybrid Image Processing System

STATISTICAL PATTERN RECOGNITION

Statistical pattern recognition methods are based on statistical decision theory, very othern Baye's decision theory, and these methods of image delineation enable to establish the boundaries between classes or features. The imaging spectrometer data used this study comes from FLI imagery taken over Deep River area in 1987. Owing to differences in the responcivity of 5 CCD's used in imaging, the camera boundaries are visible in the image. These are eliminated by uniformity correction to a certain extend by not entirely. The camera boundaries could be eliminated by a histogram equalization procedure for each camera region separately, this method could only be applied when features covered in all five cameras are almost identical.

It has been shown by Haralick (1978), Rosenfeld and Kak (1976) and Bernstein (1978) that incorporation of relative terrain models and textural analysis improves the classification in multispectral imagery. In order to reduce the dimensionality of FLI (288 channels) 8 channels have been chosen, which show maximum variation for the features sought in the scene. These 8 spectral channels of FLI data, along with textural analysis of 550 nm channel with ranked elivation map of the area, is subjected to supervised maximum likelihood classification. This classification is carried out with the use of PCI EASI/PACE software installed in a Micro-Vax computer running on a VMS operating system.

The thematic map produced from the classification along with the original 288 channels are tranfered to SUN 3/60 computer running on a UNIX operating system for further analysis of spectral signatures fo te features delineated by statistical means.

STRUCTURAL PATTERN RECOGNITION

Structural pattern recognition can be further subdivided into three groups;

1. syntectic
2. structural prototype
3. relaxation.

In the present work only the structural prototype will be discussed. Structural pattern recognition methods are important mainly because they use the full compliment of spectral signatures to determine the differences between the pixels in the image. This is accomplished in the present work with the use of Spectral Analysis Manager (SPAM), a software package developed at the Jet Propulsion Laboratory of California, USA (Mazer,1988). Here, each spectral signature is binary encoded to reduce the computational complexities encountered in using 288 spectral channels. A classification criteria based on Hamming distance (Viterbi and Omura (1979)) has been employed in delineating the features.

The Hamming distance is a sum of bit-wise exclusive-OR operation, and it is very rarely that a perfect match does occur, as an allowance for natural variability, a specific threshold distance of acceptance have been devised.

ARTIFICIAL INTELLIGENCE

Artificial intelligence has come to existance as a method of problem solving with the use of computer and logical reasoning has also been built into it to assist in logical deduction. Programs have been developed to take discrete data structures and 'prove' assertions as formulas do in mathematical logic. Artificial intelligence is hybridized with structural pattern recognition methods as they exhibit functional similarities. The main difference between artificial intelligence and structural pattern recognition is that the latter recognizes patterns by matching its parts or the whole pattern to another while te former recognizes a pattern by logical reasoning. Artificial intelligence methods could be subdivided into formal logic, production system and simantic networks but this paper will discuss only the production system.

The Production system consists of three parts:

1. rule base,
2. context,
3. interpreter,

which controls the activity of the system. Here, all the knowledge about a spectral signature is captured in a series of IF-THEN rules,

e.g.

IF $(IR - G)/(IR + G) > thresholdlevel$
THEN (conclude it is a plant)

IR - infra-red channel
G - green channel

Examples of these rules are shown in Table 1. These rules allow certain conventions. The two types premise clauses used are: IFSOME and IFALL

The 'IFSOME' type represents a Boolean 'OR' while the 'IFALL' is equivalent to a Boolean 'AND'. If any clause following an 'IFSOME' specification is true then the action (THEN) clause will also be true. The 'IFALL' type requires all precises to be true if truth is to be assigned to the action clause.

Each parameter used in this production system is allowed a certain degree of freedom to fluctuate by specifying the threshold limits based on observations.

The 'context' in the present analysis is the spectral signature or the string of reflectance values which have been subjected to analysis. The segments of this spectral signature upon which the production rules are constructed have to be present for successful completion of the production system (Vron and Feigboum, 1981).

In rule 'interpreter' the basic control stratergy for rule interpretion is a hypothesis driven backward chaining search. The rules are verified or disproved by chaining backward to the initial data. All rules are applied to each spectal signature obtained from a particular feature to determine the homogeniety of the feature. The pixels that have been found to deviate within the allowable threshold limits for that feature, then such pixels are catagorized as subclass.

TABLE 1.

Rule 1.

RDB = wavelength values in data base
RMS = wavelength values for a given sample
r = wavelength r

IFALL (RDB(r) = RMS(r))
THEN try rules (2,3,4)

Rule 2.
wavelengths = 400,500,700 nm

IFSOME $(RMS(500) - RMS(700))/$
$(RMS(500) + RMS(700)) > 20\%\ RMS(400 - > 500) > 6\%$

$GradientRMS > 10 degrees$
THEN try rules (5,6,7)

Rule 3.
IFALL RMS (400 - 500)›6%
THEN try rules (8,9,10)

Rule 4.
IFALL RMS(400-500)›20%
THEN try rules (11,12,13)

There are a series of such rules which would be tested for each pixel to allocate it to one of the four main features, such as man made substances, water , vegetative material and rock or soil. Once, a spectral signature is assigned to one of the major classes, more rules are tested in order to allocate it into particular subclass.

DISCUSSION

The hybrid system described in this paper for computer analysis of imaging spectrometer data using statistical pattern recognition, structural pattern recognition and artificial intelligence methods is radically different to conventional methods of image analysis. Efficiency in this system lies in the way the three types of data processing techniques are linked together. If only a single technique is used in classifying the whole image it would take a greater time to complete the analysis than all three methods combined in the way discussed in this paper.

In statistical analysis of multispectral images, each feature is classified according to its behavior in spectral space, but it fails to consider the relationship between one spectral point toanother. The structural pattern recognition and artificial intelligence methods on the other hand take into account the variations or trends in the whole spectral signature. Thus making them more suitable in detecting finer changes in features such as stress conditions in plants.

REFERENCES

1. Avron Barr and Edward A. Feigenboum (1981), The Handbook of Artificial Intelligence Vol.I.
2. Bernstein,A. (1978),Digital Image Processing for Remote Sensing, New York: IEEE press 1978.
3. Collins, W. et.al. (1981), Infrared airborne spectrometer survey results in the Western Nevada area . Columbia University, Aldridge Lab. Appl. Geophysics, Final report to NASA, Contract JPL 955832, 61p.
4. Goetz, A.F.H., et.al. (1981), Geologic Remote Sensing, Science, Vol.211, p 781-791
5. Haralick,R. (1978), Scene Analysis, arrangements and hormorphisms, in Computer Vision ,Hansen and Riseman,Eds. NY press.
6. Hollinger, A.B.,et.al. (1987), FLI, Imaging spectrometer for land and ocean remote sensing, In Proceeding of the Society of Photo-optical Instrumentation Engineers, 834.
7. Mazer,A. and Miki Martin (1988), Image processing software for imaging spectrometry, SPIE, Vol.834 Imaging Spectroscopy II 1987.
8. Milton, N.M.,et.al.(1981), Remote detection of metal anomalies on Pilot mountain, Randolph county, North Carolina, Econ.Geol. Vol.78, p605-617.
9. Rosenfeld and Kak (1976), Digital Picture Processing, NY, Academic Press, 1976.
10. Viterbi, A.J. and Ormura, J.K. (1979), Principles of Digital Communication and Coding, McGraw Hill Book Co. NY,(1979).

COMBINING SPECTRAL AND SPATIAL MODE IMAGERY FROM AN IMAGING SPECTROMETER

J. Douglas Dunlop

Earth Observations Laboratory, Institute for Space and Terrestrial Science
Department of Geography, University of Waterloo
Waterloo, Ontario, N2L 3G1

Abstract/Resume

Imaging Spectrometers have the capability to collect both high spatial and high spectral resolution imagery at the focal plane, but then must trade off spatial against spectral resolution to reduce the data rate for digitizing and recording. Often data is collected in two modes which compliment one another during image analysis. *Spatial mode* provides the highest spatial resolution but has reduced spectral resolution while *Spectral mode* gives high spectral and low spatial resolutions. Although the imagery from the two modes can be analyzed as independent data sets, a much more effective methodology is described which unifies them into a single representation of the original spectral radiance field.

The methods described are applicable to most of the current or planned imaging spectrometers. Simulation results are shown to demonstrate the efficacy of the method. The conclusions include a discussion of how the design of imaging spectrometers could be affected.

Keywords: Imaging Spectrometer, sampling, reconstruction, image sharpening, Fluorescence Line Imager.

Introduction

In the 1990's the spaceborne imaging spectrometers will become operational and the number of spectral bands that are available to the satellite data user will increase by over an order of magnitude from what is currently available. The list of known and planned imaging spectrometers include HIRIS, HRIS, MODIS, MERIS, AVIRIS, ROSIS, AIS and FLI. Imaging spectrometers are nearly "ideal sensors" because they have high resolution in all domains; high spatial resolution, high spectral resolution and high radiometric resolution. The problem now becomes what to do with all this data. The potential data rate from an imaging spectrometer is horrific. For example the NASA imaging spectrometer HIRIS due for launch in 1995 is capable for generating 500 Mbits/s of data, which in more understandable terms is 38,571 9-track CCTs per day (recorded at 6250 bpi) of operation. Currently it is not technologically practical to continuously digitize, transmit and record at 500 Mbit/s data rates, even if all the data was desired.

All of the above imaging spectrometers can produce data at a rate between 50 and 500 Mbits/s. It is likely that only subsets of the data from these imaging spectrographs will be downlinked to the receiving stations or recorded. Colvocoresses (1977) and Schowengerdt (1980) suggested that perhaps a mixture of critically placed low spatial resolution narrow spectral bands and a single high spatial resolution band may be the optimal way to reduce the data rates with minimal information loss.

The strong correlations which will occur between spectral bands and adjacent pixels can be exploited to reduce the volume of data. This can be accomplished with imaging spectrometer data by averaging spectrally into spectral bands with full spatial resolution and simultaneously averaging spatially into larger pixels with full spectral resolution. Two distinct representations of the spectral radiance field result, each containing some unique information. These two data sets are useful in their own right, but combining them provides a more complete description of the original spectral radiance.

This paper proposes methods by which to average and sample the imaging elements of an imaging spectrometer in a manner which best facilitates reconstruction of the original radiance field.

Image Sharpening

A number of researchers have considered the problem of how to use information from and high spatial resolution band to sharpen the edges in other lower resolution bands. The idea of combining the data from one spectral channel to sharpen the edges in a second spectral channel was first suggested by Colvocoresses (1977) and first implemented when Schowengerdt presented his technique for restoring the high spatial frequency information to Landsat MSS imagery (Schowengerdt, 1980).

In 1980 Roller and Cox studied the possibility of improving classification accuracies and the interpretability of Landsat Multispectral Scanner (MSS) imagery by increasing its spatial resolution using Return Beam Vidicon (RBV) imagery (Roller and Cox, 1980). Hallada and Cox compared the methods of Schowengerdt and of Roller and Cox (Hallada and Cox,1983) using data from an airborne Daedalus 1260 multispectral line scanner. Hallada and Cox note that the combination of mixed spatial and spectral resolutions is very adaptable to the concept of imaging spectrometers. They state "On-board data compression would be a simple matter of integrating the signals across detectors in the spatial and/or spectral domains".

Simard was first to propose that the SPOT 20 m multispectral data could be sharpened by using the 10 m panchromatic band to modulate the brightness (Simard,1982). Cliche et al. describe algorithms for integrating the SPOT panchromatic band with the multispectral bands to improve image sharpness to yield results resembling colour infrared photography (Cliche, 1985).

Tom and Carlotto presented a Least Mean Squares (LMS) approach to edge sharpening (Tom, 1984). It relies on the fact "that at sufficiently small resolution multispectral bands are correlated", so if the digital levels of the any two bands are plotted against each other in a scatterplot then all the data will fall approximately along an straight line Any particular spectral band will generally not be globally correlated with the reference band (or bands), so the regressions are performed adaptively within a sliding window which moves across the image.

Sampling Theory for Imaging Spectrometers

Often when analyzing multispectral scanner data the analogy of image planes stacked on top of one another, but distinctly separated in space, is used to visualize the spatial and spectral relationships between the spectral bands. When using imaging spectrometer data we should refrain from thinking of 2-D sampling in distinct spectral bands but rather as sampling from a 3-D dimensional space where two dimensions are spatial dimensions and the third is spectral (Figure 1).

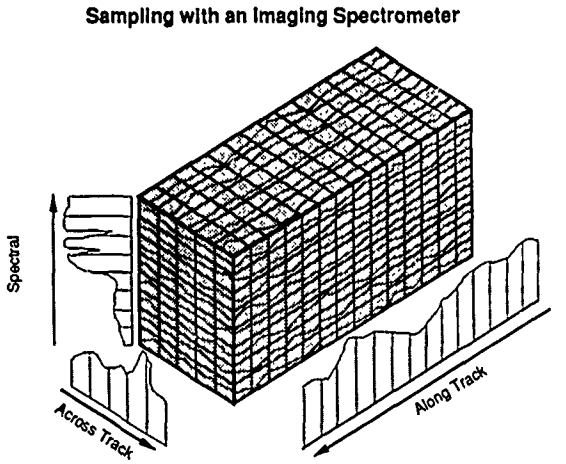


Figure 1. Sampling the radiance distribution with an Imaging Spectrometer in 3-Dimensional space.

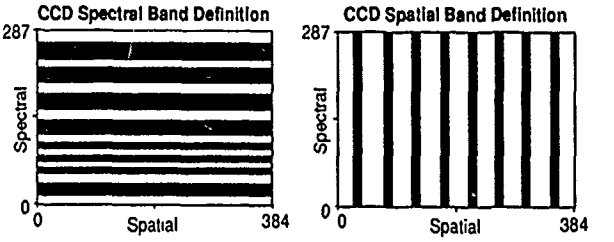
The image forming process can be thought of as a large square which is as wide as the sensor's swath and as tall as the spectral range. As the sensor's platform moves forward the square extrudes to become a parallelepiped which is as long as the flightline. Partition the parallelepiped into tiny cubes by dividing the width into the number of pixels per line, the height into the number of spectral elements and the length along track into the number of image lines. Contained within each cube is the spectral radiance from that small area of the scene. The imaging spectrometer's CCD array can be thought of as imposing just such a grid on the radiance distribution. Further, it will integrate the spectral radiance over the volume of each cube and quantize it to a radiometrically scaleable value.

The model presented here extends the concept of sampling into the spectral dimension and so what used to be spatial sampling in various spectral bands becomes spectral/spatial sampling. In the Fourier domain the distinction between the two modes of operation is well defined and may be clearly explained, but first the usage of some terms needs to be clarified.

Definition of Spatial and Spectral Modes

The terms spatial and spectral modes will be used to describe two generic methods of reducing the vast volume of data that is generated by an imaging spectrometer rather than to any specific sampling strategy. *Spatial mode* will refer to data which is recorded with the full spatial resolution of the spectrometer, but is averaged spectrally to a reduced spectral resolution. Conversely, the *spectral mode* data is recorded with full spectral resolution and is averaged spatially. Conceptually, it is also a vast simplification to neglect the along-track spatial dimension and consider a two-dimensional space which includes only the across track spatial and the spectral dimensions. In practice the along-track spatial resolution will likely be different in the two modes but including the extra dimension in the following discussion is an unnecessary complication.

As a specific example, Figures 2(a) & 2(b) show how the the Fluorescence Line Imager (FLI) defines spatial and spectral modes.



(a) (b)

Figure 2. Band definitions for spectral and spatial modes for the FLI

The rectangle represents the 2-D CCD array (288 x 385 elements) sitting in the focal plane measuring the entire spectrum of a single image line. The gray bands cover the groupings of elements that are summed into either spectral bands (Figure 2(a)) with full spatial resolution or into large pixels with full spectral resolution.(Figure 2(b)).

Fourier Domain Representation

By viewing how the data overlays the Fourier frequency space it is much easier to visualize how the two data sets can be combined in the regular domain and it elegantly demonstrates what information the combination data set *can* and *cannot* contain. Consider the discrete 2-dimensional Fourier Transform of the 2-D signal measured by an imaging spectrometer for a single image line. The name given to the Fourier spectrum of the spectral signal will be the *spectral frequency spectrum* and is analogous to the Fourier *spatial frequency spectrum*. In reference to the sampled radiance distribution, the amplitude of the spectral frequency spectrum will give the weighting required to reconstruct the spectral signal, at any given pixel location, by using a linear combination of sinusoids.

The spatial mode data has high spatial resolution and low spectral resolution and so represents full coverage of the spatial frequency spectrum but only covers the lower range of spectral frequencies (See Figure 3, "Spatial"). The spectral mode data is the converse with high spectral resolution and low spatial resolution, so

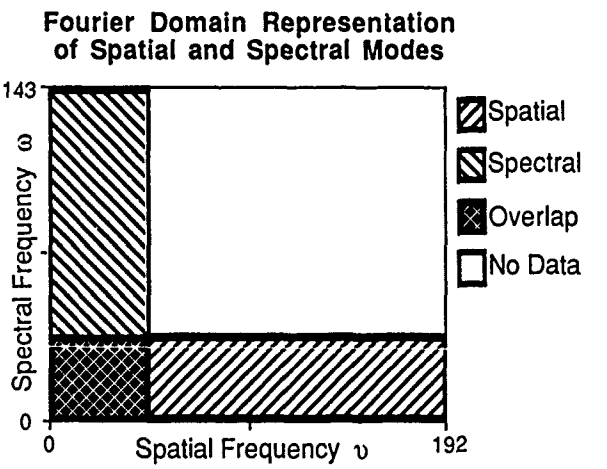


Figure 3. The Fourier domain representation of Spectral and Spatial modes. This shows sampling as the 2D Discrete Fourier Transform of the CCD array.

it covers the full spectral frequency spectrum but only the lower range of spatial frequencies (Figure 3, "Spectral"). As is evident in Figure 3, overlaying the two spatial/spectral frequency spectrum plots results in two regions which are uniquely sampled by the individual modes, a region that is sampled by both modes and a region that is not sampled at all. The region which is labeled *overlap* represents the low spatial and low spectral resolution image that could be formed either by averaging the spatial data in the spectral direction or the spectral data in the spatial direction. Thus the overlap region represents truly redundant data. The area labeled *no data* represents the portion of the signal that contains high frequency both spatially and spectrally. An example of this might be a fluorescence signal which appears in only a single pixel.

Signal Reconstruction

The ideal reconstruction filter will pass the areas of unique sampling with a weighting of unity, the redundant area will be an average of the two data sets and the area of no coverage would have zero weighting. Two points are particularly important to keep firmly in mind. First, there two data sets and so two interpolators are required because the sample spacings and sampling apertures are different. Second, the area of no-data represents an area which will contain unresolvable signal that will be entirely lost and cannot be recovered from the spectral, spatial or combined data sets. If this type of signal is present in the original radiance distribution it will be aliased in the reconstruction.

The algorithm used to reconstructed the signal is as follows:

1. Extract the redundant information from both the spectral and spatial data by lowpass filtering.
2. Subtract the spectral lowpass output from the original spectral data to get the low spatial frequency, high spectral frequency portion of the signal.
3. Subtract the spatial lowpass output from the original spatial data to get the low spectral frequency, high spatial frequency portion of the signal.
4. Average the spectral and spatial lowpass outputs.
5. Sum the results from steps 2, 3 & 4 with equal weights.

Simulation

A simulation using a spreadsheet program on a microcomputer was done to test the method explained above for integrating the spectral and spatial modes. To avoid dealing with any other characteristics of the sensor the simulation begins by generating the 2-D signal on the CCD array of a imaging spectrometer. The results are judged by how accurately they predict the digital level of each CCD element using only the values derived from perfonning spectral and spatial mode sampling on the simulated data. An array of 25 x 25 elements was used which were considered to produce 25 x 2 samples in either sampling mode. The horizontal dimension will be called X and the vertical dimension Y.

There are essentially three signal categories that need to be simulated. The first has all its energy within the area of overlap in Figure 3, is collected by in both modes. and should be the easiest to reconstruct. The signal must have frequencies no higher that 1 cycle/25 elements either spectrally or spatially. A signal which is uniform in one dimension (Y) and triangular shaped in the other (X) was used. The second comprises signals with either high spectral or high spatial frequency, but not both. The spatial frequencies must be less than 1 cycle/25 elements in one dimension and up to 1 cycle/2 elements in the other. The same signal as above was used except high digital level was inserted into a single column (all of Y). Third is the high spectral and high spatial frequency signal which not resolved by either mode. This will be any signal that has greater than 1 cycle/25 elements in both dimensions simultaneously. A uniform zero signal is used throughout the 2-D array with the exception of a single high digital level.

The results of the first test are shown in Figure 4 as a

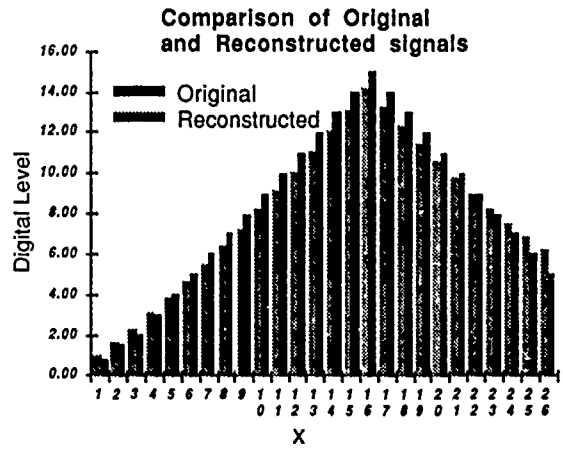


Figure 4. Simulation results from test case 1 with only low spectral and low spatial frequencies present in the signal.

comparison of the original signal to the reconstructed signal for one row of the array. The slope of the signal on both sides of the peak is slightly different, although the peak is accurately located. This is due to the averaging of the lowpass information in step 4 of the algorithm. The low-spectral/low-spatial frequencies are accurately predicted.

The second test results are shown in the same manner in Figure 5. Again there is a difference in the slopes of the triangular shape portion of the signal and there is a difference in amplitude between the original and the reconstructed spikes, but the position of the triangle peak and spike are accurate.

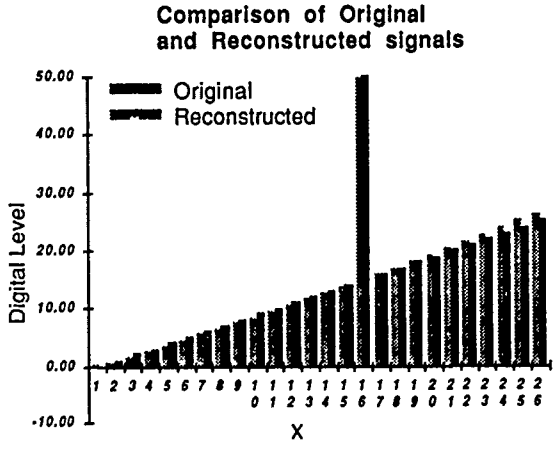


Figure 5. Simluation results for test case 2 with combine high and low frequencies.

Finally, results from the third test are shown in Figure 6 as a perspective plot of the entire reconstructed 2D array. This demonstrates the aliasing that will occur when the original signal contains Fourier frequencies which have not been sampled. This high-spectral/high-spatial-frequency signal has been aliased to both the high-spectral/low-spatial frequencies and the low-spectral/high-spatial frequencies and results in the cross pattern rather than the single spike. The original spike has an amplitude of 100 units and

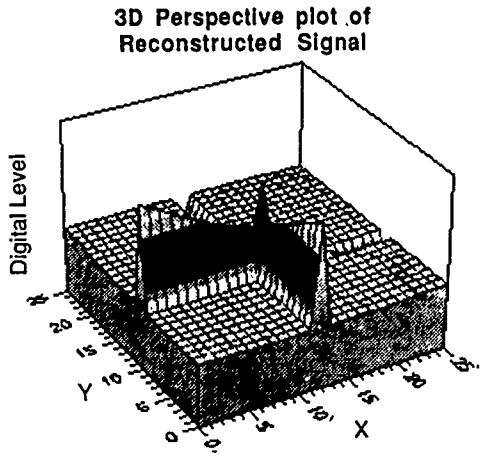


Figure 6. Simulation results from test case 3 with both high spectral and high spatial frequencies present in the signal.

was localized to the single array element at the center of the cross pattern of Figure 6. Other smaller fluctuations are present in the remaining elements although they are not readily seen.

Discussion

The spectral and spatial modes are mutually exclusive in the FLI and so must be collected on separate overpasses. Changes in the radiance distribution between the overpasses will affect the quality of the reconstruction and to combine them requires precision geometric correction (Buxton, 1988). For this method to be properly exploited then the spectral and spatial modes should be averaged from the same array of values so that they are inherently registered.

Although the information in the imagery is limited according to sampling and information theory, other sources of information can be exploited. Often the form of a spectral reflectance curve is well-known a priori and the imagery samples only serve to make subtle refinements. This is a form of interpolation which relaxes the conditions on the sampling strategy necessary to achieve accurate reconstruction. Samples can be placed at critical wavelengths to make the best use of the total data rate for acquiring new information. This is the standard technique used to select spectral bands, but put into the new context of signal reconstruction from imaging spectrometer data. This method cannot generally be used on the spatial samples, since the correlations tend to be predictable locally, but not globally, so no reliable model for interpolation other than from the samples will exist.

An important consideration when using this reconstruction technique is how much of the signal is lost irretrievably through averaging the CCD elements to produce the spectral and spatial modes. This could be computed from the full CCD array and expressed by the information density function which would give a accurate quantitative estimate of the size of the error expected in the reconstructed image data (Huck, 1985).

Conclusions

Reconstructing the full spatial and spectral resolution is not possible in the most rigorous sense. However, it is possible to combined the data from both spatial and spectral modes into a single integrated data set which may be easier to interpret. The low-spatial and low-spectral frequencies will be faithfully represented in the reconstruction. The low-spatial/high-spectral or high-spatial/low-spectral frequencies will be reconstructed with good results but aliasing from the high-spectral/high-spatial frequencies can occur. If high-spectral/high-spatial frequencies are present in the signal they will be aliased.

This reconstruction method will give good results if the energy in the signal in the high-spectral/high-spatial frequencies is a small fraction of the total power. In general use, the reconstruction will be useful if an error estimate can be predicted from the power in the aliased frequencies.

Acknowledgements

I would like to thank the people at Moniteq Ltd. for allowing me to work with the FLI which lead to this to this paper. I am indebted to John Miller and Allan Hollinger for reviewing much of the material presented. This research is supported, in part, by a Centre of Excellence grant from the Province of Ontario to the Institute for Space and Terrestrial Science, and an NSERC Operating Grant to E.F. LeDrew.

References

1. Buxton, R.A.H., "The FLI airborne imaging spectrometer: A highly versatile sensor for many applications", *Proc. ESA Workshop*, Frascati, pp 7-16, 1988.
2. Cliche, G., F. Bonn & P. Teillet, "Integration of the SPOT panchromatic channel into its multispectral mode for image sharpness enhancement", *Photogram. Eng. & Remote Sensing*, Vol. 51, pp 311-316, 1985.
3. Colvocoresses, A.P., Proposed parameters for and operational Landsat, *Photogram. Eng. & Remote Sensing*, Vol. 43, pp 1139-1145, 1977.
4. Hallada, W.A. & S. Cox, "Image sharpening for mixed spatial and spectral resolution satellite systems", *Proc. 7th Int. Symp. Remote Sensing Envir.*, pp 1023-1032, 1983.
5. Price, J.C., "Combining panchromatic and multispectral imagery from dual resolution satellite instruments", *Remote Sensing of Envir.*, Vol. 21, pp 119-128, 1987.
6. Roller, N.E.G. & S. Cox, "Comparison of Landsat MSS and merged MSS/RBV data for analysis for natural vegetation", *Proc. 14th Int. Symp. of Remote Sensing of Envir.*, pp 1001-1007, 1980.
7. Schowengerdt, R.A., "Reconstruction of multispatial, multispectral image data using spatial frequency content", *Photogram. Eng. & Remote Sensing*, Vol. 46, pp 1325-1334, 1980.
8. Simard, R., "Improved spatial and altimetric information from SPOT composite imagery", *Proc. Int. Symp. on Remote Sensing of Envir.*, 1982.

USE OF THE COMPACT AIRBORNE SPECTROGRAPHIC IMAGER (CASI): LABORATORY EXAMPLES

Gary A. Borstad, David A. Hill and Randy C. Kerr

G.A. Borstad Associates Ltd., 100-Marine Technology Centre
9865 West Saanich Road, Sidney, British Columbia, Canada V8L 3S1
Phone: (604) 656-5633 Fax: (604) 656-3646

ABSTRACT

As more and more remote sensing applications are found requiring high spectral resolution, there is a need for a small, easy to use instrument designed for experimental projects. The Compact Airborne Spectrographic Imager (CASI) is a newly developed second generation imaging spectrometer for use in small aircraft or in the laboratory. The CASI operates in the 423-946nm spectral range, with 288 spectral bands, has a variable field of view depending on the fore-optics, 12 bit dynamic range and 2.2 Gbyte recording capacity on digital cassette tape. System mass is 55 Kg and the size is about that of a personal computer. Power consumption is 250 Watts 110 Hz.

The system has a real time display of one spatial (image) band and one spectral channel and an interactive menu driven control. The 288 spectral channels are co-registered with a single 'track recovery image' which allows high accuracy in determining the ground location of spectral data. This paper describes the user interface during data acquisition, and reviews and discusses the philosophy of spectral data handling.

Keywords: digital imagery, spectrometer, spectral, spatial, CASI

INTRODUCTION

Canadian experience in non-imaging spectral resolution remote sensing (eg. Gower & Neville 1977, Gower & Borstad, 1981) led to the development of one of the first imaging spectrographs, the Fluorescence Line Imager (FLI) built for the Canadian Department of Fisheries and Oceans by Moniteq Ltd. of Toronto and Itres Research Limited of Calgary (Borstad et al., 1985). The FLI has been successfully utilized for many applications ranging from phytoplankton surveys in the oceans (Gower et al., 1986) to imaging missions over forested areas (Rock et al., 1988), and is now operated commercially by Moniteq Ltd. Building upon their considerable experience with Charged Coupled Device (CCD) imagers (Anger et al., 1987), and on Borstad Associates Ltd. field experience with airborne remote sensing and analysis of imaging spectroscopy data from the FLI, Itres Research Limited has built a second generation imaging spectrometer called the Compact Airborne Spectrographic Imager (CASI).

DESCRIPTION OF THE COMPACT AIRBORNE SPECTROGRAPHIC IMAGER

Like other pushbroom sensors, CASI and other imaging spectrometers view their target through a slit and build up an image via successive readouts as the aircraft moves forward. Unlike other sensors which use several separate sensors or linear arrays, with optical filters to obtain spectral bands, the CASI utilizes diffraction gratings and a two-dimensional Charged Coupled Device (CCD) array. The CASI uses a UT104 P9600 CCD Frame Transfer Array manufactured by the English Electric Valve Company of Chelmsford, England.

The array is illuminated so that across-track spatial information falls along its long dimension while spectral information registers across the shorter dimension providing a spectral resolution of 1.8nm over the spectral range of 423-946nm. Of the 612 cross-track array elements, 578 are illuminated, receiving spatial information emanating from the target, while the remaining elements record the dark signal and other diagnostic data. Table 1 summarizes the CASI's specifications.

The CASI records data in two fundamentally different fashions, referred to as Imaging Mode and Multi-Spectrometer Mode. In Imaging Mode the CASI acquires spatial data similar to other pushbroom imagers with the exception that the spectral band widths, positions, and number are programmable. The Imaging Mode data is translated from video tape to files which are transferred to an image processor for conventional image processing.

In Multi-Spectrometer Mode each of the 578 illuminated along-track columns of the array are literally separate spectrometers. The CASI can do a full frame readout but this is a slow process due to array read time and tape write time. For practical airborne operations the CASI limits the number of spectral columns to 39. The viewing direction along the swath and the number of spectrometers can be selected interactively by the operator.

A unique feature of the CASI Multi-Spectrometer Mode is a co-registered monochromatic spatial image acquired at the same time as the spectral data. This image, termed the 'track recovery image', is a high spatial resolution monochromatic reference image which permits the spectral data to be accurately located with respect to targets on the ground (Borstad and Hill, 1989).

INSTRUMENT OPERATION

Instrument control software allows the operator to set up the instrument by selecting the operating mode, band configuration, and integration time as well as view the incoming data in real time. In Imaging Mode the real-time display shows a black and white or colour image in one (user selectable) band. A band intensity plot showing the maximum and minimum data value of all spectral bands for each scan line is also displayed. This plot allows the operator to monitor each band for saturation. If saturation occurs the software allows the operator to change integration time or modify the spectral band width. In Multi-Spectrometer Mode the 'track recovery image' is also displayed in black and white or colour. Along the base of this image are ticks which indicate the positioning of the spectral channels. A user selected spectral look direction is monitored and for each scanline a spectrum is displayed. This allows preliminary verification of the spectral data in-flight. The CASI or an MS-DOS PC-AT, equipped with software similar to the instrument control software, can be used to review the data after a flight with software which is functionally identical to the real-time display.

DATA TRANSLATION FROM HDDT TO PC DISK FORMAT

CASI data is written on 8mm Video Tape in high density digital format, using a Exabyte Helical scan tape drive manufactured by EXABYTE Corporation, Boulder, Colorado. Each tape can record up to 2.2 gigabytes of data, approximately 1 hour recording time. Data is translated to LOGSWG format using specialized software which reads the Exabyte drive and writes MS-DOS files on hard disk. At present single band files are written out in either 8 or 16 bit format, but in the future we intend to write out pixel interleaved files.

DATA RATES

A major disadvantage of imaging spectrometers is the truly massive amounts of data they produce. The CASI has a 12 bit dynamic range and acquires 16 bits of data from each element in the array, a feature which allows simultaneous imaging of high reflectance land targets as well as dim targets such as deep water. A full array

scan will generate 352 Kbytes of data per scanline. A full 512 scene in this mode will generate 180 Mbytes of data. Even with the power of high speed work stations available for processing, the amounts of data produced by imaging spectrometers is a burden.

ANALYSIS OF SPATIAL DATA

Imaging Mode data from the CASI can be analyzed in a conventional manner using generic software on standard image processors. However the 8 bit precision available on most image processors can be a severe limitation. In our previous work with the Fluorescence Line Imager (Borstad et al., 1985), we had to write our own 16 bit software, especially for use with low radiance signals such as fluorescence. In this early stage of CASI processing, we are presently using divisions and subtractions to reduce the data to 8 bit for display. This process is not satisfactory for arithmetic operations and we are beginning to write our own 16 bit operators.

ANALYSIS OF SPECTRAL DATA

In special circumstances the CASI can acquire 288 spectral bands. The question arises of how to handle multi-spectral classification on this enormous amount of data. One answer is to determine the amount of spectral redundancy there is for a particular target and which bands best describe its spectral signature. The high spectral resolution capability of the Multi-Spectrometer Mode allows one to examine and analyze the spectral signature of a target. Analysis of this spectral signature allows determination of optimum spectral bands for use in acquiring high spatial resolution data.

In order to avoid congestion on our image processor and to allow us to carry out analysis in the field, we have developed specialized software that allows examination of spectral data on an MS-DOS AT computer. This software performs calibrations, arithmetic manipulations, stripcharts of spectral indices (defined as bands of widths and wavelengths of interest) versus time, plotting of scatterplots and histograms. Spectra can also be transported to and from LOTUS 123 files for modelling or manipulation.

To illustrate the use of the CASI we present a laboratory exercise that was conducted on April 19, 1989. This example could have just as easily been taken from the processing of airborne data. Spectral data of young, freshly picked Red Alder (Ainus rubra) and Balsam Poplar (Poplus balsamiferous) leaves were obtained outdoors under overcast illumination. Several leaves of each tree were arranged on a white vertical board, so that both the tops and bottoms were exposed. The board was then moved slowly past the CASI so as to sweep out an image.

In the first step of processing, we used the 'track recovery image', overlaid with the location of the spectral look directions, to locate a single look direction crossing an area of leaves in which the background was not showing through. Spectral data from that look direction were then displayed in image format beside the 'track recovery image' (Figure 1). This allowed us to look for small changes in the spectrum with time, and to see changes in parts of the spectrum other than that from which the 'track recovery image' was chosen. Average and standard deviation wavelength for specified times or scan lines were then extracted, divided by the spectra of the white board to form reflectance spectra. Division of the spectra from the top part of the leaves by spectra from the bottom show strong spectral variation in the vicinity of the red reflectance edge for both types (Figure 1). This may be significant where this feature is being used to detect plant stress. A comparison of spectra from the tops of the two types of leaves show that there are spectral differences which may be of use in discriminating the two tree species in airborne remote sensing operations (Figure 1). Some of these features are relatively narrow and will require the capabilities of imaging spectrometers or special optical filters for conventional devices. The beauty of imaging spectrometers like the CASI is that such laboratory studies can be quickly done to optimize spectral bands before flying. In total this exercise took two hours.

CONCLUSIONS

The Compact Airborne Spectrographic Imager is a flexible and portable instrument that can be used in both the laboratory and in the field. CASI's ability to act as a spectrometer and an imaging device at the same time greatly facilitates acquisition of high quality spectral data which are accurately located. CASI's flexibility in selecting and defining spectral bands allows it to be used to ground truth or select optimal spectral band definitions for other satellite and airborne systems.

REFERENCES

Anger, C.D., G.A. Borstad and S.K. Babey, "Progress in imaging spectroscopy", Paper presented at the 11th Canadian Symposium on Remote Sensing, Waterloo, Ont., June 22-25, 1987.

Borstad, G.A., H.R. Edel, J.F.R. Gower and A.B. Hollinger, "Analysis of test and flight data from the Fluorescence Line Imager", Can. Spec. Publ. Fish. Aquat. Sci., Vol. 83, 38 p., 1985.

Borstad, G.A. and D.A. Hill, "Using visible range imaging spectrometers to map ocean phenomena", Conference on Advanced Optical Instrumentation for Remote Sensing of the Earth's surface from Space, International Congress on Optical Science and Engineering, Paris, France, April 24-29, p. 7, 1989.

Gower, J.F.R. and G.A. Borstad, "Use of the in vivo Fluorescence Line Imager at 865nm for remote sensing of surface chlorophyll a". In: J.F.R. Gower (ed.) Oceanography From Space, Plenum Press, New York, pp. 329-338, 1981.

Gower, J.F.R. and G.A. Borstad, "On the use of the solar stimulated fluorescence signal from chlorophyll a for airborne and satellite mapping of phytoplankton", Paper presented at the 26th Plenary Meeting of Cospar, Topical Meeting A3 'Satellite Observations of Ocean Colour for Dynamic and Biological Studies', Toulouse, France, July 9-10, 1986.

Gower, J.F.R. and R.A. Neville, "A method for the remote measurement of the vertical distribution of phytoplankton in seawater", Presented to the 4th Canadian Symposium on Remote Sensing, Quebec City, May, 1977.

Rock, B. N. et al., "Comparison of in situ and airborne spectral measurements of the blue shift associated with forest decline", Remote Sensing of Environment, Vol. 24, p. 109, 1988.

Table 1
Compact Airborne Spectrographic Imager (CASI) Specifications

FIELD OF VIEW	10 to 70 degrees (depending on foreoptics)
SPECTRAL RANGE	423 to 946nm. (other ranges available)
SPATIAL RESOLUTION	512 spatial pixels (1m resolution at 833m altitude with 40.5° FOV)
SPECTRAL RESOLUTION	1.8nm.
SPECTRAL BANDS	Up to 288 selectable non-overlapping bands.
SCAN RATE	Up to 100 lines/second
DYNAMIC RANGE	12 bits
RECORDING	Digital cassette tape, 2 Gigabyte capacity, 190 KB/second maximum recording rate
OPERATOR CONTROL	Menu driven or from configuration files.
MASS	55 Kg
POWER	110V 2.5A

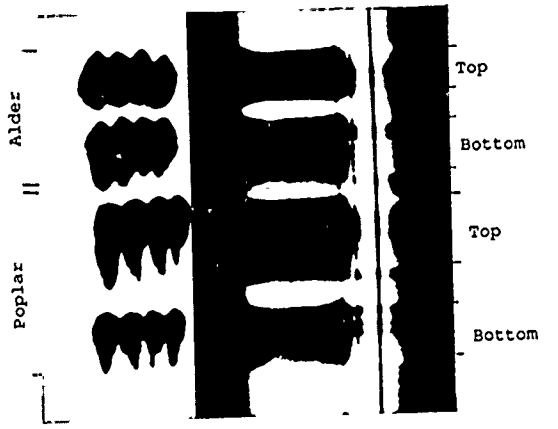
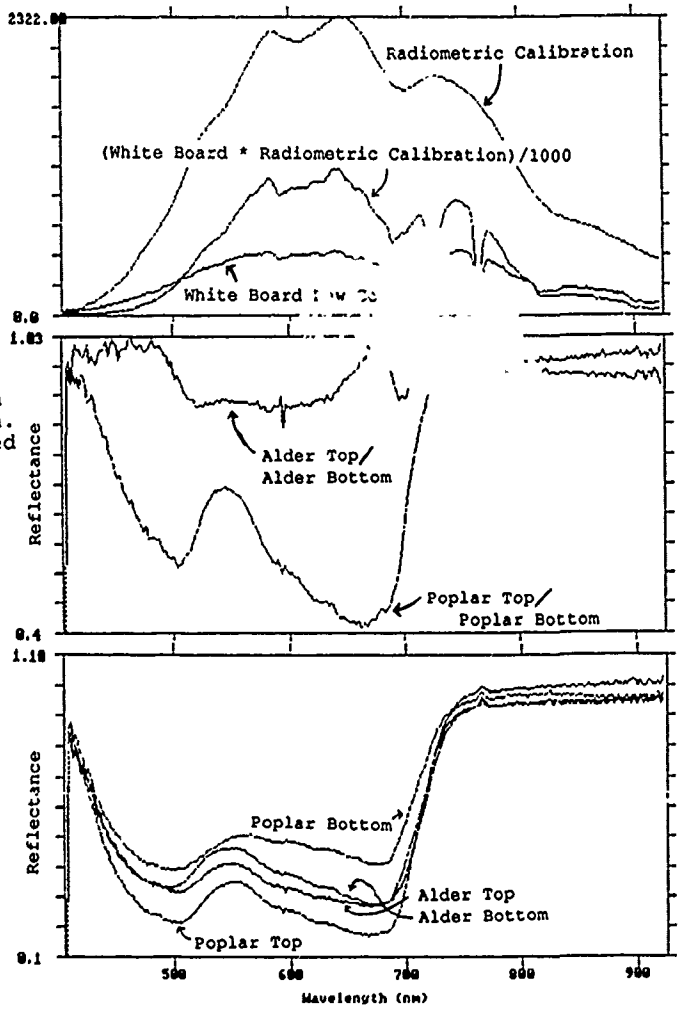


Figure 1: Track Recovery Row displayed with Spectral image of a selected spectral column. Plots are derived from pixel values extracted from indicated positions along the Spectral Image.



Changes In Canopy Leaf Area Index And Biochemical Constituents Of A Spruce Forest as Measured By The AIS-2 Airborne Imaging Spectrometer

C. Banninger

Institute for Image Processing and Computer Graphics
Graz, Austria

Abstract

AIS-2 data acquired from a Norway spruce forest in the 1900 – 2300 nm wavelength region at selected spectral bands corresponding to absorption features of lignin and nitrogen-containing protein correlated significantly with foliar lignin and nitrogen concentrations Although linear trends are apparent in the scatter-plots of combined AIS and biochemical data sets, the wide dispersion of the values resulted in rather low correlation coefficients. Needle nitrogen content showed the strongest correlation with AIS bands and canopy leaf area indices (LAI), whereas needle lignin content and canopy LAI values are poorly associated. Results from the study indicate that information related to canopy lignin and nitrogen content is obtainable from airborne imaging spectrometer data, but specialised information extraction techniques are required to properly analyse these data.

Key Words: Airborne Imaging Spectrometry (AIS), Lignin, Nitrogen, Leaf Area Index (LAI), Norway Spruce

Introduction

The ability to link canopy spectral absorption features to the concentrations of essential canopy biochemical constituents would greatly increase our ability to accurately assess the health and vitality of vegetation by remote sensing means. Infrared reflectance spectroscopy of plant material has identified numerous absorption bands in the shortwave infrared region (1200 – 2400 nm) that are associated with the harmonics and overtones of the fundamental bending and stretching frequencies of carbon-oxygen, carbon-hydrogen, oxygen-hydrogen, and nitrogen-hydrogen absorption bands and their combinations in the mid-infrared wavelength region. These absorption bands become progressively weaker at shorter wavelengths with higher orders of harmonics and overtones. Wessman at al. (1987a, b, 1988b) and Peterson et al. (1988) showed that subtle absorption features present in plant spectra in the 1200 – 1600 nm wavelength interval contain biochemical information on canopy lignin, nitrogen, starch, cellulose, and sugar content that could be estimated from high spectral resolution data acquired by NASA's Airborne Imaging Spectrometer (AIS-1). Scene noise reduction techniques (band differencing) and stepwise multiple linear regression analysis of AIS and canopy biochemical data sets were used to determine those wavelengths that provided the best ability for predicting canopy biochemical composition. These statistically defined wavelengths were then related to known absorption bands of organic compounds occurring in that wavelength region.

Foliar lignin and nitrogen concentrations are good biochemical indicators of stress conditions present in a canopy. By relating changes in the depth of laboratory-defined absorption bands of these two essential plant constituents to changes to their foliar concentrations, it may be possible to identify bands that give a good measure of their canopy content and, by inference, a measure of the health and vitality of the vegetation. The 1900 – 2400 wavelength region covered by the AIS sensor system contains information on lignin and nitrogen-containing protein that could be useful in the remote sensing of canopy vitality, if laboratory-derived lignin and nitrogen (protein) absorption bands can be shown to have a strong association with their respective canopy foliar concentrations.

Lignin and Protein Absorption Bands

A plant's health and vitality is dependent upon the uptake and metabolism of nitrogen-containing compounds to form chloroplasts, chlorophylls, proteins and amino acids. Chlorophyll pigments are contained within the chloroplasts, which comprise approximately 75 per cent of a plant's nitrogen content. Lignin is a major cell wall constituent of all vascular plants and increases in concentration with a decrease in plant nutrient availability.

The reflectance spectra of lignin and protein in the near to shortwave infrared region (800 – 2400 nm) are shown in Figures 1 and 2, with the wavelength positions of their major absorption features indicated Because of the amorphous nature of lignin and the difficulty in extracting it in its pure form from leaves (the lignin spectrum shown is of wood lignin), its actual spectrum may not be exactly as that depicted in Figure 2. Band widths for the major lignin and protein absorption features range from approximately 75 – 100 nm (measured as full band width at half maximum depth).

AIS-2 Data

The AIS-2 sensor system employs a 64 by 64 detector element array that measures the spectral region from approximately 800 – 2500 nanometres at a 10.5 nm sampling interval. Spectral data are

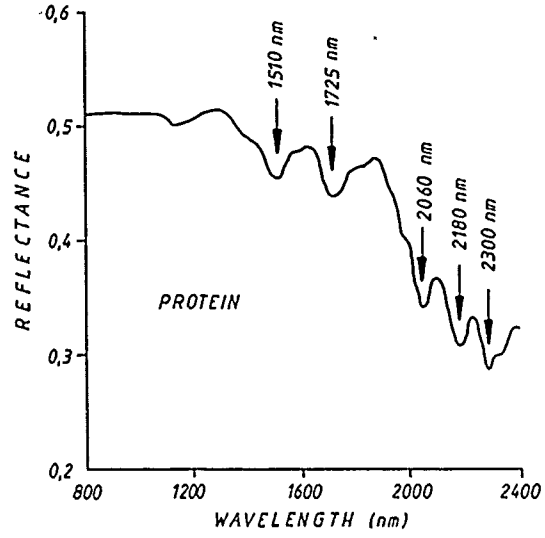


Figure 1 Protein Reflectance Spectrum (after: Norris and Barnes, 1976)

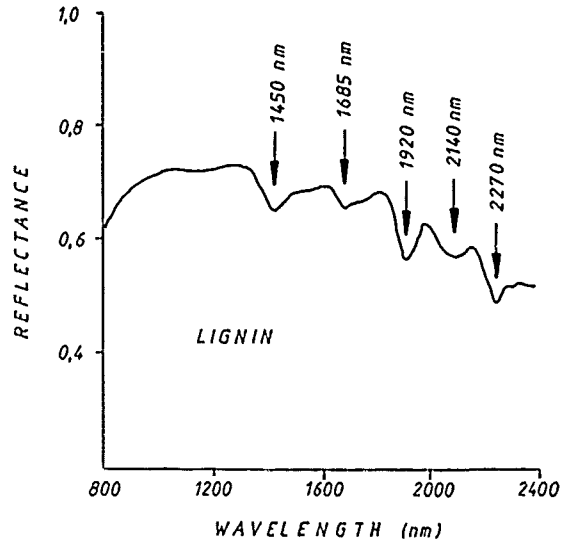


Figure 2 Lignin Reflectance Spectrum (after. Elvidge, 1987)

acquired in two separate, but overlaping wavelength regions. A so-called 'Tree Mode' set from 809 - 2143 nm and a so-called 'Rock Mode' set from 1184 - 2523 nm (Figure 3). Appropriately placed blocking filters eliminate contamination of the spectral data at longer wavelengths (up to 2400 nm) by lower-order diffracted radiation produced by the dispersion grating. Table 1 lists the pertinent information on the AIS-2 sensor system and the data collected from the test site. AIS-2 data acquired from vegetation canopies are dominated by the presence of major atmospheric water absorption bands. A logarithmic (log) residual transformation (Green and Craig, 1985) of the spectral data enhances the presence of weak spectral absorption features in the data, whilst suppressing or removing sensor, solar, and atmospheric-related scene effects, such as the two prominant water absorption bands in the shortwave infrared region (Figure 4).

Ground and AIS-2 Data

Needle samples collected from mature Norway spruce (Picea abies) trees at 50-m station intervals and AIS-2 spectral data acquired from the spruce tree canopy comprise the data sets used in the study. First year needles were excised from the twigs of lower tree branches, dried, ground, and analysed for lignin and nitrogen as a per cent weight by the infrared reflectance spectroscopy method (Wessman et al., 1988). Canopy leaf area index (LAI) values were derived from tree diametre measurements made from 5-m radius plots spaced every 50 m over the test site. AIS-2 data used in the study represent the mean intensity value of 2 by 2 pixel groupings acquired in both the 'Tree' and 'Rock Modes'. A log residual transformation of the data revealed subtle, but distinct absorption-like features in both 'Tree' and 'Rock Mode' data that corresponded to the wavelength positions of absorption features present in vegetation spectra obtained by laboratory measurements. The 10.5 nm spectral sampling interval employed by the AIS-2 sensor system is fine enough to resolve the major absorption bands of the more important plant organic compounds. Of particular interest were the 'Rock Mode' absorption features lying between 1900 - 2400 nm that showed a close correspondence to the major absorption bands of lignin and protein in this wavelength region (*see* Figures 1, 2, and 4). To clarify and better define the width and depth of these absorption features in the AIS data, a continuum is drawn to a smoothed spectrum and removed in the manner described by Clark and Roush (1984) and shown in Figure 5, with the maximum depth of the absorption feature measured from the removed continuum. A more detailed description of the AIS-2 data obtained from the test site and their processing is given in Banninger (1988).

The establishment of relationships between AIS spectral absorption bands and canopy lignin and nitrogen content and canopy LAI required the geographical positioning of each AIS canopy pixel to its corresponding test site location and the derivation of the average lignin, nitrogen, and LAI values from their derived isopleth maps for the 20 by 20 metre ground area covered by the 2 by 2 pixel window employed in the sampling.

Data Analysis and Results

A pairwise linear regression of the AIS and ground data sets helped to evaluate the usefulness of AIS-type imaging spectrometer data in measuring canopy lignin and nitrogen concentrations and canopy LAI. The results of the analysis are given in Table 2. The overall strongest correlations are between needle nitrogen content and nitrogen-related AIS-2 absorption bands and canopy LAI. The wide dispersion of the data points around the regression lines and resulting low coefficients of determination (r^2 values) show that most of the variance between the spectral and biochemical data sets is not explained by the regression parameters. Although the correlation coefficients obtained from the regression analysis are rather low, most are significant at least at the 90 - 95 per cent level and do indicate a distinct trend in the data.

Table 1 AIS-2 SENSOR SYSTEM

Number of Bands:	128 (contiguous)
Spectral Sampling Interval:	10.5 nm
Nominal Spectral Resolution:	21 nm
Spectral Range:	809–2523 nm
Tree Mode Range:	809–2143 nm
Rock Mode Range:	1184–2523 nm
Detector Array Size:	64 x 64 elements
Flying Height:	5000 m (a.g.l.)
Nominal Pixel Size:	10 m (nadir)
Nominal Swath Width:	640 m
Radiometric Range:	12 bits (summation to 16 bits)

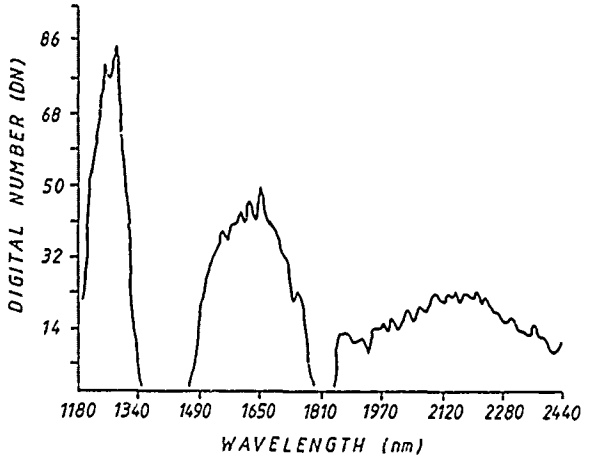


Figure 3 AIS-2 Rock Mode Spectrum from Spruce Tree Canopy

Discussion and Conclusions

The ability of remote sensing data to provide information on the health and vitality of plants by making use of spectral absorption features associated with foliar organic compounds offers the potential to rapidly survey and assess large areas of vegetation cover in an efficient and economical manner. This depends, however, on the premise that variations in these absorption features are related to changes in the chemical composition of canopy foliage at a level of accuracy commensurate with the required purpose(s) of the data. Biochemically related absorption features in the shortwave infrared region are weak and of a complex nature compared to the prominent water absorption bands of this region, and are composed of several overlapping absorption bands related to the overtones and harmonics of various leaf biochemical constituents. The mixing of multiple absorption bands causes a broadening of the absorption features and a shifting of their band centres relative to the absorption bands comprising the individual components. Although one might expect that overlapping absorption bands of biochemical constituents would respond in a similar manner to changes in a plant's condition and therefore reinforce each other, this is not necessarily the case, as is shown, for example, by the inverse relationship existing between foliar lignin and tannic acid content and foliar nitrogen content. The bandwidths of the lignin and nitrogen (protein) absorption features used in this study overlap with the absorption bands of between three and eight other leaf organic substances, thereby militating against the information content obtainable from them.

Wessman et al. (1987a, b, 1988b) used derivative spectroscopy techniques (band differencing) and stepwise multiple linear regression analysis of AIS-1 data between 1200 – 1600 nm to statistically select optimal wavelengths for predicting canopy foliar lignin and nitrogen concentrations. The wavelengths selected showed reasonably strong correlations with canopy chemistry, but were based on rather small sample sizes (n = 13 and 18), which could have unduly influenced the wavelength selection process (*see* Card et al., 1988). The selection of wavelengths by a statistical approach allows for absorption band mixing, but can result in a variability of the optimal band(s) selected, depending on the size and method of processing of the data, and difficulty in ascertaining the organic compound(s) responsible for the absorption band(s) chosen.

The deterministic approach used in this study does not account for the band mixing problem, but relies on the assumption that changes in the absorption feature being measured is dominated by the biochemical substance with that absorption band centre This may not, however, be true, if the interfering band(s) is closely centered to the absorption feature of interest and is also much stronger.

AIS data are quite noisy, due to both scene and sensor system related noise, and is not completely removed in the log residual transformation of the data. Its presence serves to further obscure the expressions of the already weak biochemical absorption features in AIS data Changes in canopy LAI also contribute to the variability in the canopy lignin and nitrogen estimations, as does the low range

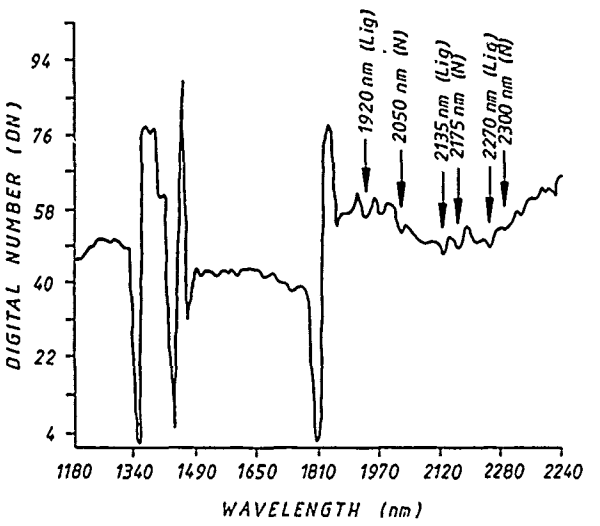


Figure 4 Log Residual Rock Mode Spectrum from Spruce Tree Canopy with Lignin and Nitrogen Absorption Bands Indicated.

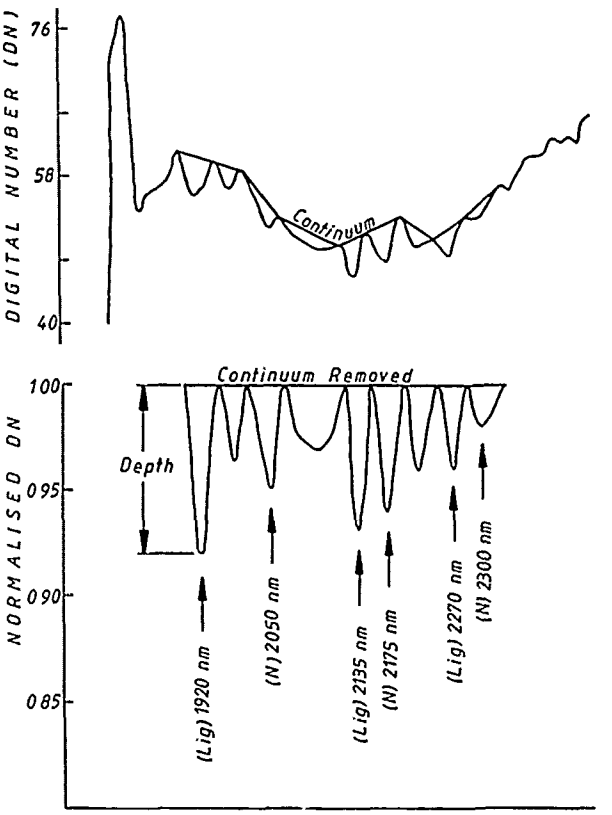


Figure 5 Removal of Continuum from Spectrum

of lignin (14.9 – 24.9 %) and nitrogen (0.748 – 1.270 %) values for the foliar samples analysed. The lack of significant correlations between the highest ranking nitrogen-based wavelengths (2050 nm and 2175 nm) with LAI values indicates that changes in these absorption features are likely related to foliar nitrogen content and not canopy LAI. Peterson et al. (1988) also found no significant correlation between LAI and AIS data in the 855 – 2036 nm wavelength interval examined.

The results of this study confirm the earlier work by Wessman and Peterson and their co-workers that information related to canopy lignin and nitrogen content is obtainable from AIS-type data, but that improved noise removal or reduction algorithms are needed, so as to allow for better extraction of the pertinent canopy information. Better information extraction techniques will also have to be employed in the analysis of AIS-type data to resolve the mixing problem caused by overlapping absorption features, either by the development of appropriate mixing models or the application of curve fitting and derivative techniques.

Acknowledgments

I thank Dr. John Aber of Complex Systems Research Center (Durham, New Hampshire) for performing the lignin and nitrogen analysis of the spruce needles and Dipl.-Ing. Wolfgang Pölzleitner of our institute for providing the statistic programs used in the regression analysis.

References

1. Banninger, C., 1988. Analysis of Airborne Imaging Spectrometer (AIS) Data for Geobotanical Prospecting, Proc. 6th Thematic Conf. Rem. Sen. for Explor. Geol., Houston, Texas, pp. 489-498.

2. Card, D.H., Peterson, D.L., Matson, P.A , 1988. Prediction of Leaf Chemistry by the Use of Visible and Near-Infrared Reflectance Spectroscopy, Rem. Sen. Environ., Vol 26, No. 2, pp. 123-147.

3. Clark, R.N. and Roush, T.L., 1984 Reflectance Spectroscopy: Quantitative Analysis Techniques for Remote Sensing Applications, J. Geophys. Res., Vol. 89, No. B7, pp. 6329-6340.

4. Elvidge, C.D., 1987. Reflectance Characteristics of Dry Plant Materials, Proc. 21st Int. Symp. Rem. Sen. Environ., Ann Arbor, Michigan, pp. 721-733.

5. Green, A.A. and Craig, M.D., 1985. Analysis of Aircraft Spectrometer Data with Logarithmic Residuals, Proc. 1st Airborne Imaging Spectrometer Workshop, Jet Propulsion Laboratory, Pasadena, California, pp. 111-119.

6. Norris, K.H. and Barnes, R.F., 1976. Infrared Reflectance Analysis of Nutritive Value of Feedstuffs, Proc. 1st Int. Symp. Feed Composition, Animal Nutrient Requirements, and Computerization of Diets, Utah State Univ., Logan.

7. Peterson, D.L., Aber, J.D., Matson, P.A., Card, D.H., Swanberg, N., Wessman, C.A., and Spanner, M., 1988. Remote Sensing of Forest Canopy and Leaf Biochemical Contents, Rem Sen. Environ., Vol. 24, No. 1, pp. 85-108.

Table 2 Correlations (r-values) Among Needle Lignin, Needle Nitrogen, Canopy Leaf Area Index (LAI), and AIS-2 Absorption Bands

n = 45	Lignin 1920 nm	Nitrogen 2050 nm	Lignin 2135 nm	Nitrogen 2175 nm	Lignin 2270 nm	Nitrogen 2300 nm	LAI
Lignin Content	0.26 **	–	1.10	–	0.27 **	–	0.14
Nitrogen Content	–	0.42 ***	–	0.34 ***	0.20 *	0.16	0.27 **
LAI	0.25 *	0.22 *	0.11 *	0.12	0.23 *	0.17	–

Level of significance: * $0.90 \leq p < 0.95$
** $0.95 \leq p < 0.98$
*** $p > 0.98$

8. Wessman, C.A., Aber, J.D., and Peterson, D.L., 1987a. Estimating Key Forest Ecosystem Parameters Through Remote Sensing, Proc. Int. Geosc. and Rem. Sen. Symp. (IGARSS'87), Ann Arbor, Michigan, pp. 1189–1193.

9. Wessman, C.A., Aber, J.D., and Peterson, D.L., 1987b. Estimation of Forest Canopy Characteristics and Nitrogen Cycling Using Imaging Spectrometry, Proc. 31st Int. Tech. Symp. Optical and Optoelectronic Appl. Sci. and Eng., S.P.I.E., San Diego, California.

10. Wessman, C.A., Aber, I.D., Peterson, D.L., and Melillo, J.M., 1988a. Foliar Analysis Using Near-Infrared Reflectance Spectroscopy, Can. J. For. Res., Vol. 18, pp. 6–11.

11. Wessman, C.A., Aber, I.D., Peterson, D.L., and Melillo, J.M., 1988b. Remote Sensing of Canopy Chemistry and Nitrogen Cycling in Temperate Forest Ecosystems, Nature, Vol. 335, No. 6186, pp. 154–156.

DETERMINING EXPERIMENTALLY INDUCED VARIATION IN CONIFEROUS CANOPY CHEMISTRY WITH AIRBORNE IMAGING SPECTROMETER DATA

[1]N.A. Swanberg and [2]P.A. Matson

[1]TGS Technology, Inc., MS 242-4, NASA Ames Research Center, Moffett Field, CA 94035, USA
(415) 694-5896, FAX (415) 694-4004, Telex 348408 NASA AMES MOFD

[2]NASA Ames Research Center, MS 239-12, Moffett Field, CA 94035, USA
(415) 698-6884, FAX (415) 694-4004, Telex 348408 NASA AMES MOFD

Abstract

Experimental treatments in a Douglas-fir forest in north-east New Mexico was carried out to determine whether differences in forest canopy chemistry could be detected using data from the Airborne Imaging Spectrometer (AIS-2). Experimental treatments consisted of nitrogen fertilizer additions, sawdust additions, and control plots.

After AIS-2 data were collected from the study site the digital number of a given pixel was extracted from each channel, yielding 128 values that were used to form a spectrum. Four spectra were extracted from each treatment plot.

Multiple stepwise linear regressions between first and second difference transformations of AIS-2 spectra and the canopy characteristics of biomass, nitrogen concentration, and nitrogen content were performed to determine which channels were most highly correlated with the measured canopy characteristics.

Results of a three term stepwise multiple linear regression showed a coefficient of multiple determination (R^2) of 0.71 between first difference AIS-2 spectra and measured nitrogen concentration in foliage. The results of this controlled study suggest that it may be possible to predict nitrogen concentration in Douglas-fir using AIS-2 spectra.

Key Words

Airborne Imaging Spectrometer, Canopy Chemistry, Nitrogen

Introduction

Chemical composition of forest canopies reflects ecosystem processes and biogeochemical cycling in forest ecosystems (Cole and Rapp, 1981; Birk and Vitousek, 1986). If the chemical composition of forest canopies could be obtained over large areas using airborne or spaceborne sensors, this information could serve as input to models of biogeochemical cycling and would allow estimation of biogeochemical fluxes on regional or global scales. The objective of this study was to determine whether experimentally induced differences in coniferous forest canopy chemical composition could be detected using data from the Airborne Imaging Spectrometer (AIS).

Study Site

This study was carried out in an even-aged stand Douglas-fir (Pseudotsuga menziesii (Mirb.) Franco) in north-east New Mexico near Mt. Taylor at an approximate elevation of 10,000 feet (Figure 1). The six experimental 25 m x 25 m plots were established by researchers in the School of Forest Resources at Northern Arizona State University. Replicate treatments included addition of 200 kg nitrogen fertilizer per hectare per year, addition of sawdust at 50 tons per hectare, and a control. The original plots were 12 m x 12 m and were treated in June of 1985. In May of 1986 additional treatment were applied to enlarge these plots to 25 m x 25 m for the purpose of this study.

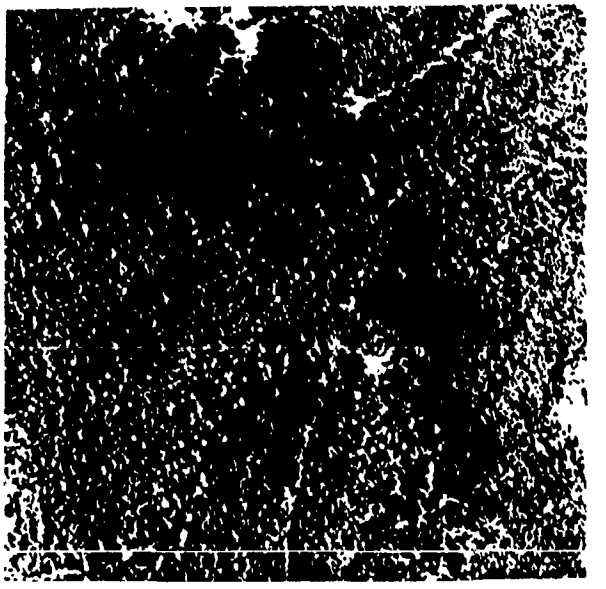


Figure 1. The study site.

Methods

On October 1, 1986, fresh foliage samples were collected from each plot. Nitrogen concentration in the foliage was measured after block digestion using a sulfuric acid/mercuric oxide catalyst. Leaves were freeze dried and ground with a 40-mesh Wiley Mill before analysis Lignin concentrations were measured using the method described in Van Soest and Wine (1968). Foliar biomass for each plot was obtained from allometric equations developed at Northern Arizona State University (Gower et al., in preparation). Canopy foliar content of a given constituent was calculated by multiplying foliar biomass by the concentration of the constituent for each plot.

AIS-2 data with a spatial resolution of 7 m were acquired over the study site with the NASA C-130 aircraft on September 19, 1986. These data were obtained in the "Rock Mode" with 128 channels, a 10.6 nm sampling interval, and a usable range between 1600 and 2400 nm (Vane, 1986). Black and white aerial photographs were also obtained.

Examination of the data revealed single drop out lines. Each pixel in these lines was replaced by the average of the corresponding pixels in the two adjacent lines. Vertical striping, striping in the along-track direction, was also observed. It was removed by applying a notch filter to the Fourier transform of the image in each band as described by Hlavka (1986).

The aspect ratio of these data was incorrect due to overscanning. The AIS had imaged the same ground area more than once giving the data a blurred appearance. The amount of overscanning was calculated as follows:

$$\frac{\text{altitude} \times \text{IFOV}}{\text{ground speed} \times \text{integration time}}$$

where:

aircraft altitude is in feet
aircraft ground speed is in feet/sec.
AIS instantaneous field of view (IFOV) is in radians
AIS integration time is in sec.

The result of this calculation was used to determine the scale to which to resample the image in the line direction using nearest neighbor interpolation. Resampling corrected the aspect ratio and produced a sharp image from which to locate the study plots.

Once each plot had been located, four pixels falling within the plot boundaries were identified. For each pixel, the digital number was extracted from each channel and these 128 values were then used to form a spectrum for each of the four pixels.

In order to determine the AIS-2 spectral position for this flight the spectra were compared to the reflectance spectra from the LOWTRAN 6 mid-summer, mid-latitude model for a hypothetical fifty percent isotropic reflector (Kneizys et al., 1983). The spectra were randomly divided into two subsets. Principle component analysis was performed on each subset to determine the number of degrees of freedom in the spectral data (Honigs et al., 1985). Results for one subset are shown in Figure 2. The rate of signal to noise in each channel of the AIS-2 data was estimated by calculating the ratio of the mean pixel value to the standard deviation in a fifty pixel homogeneous area of dense forest in the scene. Those channels with a signal-to-noise ratio of nine-to-one or lower were eliminated.

First and second difference transformations of the spectra (approximations of the first and second derivatives) were performed. A multiple stepwise linear regression between the transformed AIS-2 spectra and the canopy characteristics of biomass, nitrogen concentration, and nitrogen content (biomass x concentration) was performed to determine which channels were most highly correlated with the canopy characteristics of interest. The number of terms used in the regression equation was restricted to the estimated number of degrees of freedom in order to reduce the chance of overfitting.

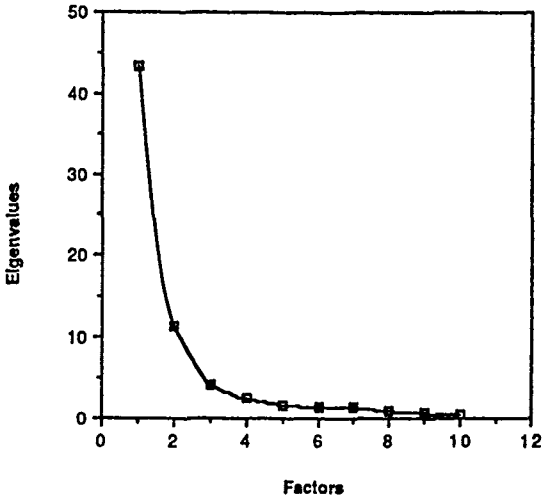


Figure 2. The results of a principle component analysis of a randomly chosen subset of 12 spectra from the study plots.

Results

Canopy biomass, nitrogen concentration and content, and lignin concentration and content for each plot are shown in Table 1. One and a half years following treatment, canopy nutrient concentrations varied among treatments, with fertilizer treatments having significantly higher nitrogen concentrations (14.43-15.69 mg/g) than did control (9.95-10.65 mg/g) or sawdust-amended plots (9.95-10.65 mg/g).

		Foliar Biomass		Nitrogen		Lignin	
				Weighted Avg Conc*	Content	Conc	Content
Treatment	Plot	kg/625m2	kg/ha	mg/g	kg/ha	%	kg/ha
Control	1	755	12080	9.86	119.17	9.73	1175
N Fertilizer	2	1048	16775	14.43	240.76	9.26	1553
Control	3	797	12763	9.59	123.26	9.19	1173
Sawdust	4	675.2	10803	10.64	115 54	9.21	995
Sawdust	9	779.9	12367	9.95	122.21	9.31	1151
N Fertilizer	10	809.7	12956	15.69	210.13	8.74	1132

Table 1 Canopy biomass, nitrogen concentration and content, and lignin concentration and content for each plot.

*Nitrogen concentrations are weighted by the percentage of needles in each age class in the canopy as calculated from the allometric equations. The three age classes of needles were current year, one plus two years old, and three years plus.

Nitrogen content showed the same pattern. Total foliar biomass was highest on the fertilized plots (809.2-1048.0 kg), while sawdust-amended plots ranged from 675.2 to 779.9 kg/ha and control plots ranged from 755.0-797.0 kg. No significant differences in lignin concentration were observed among the study plots.

After eliminating the AIS channels with a signal-to-noise ratio of less than 9-to-1, 68 channels remained. Results of a three-term stepwise multiple linear regression using AIS-2 bands centered at 1655 nm, 2025 nm, and 2115 nm showed a coefficient of multiple determination (R^2) of 0.71 ($P<0.01$) between first difference AIS-2 spectra and measured nitrogen concentration. Weaker correlations were found with nitrogen content ($R^2 = 0.65$, $P<0.01$) and biomass ($R^2 = 0.43$) which was not significant ($P>0.05$). Regressions between canopy lignin and AIS-2 response were performed, but were not significant as there were no significant differences in lignin concentrations between treatments.

Conclusions

The results of this controlled study indicate a significant relationship between AIS-2 response and nitrogen concentration suggesting that it may be possible to predict nitrogen concentration using AIS-2 spectra in Douglas-fir. A true test of prediction, rather than simply correlation, was not possible in this study as the experimental plots were too small to extract enough spectra for independent calibration and validation sets. A follow-on study is planned which will use large fertilized and unfertilized paired plots in a number of different coniferous forest types to determine if prediction of canopy nitrogen is possible and whether or not it is possible in other coniferous forest types.

Acknowledgements

We thank C. Berger for chemical analysis; C.C. Grier, F.T. Gower, and K. Elliot for setting up the treatment plots and for collecting the foliage samples; M. Baca and R. Thompson of the USDA Forest Service for providing maps and aerial photographs of the study site. We are grateful for efforts of J. Myers and C. Mahoney for decommutating the AIS data and for examining the aspect ratio problems. We thank the NASA/Ames Research C-130 crew and the Jet Propulsion Laboratory AIS team for their superb cooperation and performance in the collection of this AIS data.

Funding for this study was provided through the Biospherics Research Program of the Life Science Division at NASA Headquarters.

References

Birk, E.M. and P.M. Vitousek, "Nitrogen availability and nitrogen use efficiency in loblolly pine", Ecology, Vol. 67, No. 1, pp. 69-79, 1986.

Cole, D.W. and M. Rapp, "Elemental cycling in forested ecosystems", in Dynamic Properties of Forest Ecosystems, D.E. Reichle, ed., Cambridge University Press, Cambridge, pp. 341-409, 1981.

Gower, F.T., C.C. Grier, and K.A. Vogt, "Effects of nutrient and water availability of the carbon dynamics of Rocky Mountain Douglas-fir", (in preparation).

Hlavka, C., "Destripping AIS data using Fourier filtering techniques", Proceedings of the Second Airborne Imaging Spectrometer Data Analysis Workshop, JPL Publication 86-35, pp. 74-79, 1986.

Honigs, D.E., T.B. Hirschfeld, and G.M. Mieftje, "Near-infrared Spectrophotometric Methods Development with a Limited Number of Samples: Application to Carbonate in Geological Samples", Applied Spectroscopy, Vol. 39, No. 6, pp. 1062-1065.

Kneizys, F.X., E.P. Shettle, W.O. Gallery, J.M. Chetwynd, Jr., L.W. Abreu, J.E.A. Selby, S.A. Clough, and R.W. Fenn, "Atmospheric Transmittance/Radiance: Computer Code LOWTRAN 6", Air Force Geophysics Laboratory Environmental Research Papers, No. 846, pp. 1-200, 1983.

Vane, Gregg, "Introduction to the Proceedings of the Second Airborne Imaging Spectrometer (AIS) Data Analysis Workshop", Proceedings of the Second Airborne Imaging Spectrometer Data Analysis Workshop, pp. 1-16, 1986.

Van Soest, P.J. and R.H. Wine, "Determination of Lignin and Cellulose in Acid-detergent Fiber with Permanganate, J.A.O.A.C., Vol. 51, pp. 780-785.

RED EDGE MEASUREMENTS AND CANOPY STRUCTURE : A FIRST LOOK WITH AVIRIS DATA.

C. E. Leprieur

Laboratoire d'Etudes et de Recherches en Télédétection Spatiale,

18 Avenue Edouard Belin, 31055 Toulouse, France

ABSTRACT : Airborne measurements are evaluated with AVIRIS data, a modeling is given taking into account the atmosphere influence with the intention of quantifying red edge and red slope on vegetation sites, significance and value of the results are discussed further on. Comparaison is made with results acquired in situ and SAIL model simulation.
The behaviour of the vegetated sites concerning the relationship between the "Normalized Difference Vegetation Index" (NDVI) and the red edge position suggests that in some cases the percentage of cover is predominant and that in some other cases it is the "Leaf Overlap Index". To quantify this observation we simulated various canopy geometrical structure with the SAIL model and various types of leaves. With a standard soil spectrum, we found that for a "Leaf Area Index" (LAI) smaller than 3, the NDVI increases up to a given value (depending on the canopy) and then saturates while the red edge position decreases when the LAI varies from 0 to 1 and then increases with LAI. It has been noticed that when the NDVI saturates (LAI > 3) the red edge position is a linear function of the LAI.

KEYWORDS : *high spectral resolution, AVIRIS, vegetation.*

INTRODUCTION,OBJECTIVES

Most of studies concerning remote sensing and agriculture are based on the relationship between biological parameters and spectral data with the aim of using models (growth and yield prevision). Measurements performed on vegetation with broad spectral bands are generally related to phenology or biomass, the information being mostly present in the red and near infrared domains. This approach is justified if spectral characteristics are not modified in broad bands, this is not the case between 670 and 760 nm because one major feature of vegetation spectra is the steep slope between the low visible reflectance and the higher near-infrared reflectance : the "red edge".
Measurements performed on leaves with high spectral resolution laboratory spectrophotometers contain information correlated to vegetation properties : it has been shown that the shape of the red edge of leaves is correlated with chlorophyll content and fonction of the leaf area (Docray et al. 1981; Horler et al., 1983), it has also been noticed that the "red slope", slope of the leave reflectance curve between the maximum in the green (550nm) and the minimum in the red (670 nm) is related to chlorophyll content and composition (Gates et al. 1965).
These measurements take into account single leaf properties, and it is difficult to extrapolate laboratory results to in-situ fields measurements or remotely sensed imagery. The spectral behaviour of a canopy is notably different from that a single leaf because of its geometrical structure and the acquisition (Guyot, 1984; Verhoef W., 1984).

Few studies have been conducted on canopies in- situ or from space with narrow band spectrometers or spectroradiometers. Distinct differences have been detected on red edge position for peas and cucumbers spectra taking into account leaf stacking (Ferns et al. 1984). Phenological stages or different stresses on wheat canopies with identical "Normalized Difference Vegetation Index" (NDVI) have been discriminated (Baret et al. 1987). However studies based on in situ measurements are difficult to extrapolate to airborne sensors (Rock et al. 1988).
From these informations we investigated airborne spectral measurements acquired by AVIRIS, our main goals being :
– to determine a method detecting red edge positions and red slope values from space measurements, taking into account atmospheric parameters and sun position.
– to estimate the additionnal information given by narrow bands.
– to quantify the influence of selected factors on red edge positions from space measurements (sun position, soil response, canopy geometrical structure, leaf optical behaviour), the SAIL model (Verhoef, 1984) is used in the simulation part of this paper.

1) DATA DESCRIPTION AND METHOD

A – AVIRIS data description

The AVIRIS instrument consists in 4 spectrometers that discrimates a 20 square meters area into 224 spectral bands. A 614 pixels wide image is created by the scanner. The sampling interval is 9.6 to 10 nm, the spectral coverage is 400 to 2400 nm and the measured signal-to-noise ratio was 150 to 1 at 700 nm (Vane, 1987). After resampling, the number of bands is reduced to 210; each 9.8 nm wide. This being radiometric calibrated data supplied by JPL (Vane, 1987). The radiance values are in this step converted to 16-bit integers by multiplying each radiance value by 100. The data we have been working on were acquired over a large vegetated area near San Francisco at Moffett Field in summer 1987.

B – Test site presentation

The Moffett field test site includes San Francisco bay in the northeast corner, a transition zone between sea and land taken by some ponds limited by dikes and filled with salt water. Agriculture land is located near the airport, some golf courses and parks are present. The west half of the image is urban area, houses have watered grass and usually many trees. We selected 8 differents vegetated sites in this area.

C – Preprocessing and spectra description

We call R the "apparent reflectance" without atmospheric corrections defined by : $Ri=\pi.Li/Esi.\cos\Theta s$ where Li is the upward scattered radiance when the ground is illuminated by a solar incidence beam (Irradiance: Esi, incident zenith angle : (Θs), with i

being the band index). The integrated value of Es for each AVIRIS band is computed using linear interpolation. The data's unit is in microwatt/cm^2/nm/sr (x factor 100). They are then converted into apparent reflectance.
Presented in figure 1 are spectral plots of 7 sites. The wavelenghts of the AVIRIS channels, number [6...50], are in the [449.0, 880.2] nm spectral range.
The spectra look globally alike in the green and red domain. Levels of radiance are nearly the same except for site 4 whose response is high. Radiance generally increases from band 13 to 17 except for sites 5 and 7, and decreases for all sites from band 19 to band 27. In the near infrared, the plots show large differences in reflectance levels. Absorption peaks create radiance variations which are evident.
Of particular interest is the observation of the increase and decline patterns because they vary between sites. The slopes between the maximum in the yellow-green (band 16: 556.8 nm) and the minimum in the red (band 29: 674.4 nm) are slightly different. In spite of atmospheric absorption noise, differences among shape of spectra during the transition red/near infrared "the red edge" can be noticed.

D – Data processing

The shape of the reflectance spectra obtained on leaves and canopy in-situ exhibits an inflexion point between 670 and 760 nm, the position of which is usually determined by analytic methods (Gates, 1965). It has been shown that on vegetation reflectance spectra, the red edge has a local symetry center; the associated reflectance being the average of red (670 nm) and near infrared (760 nm) response (Guyot et Baret, 1988). According to these authors it is possible to approximate the reflectance curve between 700 and 740 nm, which is the usual range of variation of the red edge in laboratory or in-situ, with a straight line.
Using AVIRIS data, we plotted the reflectance curves beetween 600 and 850 nm to estimate whether the linear approximation is correct. We have computed the gaseous transmittance with the 5S model (Tanre et al., 1986) for each AVIRIS band (center wavelenght). This is represented on figure 2, with two models, midlatitude summer and midlatitude winter atmosphere.
Atmospheric absorption created obstacles to compute the red edge position : noise between channel 29 (674.4 nm) and 40 (782.2 nm) was a consequence of gaseous absorption : water in channel 34 and oxygen in channel 38 Hence we selected some points which are located out of the absorption bands and checked that the spectra was linear between channel 31 and 36.
To determine the local symetry center position (RE) we selected channel 29 and channel 40 respectively significant of maximum chlorophyllian absorption and near infrared vegetation response.
For each pixel with :
r_i : Reflectance in AVIRIS channel i
wl_i : wavelength center AVIRIS channel i

$$r_{(RE)} = (r_{29} + r_{40}) / 2$$

$$RE = wl_{31} + (r_{(RE)} - r_{31}) \times (wl_{36} - wl_{31}) / (r_{36} - r_{31}) \quad (1)$$

With the same notation we have computed the red slope value (RS) between yellow-green (channel 19) and red (channel 27), the atmospheric perturbations being very low in this spectral domain :

$$RS = (r_{19} - r_{27}) \times 10^2 / (wl_{27} - wl_{19}). \quad (2)$$

2) RESULTS AND DISCUSSION

A – results on AVIRIS data :

For each pixel, NDVI, RE and RS have been calculated; statistical description is performed for the eight selected sites (table 1).
Apparent reflectance values were also collected (bands 16, 29 and 40) in order to establish relationships with red slope and red edge values.

a) Red slope
The range of variation for the red slope is 3.03 ($\Delta\rho^*.10^2$) for all the sites and 7.10 for the whole image.
Modifications of the slope between green-yellow and red observed on figure 1 are quantified. Red slope tends to increase with NDVI on vegetated terrestrial sites, the behaviour of site 8 (aquatic plants) is specific.
The relationship between RS and spectral bands : Band 16 (547 nm) and Band 29 (674.4 nm) varied with sites suggesting small variations in canopy optical properties. Slight differences have to be related to $Ne\Delta\rho$ which is unknown.
b) Red edge
The range of variation of the red edge position was about 23 nanometers for all the sites. On the whole image the minimum value was 696 nm and the maximum 735 nm; so, we nearly used the possible range [695,740] nm.
We found that the relationship between spectral values (bands 29 and 40) and RE was non linear and varied slightly with sites.
In spite of the non-linearity, least squares linear regression analysis gave an indication of correlations between spectral parameters. There were several evident relationships.
On site 1 and 6 (high NDVI) correlations between RE and C29 (674.4 nm) were negative (-0.65, –0.53) and correlations between SY and C40 positive (0.60, 0.35); this suggested that the control of the red edge was partially exerted by leaf chlorophyll content and folliage scattering properties. However the standard deviation order of magnitude of band 40 on the sites induced a preponderant effect of the folliage scattering on red edge values (canopy structure).
relation with NDVI :
Optical properties of a canopy depend mostly of the LAI. NDVI being a good indicator of LAI, we have studied the relationship between RE and NDVI.
We have plotted RE as a function of NVI for each pixel of 5 vegetated sites (figure 3). RE increases with NDVI, the correlation being positive on each site in the (0.4,0.7) range. This evolution has already been described for wheat and related to phenology (Baret et al., 1988). Our sites are different, including various vegetation types; however we suggest that a commun factor has an influence on the two variables in our data set. We have studied the red edge behaviour, mechanisms which are involved in spectral shifts and the relationship with NDVI.

B – comparaison with Sail canopy model simulation, interpretation

Significatives variations of the red edge position on canopies have been already measured (Ferns et al., 1984; Curran et Milton, 1983), they may be caused either by folliage properties (foliar index, leaves and soil optical behaviour) or by external factors (sun position). Our data being acquired at the same hour, the variations of the red edge position are linked to the canopy. Experimental results acquired on leaves coupled with simulations have allowed us to determine that the main factors influencing the shape of the red edge are chlorophyll concentration in the leaf and LAI; sun position and soil optical behaviour being less important (Guyot,1988).
LAI depends on two parameters : ground cover percentage and leaf overlap index. For a given LAI these factors depend on the canopy geometry.
The specific influence of these two variables has been studied by Horler (1983) on corn leaves with a spectrophotometer, the red edge remained identifiable and the value did not change as leaf cover was reduced from 80% to 20%. The effect of leaf stacking produced a shift of the red edge towards longer wavelenghts. This positive shift has also been detected on cucumber plants over a background of dry soil when the "Leaf Overlap Index" increased from 1 to 5 (Ferns et al.,1984).
To quantify the influence of selected factors on red edge position and to give possible explanations to the described evolutions on AVIRIS data, we have computed RE with formula (1). The reflectance values in the four selected wavelenghts (675,695,745 and 780 nm) have been computed with a canopy radiative transfer model : SAIL (W.Verhoef, 1984) which states reflectance as a function of a canopy layer morphology, leaves and soil optical properties and sun/target/observer position

For a given solar zenithal angle (0°) and a nadir sensor, two different types of leaves have been chosen : corn and sunflower with two soil responses and three different Leaf Index Distribution Function (LIDF) : planophile, spherical and erectophile. The LAI increases from 0.5 to 12, the canopy height is one meter.
Afterwards, RE has been computed with (1) and plotted with NDVI in function of LAI.

a)description of the modelled general relationship :
As represented on figure 4, we found that for a LAI smaller than 3, the NDVI increases up to a given value (depending on the canopy) and then saturates while the red edge position decreases when the LAI varies from 0 to a minimum around 1 and then increases with LAI. It has been noticed that when the NDVI saturates (LAI > 3) the red edge position in function of the LAI is nearly a linear function.
We noticed that red edge position and NDVI increase simultaneously in only one interval of LAI variation whose precise limits depends on folliage geometry in our example.

The leaves spectral behaviour described above, shows that vegetation cover variation and stacking have different effects on the red edge shift. Examining the various evolutions of the red edge with LAI, we have studied the relationship linking red edge position with vegetation cover percentage and Leaf Overlap index.

b) cover percentage/Leaf Overlap Index
For a given LAI and LIDF, it is possible to compute the vegetation cover percentage and then the LOI, the latter being equal to the ratio of LAI when sensor is "nadir looking" with percent cover (Ferns et al., 1984). Vegetation cover percentage (%C) can be derived from the SAIL model for a given LAI and canopy height (h), using the extinction coefficient (k) for direct solar flux, the solar zenith angle being equal to zero.

$$\%C = (1 - e^{-k(LAI/h)}) \times 10^2$$

Maximum cover percentage was reached first by planophile canopy (LAI ~3) then by erectophile canopy (LAI~5), the LOI decreased when the LAI varied from zero to a minimum around 1 and then increased.
Comparison with figure 4 suggested that red edge and NDVI evolution varied differently with cover percentage and LOI.
By increasing LAI [.5 to 12], red edge and NDVI values have been plotted in function of cover percentage and LOI, red edge behaviour is presented on figure 5.
The mechanisms which are involved in RE and NVI evolution are different : NDVI depends mostly of vegetation cover percentage and RE position from LOI, with little variations depending on LAI value.
The remark about figure 3 can be clarified, by making the hypothesis that the commun increase of RE and NDVI on AVIRIS data is induced by a common increase of cover percentage and LOI in the study interval.

c) sensitivity study
In the last part of this paper we quantify red edge variations induced by canopy and external factors. Some factors influence slightly red edge position:
– canopy geometry as showed on figure 4, excepted for high LAI (shift > 1 nm for LAI > 8);
– sun position : when solar zenithal angle varies within the range (0°,60°) the shift reaches 1nm for LAI=8;
– soil reflectance: when it increases of about 8% the red edge variation is all the more important as LAI tends to zero (2nm for LAI=1).
Besides the LOI, the preponderant factor seems to be optical leaf behaviour : a corn canopy whose leaf has higher reflectance and lower transmittance than a sunflower canopy will exhibit a significatively higher RE value for a given LAI (shift > 5 nm for LAI > 8).
Guyot and Baret (1988) have shown that red edge position determination is slighly influenced by atmosphere optical depth, however this result didn't integrate atmospheric absorption. Besides the selection of the four AVIRIS bands in correct atmospheric windows, we have evaluated the dependance of red edge measurement from gaseous transmission phenomena. For two corn canopies with different LAI (2 and 12) and a spheric LIDF, we have simulated the red edge shift associated with the LAI increase. This has been done at the top and at the bottom of the atmosphere for a middle latitude winter (MLW) and a middle latitude summer (MLS) atmosphere model (Tanre et al. 1986).
For an AVIRIS bandwidth, the red edge position shifted positively of 1.5 nm (MLW) and 2.5 nm (MLS), the shift induce by LAI increase being of the same magnitude at the top and at the bottom of the atmosphere.

CONCLUSION

It is possible to calculate red slope value and red edge position on AVIRIS data with a correct band selection. In AVIRIS case, we nearly use the possible range of variation of the red edge position (40 nm in the whole area and 23 nm for all the selected sites).
The order of magnitude of these spectral features variations are about the same on AVIRIS data, simulations and referenced litterature.
Red edge position computed on simulated canopy under the assumption of the SAIL model, depends mostly of the leaf overlap index and leaves optical properties. It appears to be complementary of the normalised difference vegetation index which is mostly dependant of vegetation cover percentage.
However, a sensitivity study including the influence of NeΔρ and detailed atmospheric perturbations is required; the simulations show that the red edge position is significatively modified because of the gaseous absorption.

site (number of pixels)		1 203	2 164	3 59	4 24	5 146	6 309	7 292	8 118
NDVI	mean	.74	.65	.51	.31	.32	.72	.37	.55
	σ	.04	.03	.05	.06	.05	.03	.05	.01
Red edge position (nm)	mean	721.4	719 3	716.4	712.3	711.7	720.1	715.2	705.1
	σ	1.2	.97	2.2	1.4	2 8	.98	2.4	1.3
Red slope $\Delta\rho.10^2$	mean	2.19	1.31	1.03	<=0	1.11	2.13	<=0	2.38
	σ	.31	.32	.25		.21	.36		.18
Band 16	mean	.084	.079	074	.098	.071	.087	.076	.089
	σ	.005	.005	.009	.002	.004	.004	.002	.002
Band 29	mean	.056	.061	.063	.112	.059	.061	.077	.058
	σ	.026	.027	.023	.011	.012	.030	.016	.005
Band 40	mean	.386	.292	.199	.213	.115	.384	.168	.200
	σ	.007	004	.006	005	.005	.006	.005	.002

Table 1 : Mean values/Standard deviations for selected sites (AVIRIS data)

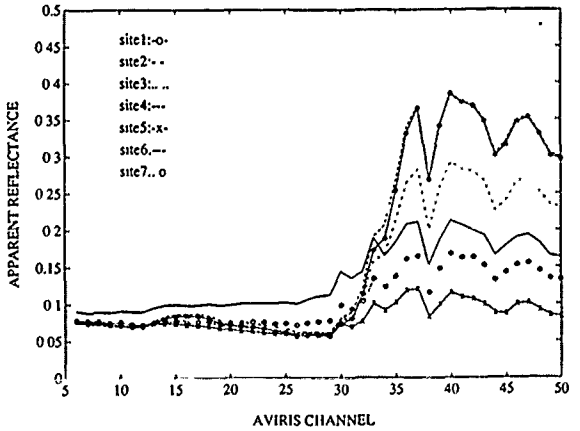


Figure 1 : Apparent reflectance spectra of selected vegetated sites (AVIRIS data)

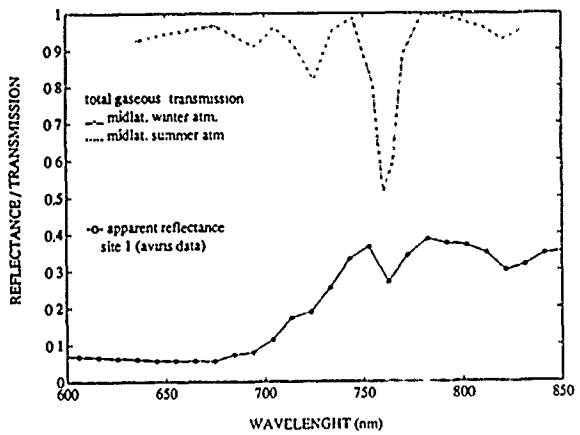


Figure 2 : Gaseous transmission influence on simulated Aviris bands, comparison with a measured vegetation response

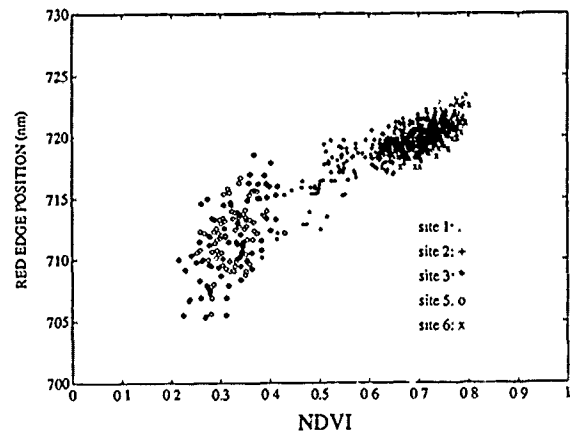


Figure 3 : Plot of the red edge position as a function of NDVI for five vegetated sites (AVIRIS data)

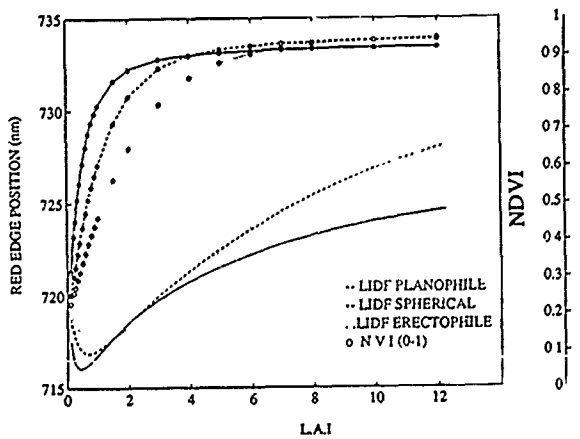


Figure 4 : Simulation of the red edge position and NDVI values as a function of LAI [0.1, 12] for a corn canopy with different LIDF

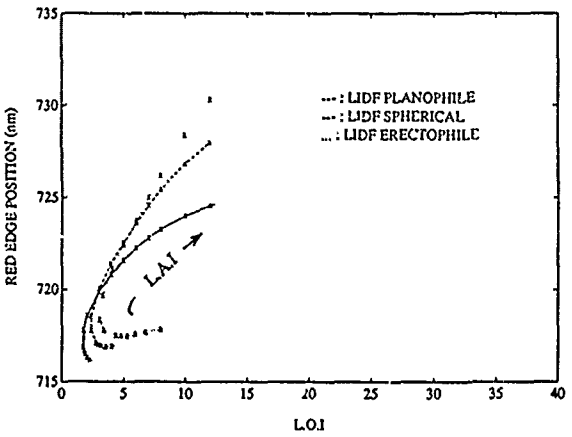


Figure 5 : Simulation of the red edge position as a function of the "Leaf Overlap Index" for a corn canopy with increasing LAI [0.5, 12]

Bibliography

1. Baret F.,Champion I., Guyot G. & Podaire A., : Monitoring wheat canopy with a high spectral resolution radiometer. Remote Sens. Env.22 : 367-378., 1987.
2. Curran P.J. & Milton E.J., : The relationship beetween the chlorophyll concentration, LAI and reflectance of a simple vegetation canopy. I.J.R.S. 4(2) 247-255, 1983.
3. Colli V; : Remote sensing of crop type and maturity. Photog. Eng. Re : Sensing 44 : 43-55., 1978.
4.Docray ., : Verification of a new method for determining chlorophyll concentration in plant by remote sensing. M. Sc. Thesis, Imperial College, University of London, 1981.
5. Ferns D.C., Zara S.J. & Barber J., : Application of high spectral resolution spectro-radiometry to vegetation. Photog. Eng. Remote Sensing 50(12) : 1725-1735, 1984.
6. Gates D.M., Keegan H.J., Schleter J.C., Weidner V.R., : Spectral properties of plants. Applied Optics 4(1) : 11-20, 1965.
7. Guyot G., : Caractérisation spectrale des couverts végétaux dans le visible et le proche infrarouge- Application à la télédétection, Bul. S.F.P.T. 95:5-22, 1984.
8. Guyot G. & Baret F., : Utilisation de la télédetection à haute résolution spatiale et spectrale. 5eme Symposium A.Q.T., 1988.
9. Horler D.N.H., Docray M., Barber J, : The red edge of plant reflectance. Int. J. Remote Sens. 4(2) : 27-228, 1983.
10. Rock B.N., Hoshizaki T. & Miller J.R. : Comparison of in situ and airborne spectral measurements of the blue shift associated with forest decline, Remote Sens Env. (24) : 109-127, 1988.
11. Tanre D., Dahaut P., Herman M., Morcrette J.J., Perbos J., Deschamps P.Y., : Effets atmospheriques en télédétection. Logiciel de simulation du signal satellitaire dans le spectre solaire. In Proc. 3rd Int. Coll. on Spectral Signatures of Objects in Remote Sensing. Les Arcs (France). ESA SP-247 : 315-319, 1986.
12. Vane G., Chrien T.G., Miller E.A. & Reimer J.H., : Spectral and radiometric calibration of the Airborne Visible/Infrared Imaging Spectrometer. AVIRIS JPL . Gregg Vane editor. 73-88, 1987.
13. Verhoef W., : Light scattering by leaf layers with application to canopy reflectance modeling : the SAIL model. Remote Sensing Env. 16 : 125-141, 1984.

Special thanks must be given to G. Vane and R.O. Green of the Jet Propulsion Laboratory who provided AVIRIS data and technical informations. Also thanks to V.C. Vanderbilt of the Ames Research Center for detailed test site descriptions.

Soumis à l'International Geoscience and Remote Sensing Symposium (IGARSS'89) et au 12e Symposium Canadien sur la Télédétection , Vancouver, C.B., 10-14 juillet , 1989

SYSTÈME DE MESURES DE RÉFLECTANCES POUR LES SPECTROMÈTRES IMAGEURS

A. Royer (1), *N.T. O'Neill* (1), *D. Williams* (2), *P. Cliche* (1) et *R. Verreault* (3).

(1) Centre d'applications et de recherches en télédétection, Université de Sherbrooke, Sherbrooke, Québec
(2) Centre de recherches sur les communications de Sherbrooke (CRCS), Université de Sherbrooke
(3) Laboratoire de physique atmosphérique, Université du Québec à Chicoutimi, Chicoutimi, Québec

RÉSUMÉ

Cet article présente quelques résultats des premiers essais en vol d'un prototype du spectromètre imageur CASI (Compact Airborne Spectrographic Imager, ITRES Research Limited, Calgary) durant l'été 1988. Les luminances CASI mesurées sont normalisées par les valeurs d'éclairement total incident mesurées simultanément avec le spectroradiomètre de terrain SPECTRON. On montre qu'il est ainsi possible d'obtenir des spectres de réflectance, propriété intrinsèque des surfaces indépendantes des conditions de mesure. La calibration des appareils est aussi discutée.

ABSTRACT

In this paper we present preliminary results for data acquired by the airborne CASI (Compact Airborne Spectrographic Imager, ITRES Research Limited, Calgary) during the summer of 1988. The radiance levels acquired by the CASI were normalized to the total downwelling incident irradiance acquired by a (SPECTRON) spectro-radiometer device mounted in the roof of the aircraft. This configuration permits the normalization of radiance data to spectral reflectance data hence to a quantity which is in principle independent of the condition of measurement. The calibration of the reflectance system is also discussed.

1. INTRODUCTION

Les spectromètres imageurs constituent la dernière génération des capteurs de télédétection dans le domaine du visible et proche infrarouge (O'Neill et *al.*, 1988; *Remote Sensing of Environment: Special Issue on Imaging Spectrometry, vol. 24, n° 1, February 1988*). Il existe actuellement au Canada deux spectromètres imageurs opérationnels aéroportés: le FLI (Fluorescence Line Imager; Borstad et *al.*, 1985) et le CASI (Compact Airborne Spectrographic Imager; *ITRES Research Limited, Calgary, Alberta, Canada*). Durant l'été 1988, nous avons réalisé les premiers essais en vol d'un prototype du CASI et nous en présentons ici une analyse préliminaire. Nous avons, en particulier, mis l'accent sur la calibration relative des données transformées en réflectance par normalisation des luminances par rapport à l'éclairement incident.

2. ACQUISITION DES DONNÉES

Le CASI est un spectromètre imageur constitué d'une matrice CCD (Charge Coupled Device) de 576 x 288 pixels illuminée par un réseau de diffraction. La souplesse du système permet l'acquisition de données dans une gamme de configuration variée, soit essentiellement spectrale: 288 bandes avec une faible résolution spatiale, jusqu'à une configuration spatiale de 512 pixels, avec une résolution spectrale dégradée. Le champs de visée total est de 30°, la

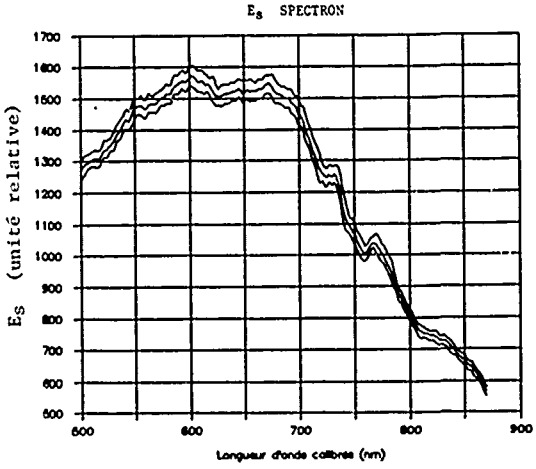


Figure 1: Éclairement solaire total incident mesuré à l'aide du spectroradiomètre SPECTRON installé sur le toit de l'avion et équipé d'un diffuseur intégrant (minimum, maximum et moyenne de spectres obtenus pendant le survol).

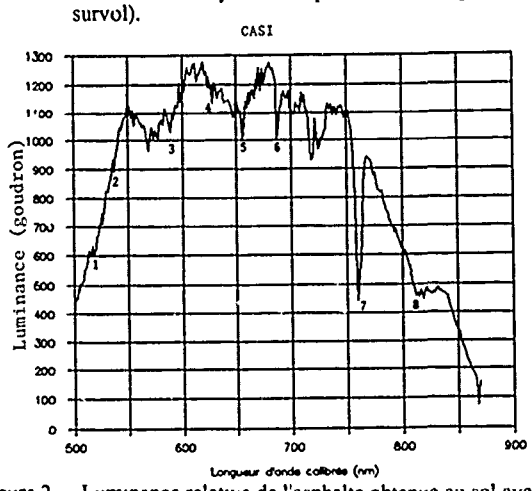


Figure 2. Luminance relative de l'asphalte obtenue au sol avec le CASI rapportée aux longueurs d'onde calibrées par rapport aux raies d'absorption de l'oxygène et de la vapeur d'eau:

1: 527 nm, 2: 544 nm, 3: 589 nm, 4: 628 nm,
5: 656 nm, 6: 687 nm, 7: 760 nm, 8: 812,5 nm.

couverture spectrale est comprise entre 450 et 900 nm, et la résolution radiométrique est de 12 bits. Dans sa version prototype (été 1988), l'instrument a permis l'acquisition d'images spectrales à 288 canaux répartis sur 9 pixels spatiaux. L'instrument et le système d'enregistrement des données ont été installés sur un monomoteur Piper PA-28 (de la compagnie Air Focus, Chicoutimi, Québec, Canada). L'image analysée a été acquise le 19 août 1988 au-dessus du lac d'Argent dans la région de Sherbrooke (altitude de vol: 3 000' par rapport au sol, vitesse de l'avion 112 noeuds).

Simultanément aux mesures CASI, des mesures de l'éclairement solaire total incident ont été réalisées à l'aide d'un spectroradiomètre (SPECTRON SE 590, *SPECTRON ENGINEERING INC., Denver, Colorado, USA*) installé sur le toit de l'avion et équipé d'un verre diffuseur (2 π sr). Le SPECTRON permet d'analyser le spectre incident entre 368 et 1 113 nm sur 252 canaux, soit avec une résolution 3 fois moins fine que celle du CASI. La Figure 1 montre le signal moyen obtenu avec le SPECTRON durant les 5 mn de vol (H GMT 14:24 - 14:29) ainsi que la variabilité observée (ensemble des valeurs maximum et minimum) ($\pm$ ~ 2-3%). Le signal reste relativement stable.

3. CALIBRATION RELATIVE DES SPECTROMÈTRES

Pour la calibration relative des spectres en longueur d'onde et afin de pouvoir les normaliser, nous avons effectué une série de mesures simultanées CASI / SPECTRON au sol au-dessus d'une surface homogène (goudron-asphalte). La Figure 2 montre le spectre CASI de goudron corrigé à partir de 8 longueurs d'onde de références correspondant aux raies d'absorption de l'oxygène et de la vapeur d'eau. Le signal, dans ce cas, varie de 501 nm à 868,8 nm, soit avec une résolution de 1,3 nm. Les raies d'absorption sont très bien identifiées et leur position est très stable (écart type de 0,58 nm pour le pic à 760 nm par exemple, pour un ensemble de 50 valeurs).

Les données SPECTRON, moins faciles à calibrer du fait de la faible résolution spectrale, sont comparées au CASI dans la Figure 3. Il apparaît que les variations du spectre CASI lissées avec une fenêtre de $\pm$ 11 pixels CASI correspondent approximativement bien à celles observées avec le SPECTRON. Ceci suggère un pouvoir de résolution effectif du SPECTRON de l'ordre de 14 nm, soit très inférieur à la valeur nominale de 3 nm. La calibration en longueur d'onde a fait apparaître un décalage positif de l'ordre de + 10 nm (λ corrigé - λ nominal) au début du spectre (500 nm) et un décalage négatif du même ordre de grandeur (- 10 nm) vers les plus grandes longueurs d'onde. Afin de simuler des spectres de réflectances que l'on obtiendrait avec le CASI, nous avons ramené le SPECTRON à la résolution CASI. Pour celà, une fonction d'interpolation du type (sin x/x) a été appliquée au spectre d'éclairement solaire du SPECTRON. La mise en concordance des spectres a ainsi permis d'établir un coefficient d'étalonnage relatif (K_λ) entre le CASI et le SPECTRON (Figure 4). Celui-ci a été défini par rapport à la réflectance d'un panneau blanc de référence (Kodak) suivant la relation:

$$K_\lambda = \left[L_s(\text{asphalte}) / L_s(\text{panneau ref.}) \right] / \left[L_c(\text{asphalte}) / E_s \right]$$

$$= \left[\rho_s(\text{asphalte}) / \rho_c(\text{asphalte}) \right]$$

où les indices s et c réfèrent au SPECTRON et au CASI respectivement, et L correspond à une luminance, E à un éclairement et ρ à une réflectance.

Les pics de haute fréquence qui apparaissent dans les variations observées (Figure 4) sont dus à la différence de résolution des deux instruments utilisés. Les fortes valeurs au-delà de 860 nm résultent du bruit instrumental.

Ce coefficient d'étalonnage nous permet donc de transformer les luminances relatives CASI en réflectances normalisées, propriété intrinsèque de la surface et indépendante des conditions d'éclairement (si la surface est Lambertienne).

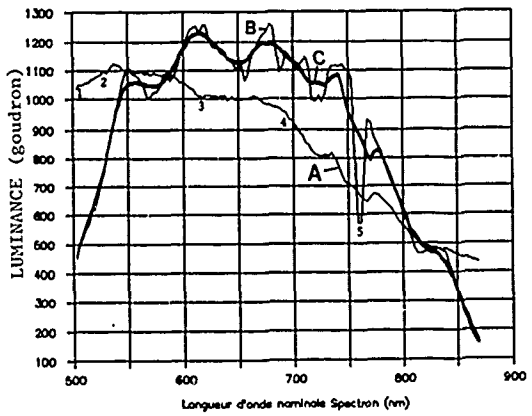


Figure 3. Comparaison des luminances relatives de l'asphalte obtenues avec le SPECTRON et le CASI:
A) Données Spectron brutes
B) Spectre CASI lissé avec une fenêtre de $\pm$ 2 pixels CASI
C) Spectre CASI lissé avec une fenêtre de $\pm$ 11 pixels CASI

Les données SPECTRON ont été par la suite calibrées en longueurs d'onde à partir des points:
1: 517 nm, 2: 544 nm, 3: 628 nm, 4: 687 nm et 5: 760 nm.

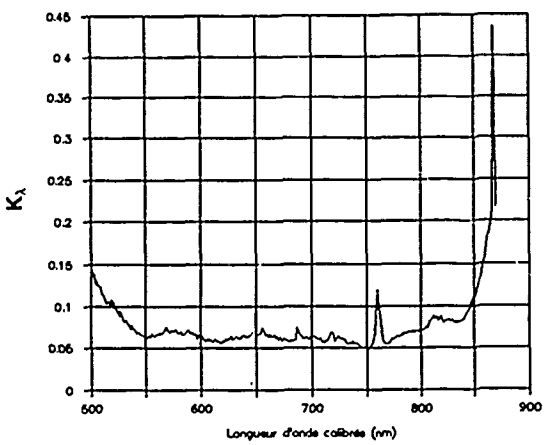


Figure 4: Coefficient d'étalonnage CASI/SPECTRON (K_λ) pour le calcul des réflectances obtenues à partir de mesures simultanées de l'asphalte, d'un panneau blanc de référence et de l'éclairement incident.

4. RÉSULTATS: EXEMPLE DE SPECTRES DE REFLECTANCE

La Figure 5 présente une petite portion de l'image analysée correspondant à 9 pixels x 52 lignes de laquelle on a extrait 3 spectres représentatifs de la végétation, de l'eau et d'une surface de sol nu. Un problème dans le réglage du temps d'intégration des signaux enregistrés a provoqué une saturation des données dans les fortes valeurs de luminances. D'autre part, une comparaison des signaux en colonnes a fait apparaître des différences significatives dues au vignetage des lentilles d'entrée. Tous ces problèmes ont été réglés sur la version actuelle du CASI.

Les réflectances équivalentes à ces luminances ont donc été déduites à partir du coefficient d'étalonnage $K\lambda$:

$$\rho_\lambda = K_\lambda\left(L_c(\text{cible}) / E_s\right)$$

Les pics d'absorption ont été largement éliminés et les spectres obtenus présentent des variations d'amplitudes typiques des surfaces analysées. Pour l'eau, il subsiste cependant des variations de hautes fréquences dans les faibles valeurs de réflectance (0,01 - 0,005). Celles-ci peuvent résulter de la non linéarité de la réponse du signal associé entre autre au problème du réglage du temps d'intégration.

5. CONCLUSION

Nous avons analysé les spectres de luminances relatives obtenus avec le spectromètre imageur CASI. La fine résolution spectrale (1,3 nm) permet un recalage très précis de la calibration en longueur d'onde.

La calibration radiométrique peut être obtenue par normalisation des luminances par rapport à l'éclairement total incident. Les tests effectués avec un deuxième spectroradiomètre (SPECTRON) enregistrant l'éclairement simultanément sont très concluants. Un système aéroporté intégrant les deux mesures de façon synchronisée (système par fibre optique par exemple) permettrait une analyse quantitative des spectres de réflectance et augmenterait ainsi considérablement le potentiel d'analyse des données obtenues.

Nous tenons à remercier monsieur Luc St-Pierre pour sa précieuse collaboration dans la réalisation du projet, ainsi que la compagnie Air Focus de Chicoutimi. Ce projet a été financé en partie par le CRSNG (projet A 8643 et A 1765), par l'Institut scientifique du Saguenay (Chicoutimi) et par le ministère des Pêches et Océans Canada.

RÉFÉRENCES

Borstad, G.A., Edel, H.R., Gower, J.F.R. et A.B. Hollinger, "Analysis of test and flight data from the Fluorescence Line Imager", *Canadian Special Publication of Fisheries and Aquatics Sciences 83, Pêches et Océans Canada*, 38 p. 1985.

O'Neill, N.T., Royer, A. et M.C. Mouchot, "Les spectromètres imageurs: leur potentiel en télédétection", *6ème Congrès de l'AQT, Sherbrooke, Québec, Canada*, 4-6 mai 1988, 13 p.

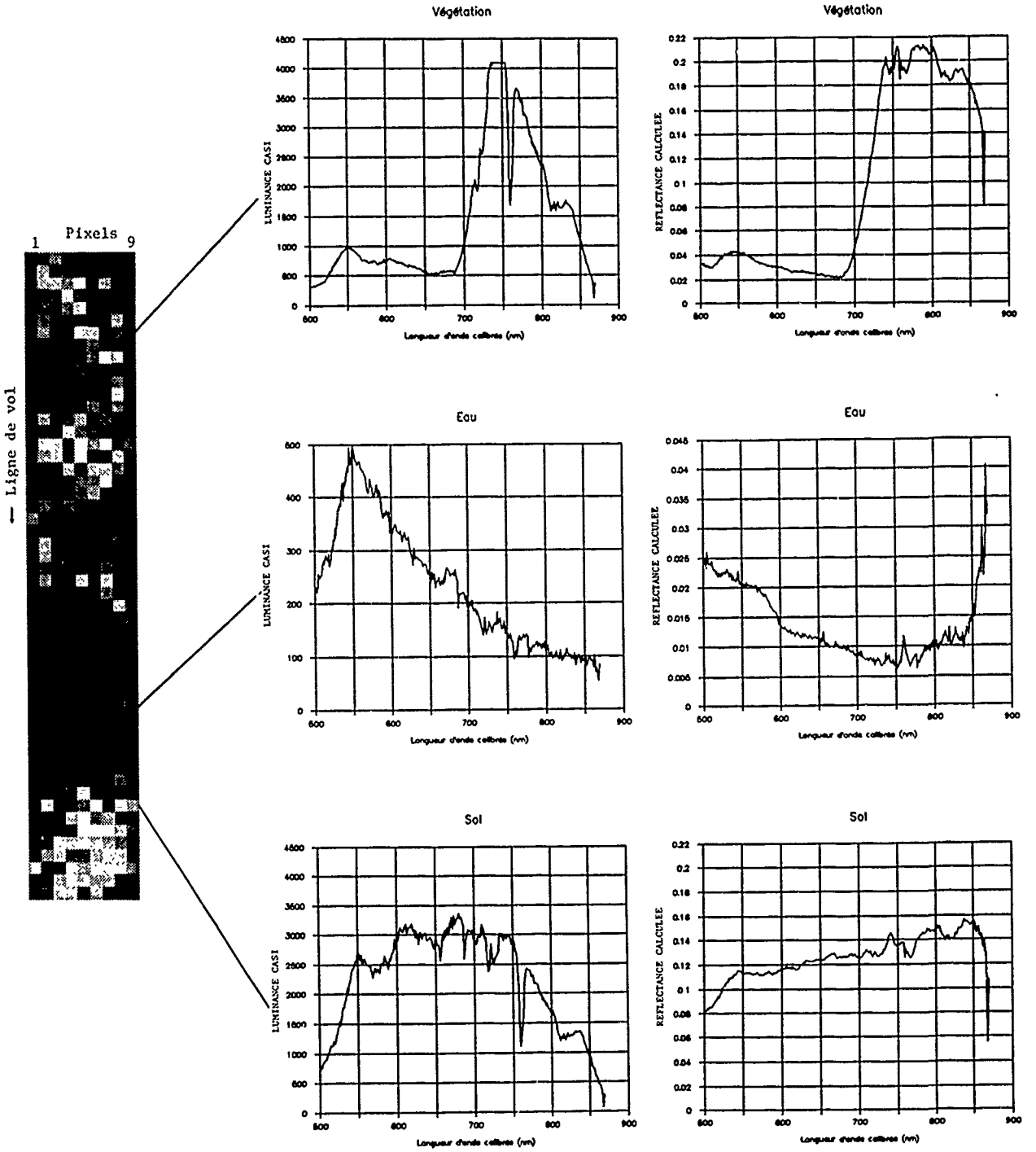


Figure 5: Image spectrale du lac d'Argent, Québec du 19 août 1988. Luminances et réflectances spectrales des pixels 9, ligne 7 pour la végétation, ligne 37 pour l'eau et ligne 45 pour le sol nu. Les réflectances ont été obtenues par normalisation par rapport à l'éclairement solaire incident mesuré simultanément sur le toit de l'avion.

ANALYSIS OF SCATTERING BEHAVIOR FROM A KNOWN RANDOM SURFACE USING FREQUENCY DIVERSE PROJECTION IMAGING

C. E. Nance & A. J. Blanchard

Wave Scattering Research Center
University of Texas at Arlington
Arlington, Texas, U.S.A.

ABSTRACT

The scattering from randomly rough surface targets has been studied in the past by using theoretical modeling, numerical simulation, and experimental measurement. The experimental scattering behavior for a known randomly rough surface has also been studied (C. E. Nance, A.J. Blanchard, and Chen, 1988). The surface target was a metalized randomly rough surface constructed from known surface statistical information. The two dimensional projected image of this artificial rough surface will be examined. The images will be compared with the known "ground truth" or surface coordinate geometry to try to interpret the information which can be obtained by the projection imaging of surface structured targets.

Keywords: Imaging, Rough Surface, ISAR, Radar, Microwave

Introduction

The methods used in the formation of the images have been presented in previous work (Blanchard & Dolaty, 1988)(Farhat,Werner,& Chu, 1984). The images are two dimensional projected images of the three dimensional target. The two dimensional projection of the object function is found by integrating the z dependance out of the relation.

$$\gamma_{proj}(x,y) = \int_{-\infty}^{+\infty} \iiint \Gamma(k_x,k_y,k_z)\, e^{j(k_x x+k_y y+k_z z)}\, dk_x dk_y dk_z dz$$

This reduces to the two dimensional inverse Fourier transform relation, which is used to produce the two dimensional projected object function onto the x-y plane. The scatterers which do not lie in the x-y plane are defocused.

$$\gamma_{proj}(x,y) = \iint \Gamma(k_x,k_y)\, e^{j(k_x x+k_y y)}\, dk_x dk_y$$

Experiment Description

The images of a Gaussian rough surface were taken in an indoor anechoic range. The scattering data was processed off line on a workstation. The rough surface patch was a circular target, (1.2m diameter), with known surface height deviation and correlation lengths. The surface was machined from a four foot square structural foam sheets. The surface was then coated with aluminum flake paint to provide a conducting surface. The standard height deviation of the surface is 3 cm, and the correlation length is 7 cm. This places the scattering behavior in the Kirchoff region. The roughness scale of the surface is therefore considered large. Figure-1 plots the surface grid of the rough surface target with the units of the figure in centimeters. The rough surface target is the 1.2 meter circle cut from the 1.2 meter square section.

The measurements were made in the Bistatic Cross Section Range at the University of Texas at Arlington. The measurement hardware is a FM swept system consisting of a Hewlett/Packard 8510 Network Analyzer, 8341A Sweeper, and a 8511A Test Set.

The measurements for the monostatic image were taken at an incidence angle of 45 degrees. The azimuth position was varied from 0 to 90 degrees in 0.5 degrees increments. The data was acquired for 401 frequency steps between 4 and 18 GHz.

Results

Figure-2 is the monostatic two dimensional projected image of the rough surface. The outline of the circular edge of the target can be seen, but the diameter of the image is smaller than the actual size of the target due the projection of the object function. This projection is a linear compression, therefore it is viewed as a scaled image. The scaling factor is around 0.7 which is close to the cosine of the look angle of the measurements. The image does not represent the actual source distribution on the surface because the method of which the data was acquired. To obtain an actual source distribution for a certain illumination direction, the transmitter position relative to the target would have to remain fixed, and the receiver would have to move to sample the scattered fields. This image represents an average source distribution because the transmitter and receiver positions changed with respect to the target. Figure-3 is the two dimensional monostatic image containing only the data of the first 10 degrees in the azimuth rotation.

This paper has illustrated the type of images that would be obtained from random large scale surfaces when using two dimensional projection imaging. The method of target rotation, with the transmitter and receiver fixed to acquire the spatial frequency data needed for the image construction produces an image of the average source distribution of the target over the look angles. A method of holding the target fixed and scanning with the transmitter and receiver in azimuth would be equivalent. This type of data acquisition is similar to the manner in which data is acquired in conventional SAR systems.

References

[1] C.E. Nance, A.J. Blanchard, and M. Chen, "Polarimetric Microwave Scattering from Known Randomly Rough Surfaces", *Proceedings of IGARSS 88 Symposium*, Edinburgh, Scotland, 1988.

[2] A.J. Blanchard, M. Dolaty, "Bistatic Frequency Diverse Imaging of Complex Scattering Targets", *Proceedings of IGARSS 88 Symposium*, Edinburgh, Scotland, 1988.

[3] Farhat,Werner, and Chu, "Prospects for Three Dimensional Projective and Tomographic Imaging Radar Networks", *Radio Science*, Vol. 19,pp. 1347-1355.

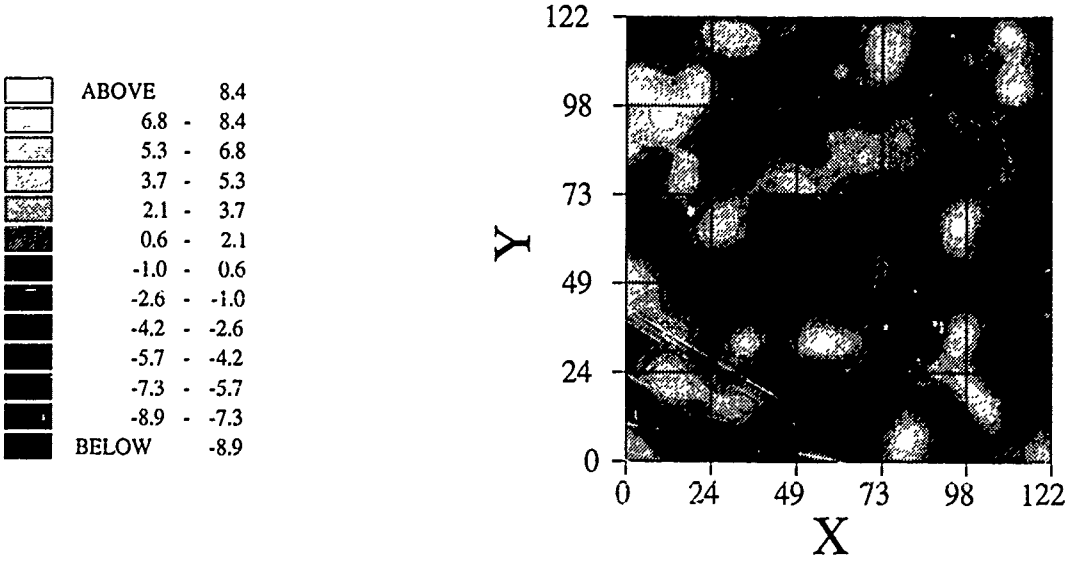


Figure-1: Surface Grid Plot For Gaussian Random Surface
(units for grid and heights in centimeters)

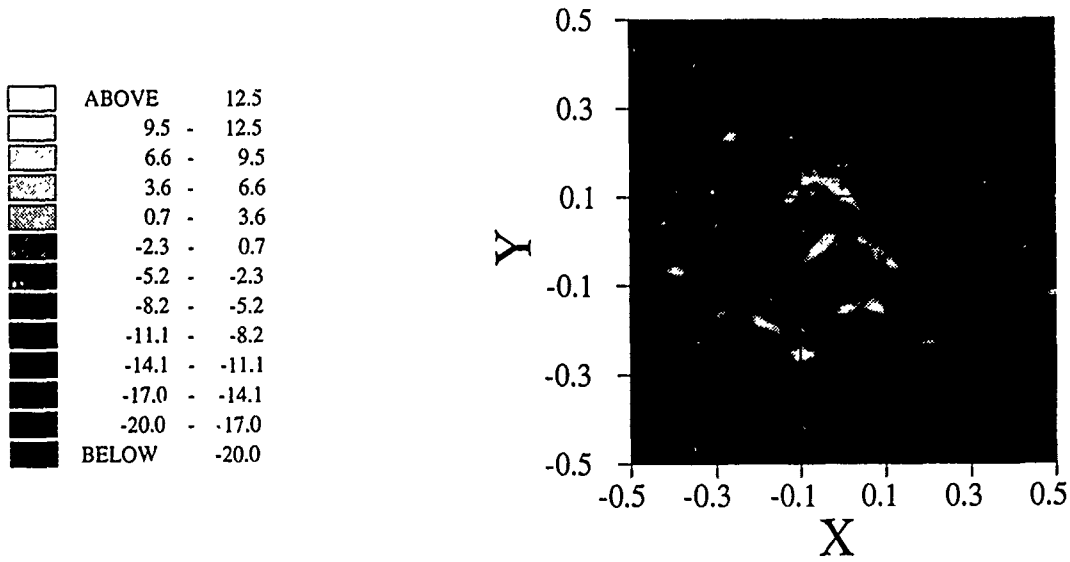


Figure-2: Two Dimensional Image of Gaussian Random Surface
(4-18 GHz, monostatic, 45 degree look angle, 90 azimuth sample)

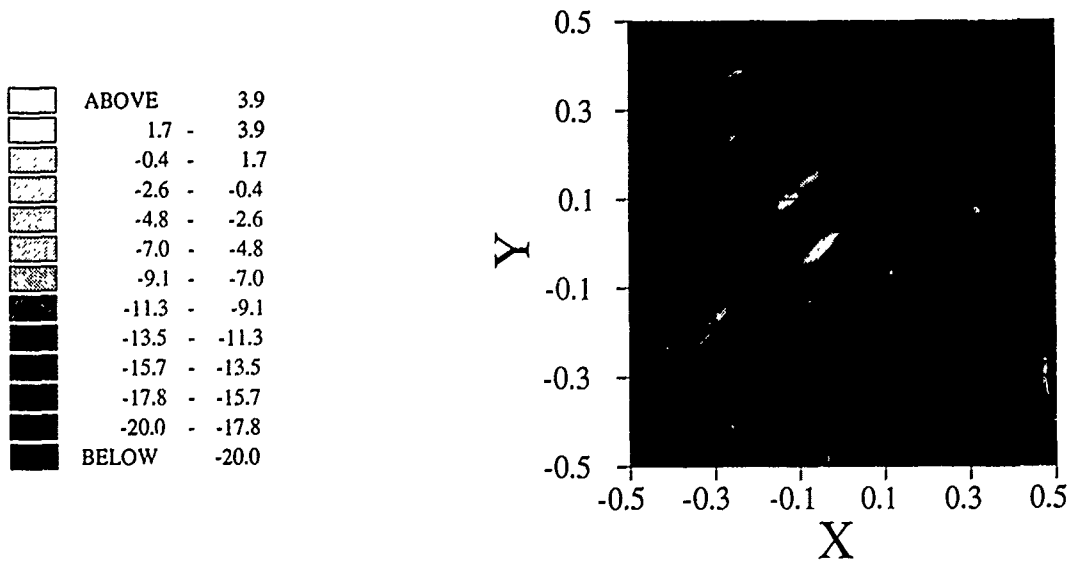


Figure-3: Two Dimensional Image of Gaussian Random Surface
(4-18 GHz, monostatic, 45 degree look angle, 10 azimuth sample)

MULTIPLE VOLUME SCATTERING EFFECTS IN MICROWAVE POLARIMETRIC REMOTE SENSING

Leung Tsang, Kung-Hau Ding, Boheng Wen, and Dale P. Winebrenner
Department of Electrical Engineering, FT-10
University of Washington
Seattle, WA 98195

ABSTRACT

The multiple volume scattering effects in microwave polarimetric remote sensing are studied. For terrain media consisting of sparse concentrations of nonspherical scatterers, the polarimetric signatures are calculated by solutions of the vector radiative transfer equations. For nonspherical particles, the extinction matrix is used giving two effective propagation constants with two characteristic polarizations. First-order and second-order solutions are calculated and compared. For terrain media consisting of a dense distribution of particles, the dense medium radiative transfer equations are used. The dense medium radiative transfer theory takes into account the effects of correlated scattering. The theoretical results of dense media transfer equations agree well with controlled laboratory experiments and give the correct functional dependence of scattering versus concentrations of particles. Numerical results are illustrated for copolarization signatures for the backscattering return. It is found that the effects of multiple scattering can significantly decrease the degree of polarization of the return signal.

Introduction

There has been increasing interest in the applications of polarimetry to microwave remote sensing of geophysical terrain (*Evans et al., 1988; Wen et al., 1988*). Some of the prominent features in terrain signatures are a pedestal in copolarization returns, exhibition of partial polarization on averaging of the return signals, and a phase difference between *vv* and *hh* polarized waves. In this paper, we model the terrain as discrete scatterers that are randomly distributed and are governed by orientation and size distributions. The polarimetric signatures are calculated by using the vector radiative transfer theory. Two cases are considered: Sparse nonspherical particles and dense media with multiple species of particles.

For the case of sparsely distributed nonspherical particles, we use the vector radiative transfer theory with a general nondiagonal extinction matrix and a phase matrix that are averaged over particle sizes and orientations (*Tsang et al., 1985*). The Mueller matrix as calculated in this manner is an ensemble-averaged quantity due to the inherent nature of the radiative transfer theory. Numerical results of copolarization and phase difference are illustrated using the first-order and second-order solutions of the Mueller matrix. The inclusion of the nondiagonal extinction matrix yields an appreciable phase difference between *vv* and *hh* polarizations particularly for aligned scatterers. The phase difference decreases with increasing variance of orientation of the scatterers. There generally exists a pedestal in the copolarization return.

For the case of dense media with multiple species of particles, we apply the dense medium radiative transfer theory. The dense medium radiative transfer equations have been derived to take into account correlated scattering effects (*Tsang and Ishimaru, 1987; Tsang, 1988*). The pair distribution functions are calculated by using the Percus-Yevick approximation (*Ding and Tsang, 1988*). We have used the dense medium transfer theory to explain the phenomenon observed in laboratory experiments (*Kuga and Ishimaru, 1984*) which indicate that the bistatic intensity first increases with the volume fraction of particles until a maximum is reached, and then decreases when the volume fraction further increases. Numerical results of radar polarimetry are illustrated. It is shown that multiple scattering also gives rise to a pedestal in the copolarization signature.

Vector Radiative Transfer Theory for a Layer of Sparsely Distributed Nonspherical Particles

Consider a collection of sparsely distributed nonspherical particles with permittivity ϵ_s embedded in a background medium with permittivity ϵ (region 1) overlying a homogeneous dielectric of permittivity ϵ_2 (region 2). An incident wave is launched from region 0 in direction $(\pi-\theta_o, \phi_o)$. Region 0 has permittivity ϵ which is the same as that of the background medium of region 1 (Figure 1).

The vector radiative transfer equation in region 1 is, for $0 \le \theta \le \pi$, $0 \le \phi \le 2\pi$ (*Tsang et al., 1985*)

$$\cos\theta \frac{d}{dz}\bar{I}(\theta,\phi,z) = -\bar{\bar{\kappa_e}}(\theta,\phi)\cdot\bar{I}(\theta,\phi,z) + \int_0^{2\pi} d\phi' \int_0^{\pi} d\theta' \sin\theta' \bar{\bar{P}}(\theta,\phi;\theta',\phi')\cdot\bar{I}(\theta',\phi',z) \quad (1)$$

where $\bar{I}(\theta,\phi,z)$ is the Stokes vector in direction (θ,ϕ) and $\bar{\bar{P}}(\theta,\phi;\theta',\phi')$ is the phase matrix. The extinction matrix $\bar{\bar{\kappa_e}}(\theta,\phi)$ is generally nondiagonal and direction dependent. It not only describes the extinction rates, but also contains the propagation constant difference between the vertical and horizontal polarized waves. The elements of the extinction and phase matrices are expressed entirely in terms of the scattering function amplitudes that are averaged over size and orientation distributions of particles. It is important to note that the scattering amplitude has to be evaluated to sufficient accuracy. The rule of thumb is to calculate them up to the leading terms of the real part and imaginary part.

The scattered Stokes vector, $\bar{I}_s(\theta_s,\phi_s) = \bar{I}(\theta_s,\phi_s,z=0)$, is proportional to the incident Stokes vector $\bar{I}_o$ with the coupling described by the Mueller matrix $\bar{\bar{M}}(\theta_s,\phi_s;\pi-\theta_o,\phi_o)$ (*Evans et al., 1988*) as

$$\bar{I}_s(\theta_s,\phi_s) = \bar{\bar{M}}(\theta_s,\phi_s;\pi-\theta_o,\phi_o)\cdot\bar{I}_o \quad (2)$$

We note that the Mueller matrix as calculated here is an ensemble-averaged quantity due to the ensemble-averaged nature of transport theory. The phase difference between vv and hh waves ϕ_{vh} is defined by

$$\phi_{vh} = \tan^{-1}\left(\frac{M_{43} - M_{34}}{M_{33} + M_{44}}\right) \tag{3}$$

where M_{ij} is the ij element of $\overline{\overline{M}}$.

Vector Radiative Transfer Theory for Dense Media with Multiple Species of Particles

Consider the same geometric configuration as Figure 1 with region 1 now occupied by densely distributed multiple species of spherical particles (Figure 2). The propagation and scattering of waves in region 1 is governed by the dense medium radiative transfer equations. The important parameters for this theory are effective propagation constant, extinction rate and albedo. They can be expressed in terms of the physical parameters of the medium: sizes, concentrations, and dielectric properties of particles. This allows us to study the electromagnetic responses of media directly through their characteristics.

Let there be L species of particles in the medium and let s_j be the species index. The volume fraction $f_{s_j} = 4\pi n_{s_j} a_{s_j}^3/3$ of the s_j species can be appreciable and the wavenumber k_{s_j} of species s_j can be substantially different from the background wavenumber k. The quantities a_{s_j} and n_{s_j} are, respectively, the radius and the number density of particles of species s_j. Generally, the effective propagation constant K is complex and $K' \equiv Re(K) \gg Im(K) \equiv K''$. The imaginary part K'' accounts for the attenuation of a coherent wave due to both absorption and scattering. For small spherical particles, $ka_{s_j} \ll 1$, the effective propagation constant under the quasicrystalline approximation with coherent potential is given by (*Ding and Tsang, 1988*)

$$\begin{aligned} K^2 &\approx K'^2 + i2K'K'' \\ &= k^2 + \frac{3K_o^2}{D(K_o)} \sum_{s_l=1}^{L} f_{s_l} y_{s_l}(K_o) \left\{ 1 + i \frac{2K_o^3}{3D(K_o)} \right. \\ &\quad \left. \cdot \left[a_{s_l}^3 y_{s_l}(K_o) + \sum_{s_j=1}^{L} y_{s_j}(K_o) a_{s_j}^3 n_{s_j} 8\pi^3 H_{s_j s_l}(p=0) \right] \right\} \end{aligned} \tag{4}$$

where K_o indicates the propagation constant when the scattering attenuation is ignored. In (4), $y_{s_j}(K_o) = (k_{s_j}^2 - k^2)/(3K_o^2 + k_{s_j}^2 - k^2)$ and $D(K_o) = 1 - \sum_{s_j=1}^{L} f_{s_j} y_{s_j}(K_o)$ are functions of K_o. $H_{s_j s_l}(p)$ is the Fourier transform of the correlation function $h_{s_j s_l}(r) = g_{s_j s_l}(r) - 1$ and $g_{s_j s_l}(r)$ is the pair distribution function of two particles of species s_j and s_l separated by a distance r.

The extinction rate κ_e is equal to $\kappa_e = 2K''$ and the albedo $\tilde{\omega}$ is given as (*Tsang, 1988*)

$$\begin{aligned} \tilde{\omega} &= \frac{2|K_o|^4}{\kappa_e |D(K_o)|^2} \sum_{s_l=1}^{L} f_{s_l} y_{s_l}(K_o) \\ &\quad \left\{ a_{s_l}^3 y_{s_l}^*(K_o) \left[1 + \sum_{s_j=1}^{L} n_{s_j} a_{s_j}^3 y_{s_j}^*(K_o) 8\pi^3 H_{s_j s_l}(p=0) \right] \right\} \end{aligned} \tag{5}$$

The dense medium radiative transfer equation in region 1 is, for $0 \le \theta \le \pi$, $0 \le \phi \le 2\pi$ (*Tsang, 1985*)

$$\begin{aligned} \cos\theta \frac{d}{dz} \bar{I}(\theta, \phi, z) &= -\kappa_e \bar{I}(\theta, \phi, z) + \frac{\kappa_e \tilde{\omega}}{4\pi} \\ &\quad \times \int_0^{2\pi} d\phi' \int_0^{\pi} d\theta' \sin\theta' \overline{\overline{P}}(\theta, \phi; \theta', \phi') \cdot \bar{I}(\theta', \phi', z) \end{aligned} \tag{6}$$

where the phase matrix $\overline{\overline{P}}(\theta, \phi; \theta', \phi')$ is identical to the Rayleigh phase matrix (*Tsang et al., 1985*). This is because of the small particle assumption and that the pair distribution functions are only correlated for the range of a few diameters which are less than a wavelength.

The dense medium radiative transfer equation (6) has a similar form to that of the conventional radiative transfer equation except that the expressions for κ_e and $\tilde{\omega}$ are different. Thus, the same numerical technique, discrete ordinate-eigenanalysis method (*Tsang et al., 1985*), can be used to solve for the dense medium radiative transfer equations including Stokes vectors and Mueller matrices. In solving the radiative transfer equations numerically, the total intensity is decomposed into reduced intensity and diffuse intensity (*Wen et al., 1988*). All orders of multiple scattering within radiative transfer theory are included in this numerical solution.

Comparison with Laboratory Experiments

To validate our theoretical models, the dense medium radiative transfer theory is used to explain the phenomena observed in the controlled laboratory experiments (*Kuga and Ishimaru, 1984*). The laboratory discrete random medium is comprised of a collection of latex spheres suspended in a cell with water. The experiments were performed by measuring bistatic intensities over a wide range of concentrations. An equivalent diameter of $0.078\mu m$ is used for the present theoretical computations. A loss tangent of 10^{-6} is introduced to account for a slight absorption loss in the latex spheres. For comparison, we also show the numerical results based on the conventional radiative transfer theory which assumes independent scattering. The measured and calculated bistatic scattered intensities as a function of fractional volume of particles are illustrated in Figure 3. It is found that the measured bistatic intensities first increase with the volume fraction of particles until a maximum is reached, and then decrease as the volume fraction further increases. Comparison with both conventional and dense medium radiative transfer theories indicates that the results of dense medium transfer theory contain these experimental features while the results of conventional transfer theory predict a monotonic increase with concentration.

Numerical Results of Radar Polarimetry

By making use of the eigenmatrix of the extinction matrix, the vector radiative transfer equation (1) can be cast into integral equations containing boundary conditions (*Tsang et al., 1985*). The first-order and second-order iteration of the integral equations give, respectively, the first-order and second-order solutions of the Mueller matrix. In Figures 4 and 5, the copolarization signatures are illustrated as a function of ellipticity angle χ and orientation angle ψ. The results in Figure 4 are based on the first-order theory for the case of randomly oriented spheroids. The copolarization return exhibits a pedestal which illustrates the effects of heterogeneity of scatterers. In Figure 5, we show the results of second-order theory for a half-space medium with less lossy aligned scatterers.

In Figure 6, the phase differences ϕ_{vh} are plotted as a function of Euler angle β_2 with $\beta_1 = 0$. We note that there exists a finite phase difference for aligned scatterers ($\beta_2 = 0$). This is because for aligned spheroids, there exist two effective propagation constants that are appreciably different. The phase difference is also larger for a larger angle of incidence. However, as β_2 increases representing an increase in the variance of

orientation, the phase difference decreases which implies that the effective anisotropy of the medium decreases.

In the studies of remote sensing of geophysical terrain, very often the ground truth data are given in terms of a histogram of particle size distribution. Therefore, the multiple species model is applicable to a medium with size distribution. A general expression for particle size distribution is the modified gamma distribution $n(a) = K_1 a^{\gamma} \exp(-K_2 a^{\varrho})$, where a is the size of particle and $n(a)da$ gives the number of particles per unit volume having sizes between a and $a + da$. For a medium with particle size distribution, we can discretize the continuous size distribution into a histogram of L different sizes to incorporate the results shown above. In Figure 7, we illustrate the numerical results of copolarization signatures for a slab of snow medium with ice particles having size distributions. We note that the horizontal polarization produces a larger return than the vertical polarization, while circular polarizations give minimum return which is independent of orientation angle. The copolarization also shows a pedestal that has been observed in SAR polarimetric radar data (*Evans et al., 1988*).

CONCLUSIONS

In this paper, we have demonstrated the usefulness of the vector radiative transfer theory for studying polarimetric signatures of random discrete scatterers. In the case of a sparse concentration of nonspherical particles, the use of a nondiagonal extinction matrix permits the explanation of the phase differences between *vv* and *hh* signatures that have been observed in polarimetric remote sensing. We have also studied the effects of orientation distributions and second-order scattering. For the case of dense media, we have included the effects of correlated scatterers and have corrected the deficiencies in the conventional radiative theory. The dense medium radiative transfer theory explains the features of dependence on concentration observed in laboratory controlled experiments. We have studied the effects of multiple scattering and effects of correlated scatterers governed by particle size distribution. We have shown that generally multiple scattering can cause a pedestal in the return signal.

REFERENCES

Ding, K. H., and L. Tsang, "Effective propagation constants of dense nontenuous media with multi-species of particles," *J. Electro. Waves Applic.*, Vol. 2, pp. 757-777, 1988.

Evans, D. L., T. G. Farr, J. J. van Zyl, and H. A. Zebker, "Radar polarimetry: analysis tools and applications," *IEEE Trans. Geosci. Remote Sensing*, Vol. GE-26, pp. 774-789, 1988.

Kuga, Y., and A. Ishimaru, "Retroreflectance from a dense distribution of spherical particles," *J. Opt. Soc. Am. A.*, Vol. 1, pp. 831-835, 1984.

Tsang, L., J. A. Kong, and R. T. Shin, *Theory of Microwave Remote Sensing*, New York: Wiley-Interscience, 1985.

Tsang, L., and A. Ishimaru, "Radiative wave equations for vector electromagnetic propagation in dense nontenuous media," *J. Electro. Waves Applic.*, Vol. 1, pp. 52-72, 1987.

Tsang, L., "Dense media radiative transfer theory for dense discrete random media with particles of multiple sizes and permittivities," submitted to *Progress in Electromagnetics Research*,, 1988.

Wen, B., L. Tsang, D. P. Winebrenner, and A. Ishimaru, "Dense medium radiative transfer theory: comparison with experiment and application to microwave remote sensing and polarimetry," submitted to *IEEE Trans. Geosci. Remote Sensing*,, 1988.

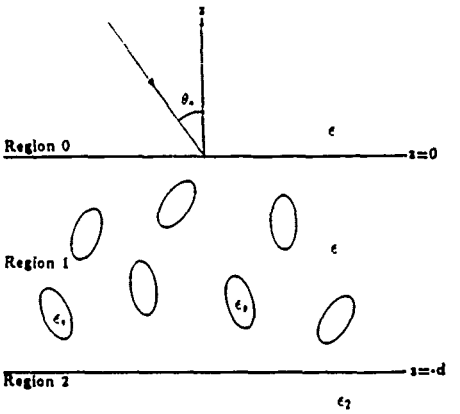


Figure 1. An incident plane wave impinging upon a layer of nonspherical particles overlying a homogeneous half space of permittivity ϵ_2.

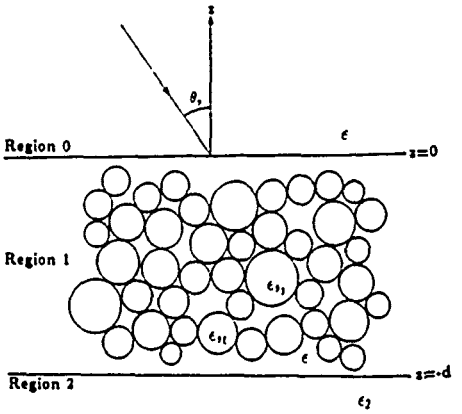


Figure 2. An incident plane wave impinging upon a layer of densely distributed spherical particles overlying a homogeneous half space of permittivity ϵ_2.

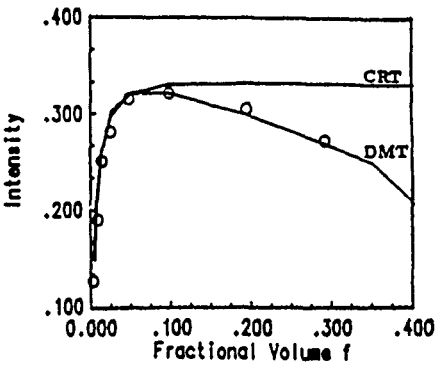


Figure 3. Normalized bistatic intensities as a function of fractional volume of Latex spheres. The dense medium radiative transfer theory (DMT) and conventional radiative transfer theory (CRT) results are compared with the experimental data (o). The wavelength of laser light is $0.6328\mu m$, the index of refraction is 1.588 plus a 10^{-6} loss tangent for Latex spheres and 1.332 for water, the equivalent diameter of Latex spheres is $0.078\mu m$, and the slab width is $10mm$.

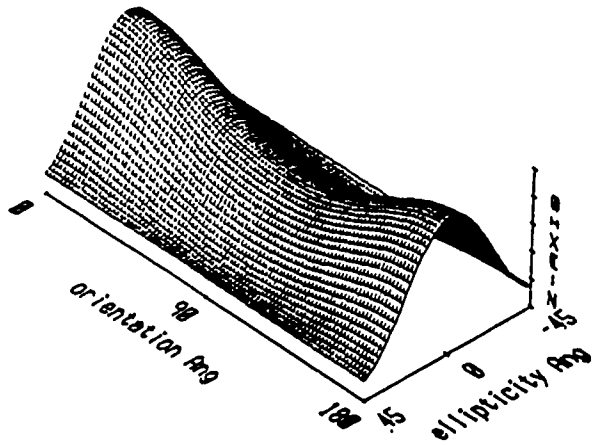


Figure 4. Copolarization signatures of a layer of randomly oriented spheroids based on first-order theory. *frequency* = $1.225GHz$, $\epsilon_s = (6.5+i0.5)\epsilon_o$, $a = 2.5cm$, $c = 0.25cm$, *fractional volume* = 0.0055, $d = 2.5m$, $\epsilon_2 = (15+i2)\epsilon_o$, and $\theta_o = \theta_s = 35°$. Orientation distribution is governed by $\beta_1 = 0°$, and $\beta_2 = 90°$. $\sigma = 5.161 \times 10^{-2}$ at $\psi = 180°$, $\chi = 0°$.

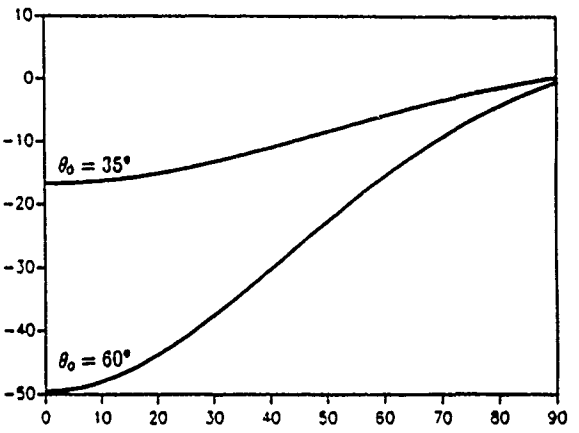


Figure 6. Phase difference ϕ_{vh} based on first order theory of a layer of spheroids as a function of orientation distribution β_2 for $\beta_1 = 0°$. At $\beta_2 \rightarrow 0°$ spheroids are aligned. Two angles of incidence $\theta_o = 35°$ and $\theta_o = 60°$ are considered. *frequency* = $1.225GHz$, $\epsilon_s = (6.5+i0.5)\epsilon_o$, $a = 2.5cm$, $c = 0.25cm$, *fractional volume* = 0.0055, $d = 2.5m$, and $\epsilon_2 = (15+i2)\epsilon_o$.

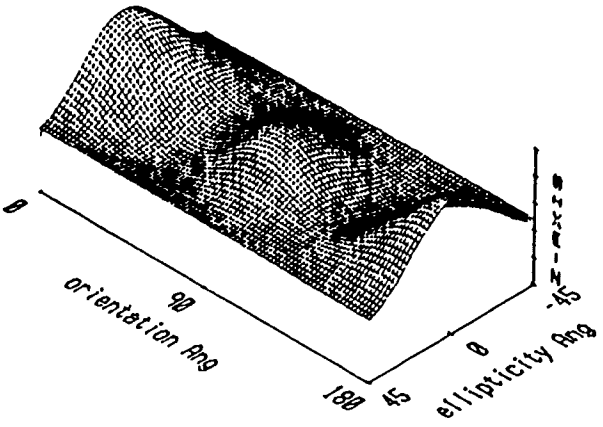


Figure 5. Copolarization signatures of a half-space of aligned spheroids based on second-order theory. *frequency* = $1.225GHz$, $\epsilon_s = (6.5+i0.05)\epsilon_o$, $a = 2.5cm$, $c = 0.25cm$, *fractional volume* = 0.0055, $d = 2.5m$, $\epsilon_2 = (15+i2)\epsilon_o$, and $\theta_o = \theta_s = 35°$. Orientation distribution is $\beta_1 = 0°$, and $\beta_2 \rightarrow 0°$. $\sigma = 0.649$ at $\psi = 180°$, $\chi = 0°$.

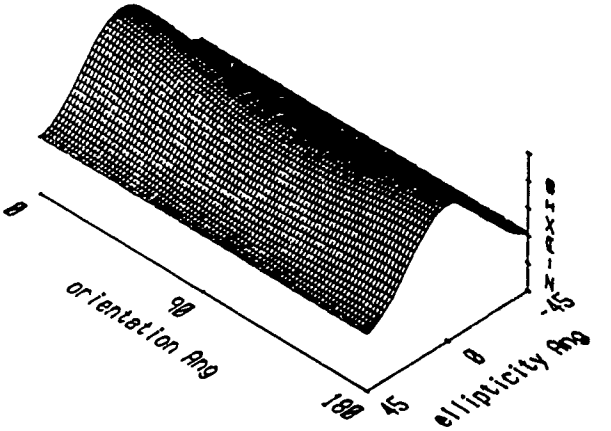


Figure 7. Copolarization signatures of a layer of densely distributed spherical particles based on the dense medium radiative transfer theory. *frequency* = $17GHz$, $\epsilon_s = (3.2+i0.002)\epsilon_o$, $d = 27cm$, $\epsilon_2 = (6+i0.6)\epsilon_o$, $\theta_o = 40.8°$, and $\theta_s = 35°$. The particle size distribution is described by $P = 6$, $Q = 4$, $a_c = 0.125cm$, and $f_{tot} = 0.2$. $\sigma = 0.440$ at $\psi = 180°$, $\chi = 0°$.

INFERENCE OF GEOPHYSICAL PARAMETERS FROM MULTIFREQUENCY POLARIMETRIC RADAR OBSERVATIONS AND MODEL INVERSION

J. J. van Zyl, C. F. Burnette and T. G. Farr

Jet Propulsion Laboratory
California Institute of Technology
Pasadena, CA 91109

ABSTRACT. During the summer of 1988 an intensive field experiment was conducted in the vicinity of Pisgah lava flow in the Mojave Desert in southern California. As part of the experiment, physical properties such as microtopography, composition, soil moisture and dielectric constant at five different sites representing surfaces with r.m.s. heights varying from less than one centimeter to tens of centimeters were measured. In addition, polarimetric radar images at P-band (67 cm wavelength), L-band (24 cm) and C-band (5.6 cm) were acquired at three different incidence angles with the NASA/JPL airborne imaging radar polarimeter. Using trihedral corner reflectors deployed in the area prior to imaging, the L- and C-band images were calibrated to provide σ^o values for each resolution element in the scene. This paper reports on the results of inferring geophysical parameters such as r.m.s. surface height and correlation length of the surfaces covered by our measurements by fitting the observed signatures with those predicted by the small perturbation model. The results obtained show that, for smoother surfaces, the r.m.s. height values inferred are in good agreement with *in situ* measurements.

Keywords. Radar Polarimetry, Radar Scattering, Model Inversion.

INTRODUCTION

Radar backscatter from rough surfaces is influenced by both the geometrical and electrical properties of the surface. The ability to characterize these physical properties of a rough surface is fundamental to answering many questions about the physical nature of the earth's surface. For example, the roughness of lava flows is usually a function of composition, eruption characteristics and age. The dielectric constant of land surfaces as measured at a particular microwave frequency is a function of, among other things, composition and soil moisture content.

During the week of 29 May - 3 June 1988, an extensive field experiment, the Mojave Field Experiment (MFE) was conducted in the Mojave desert of southern California. The overall objective of this experiment was to collect simultaneous data from airborne remote sensing instruments of geologically interesting targets, while at the same time measuring a number of characteristics of those targets on the ground.

This paper describes the results of inverting the radar data for surface power spectrum and r.m.s. height. Here, we shall restrict our discussion of results to inversions using the small perturbation model. First, we describe the ground data collected during the field experiment. This is followed by a description of the inversion technique used, and finally the results of the inversion are compared to the *in situ* measurements.

GROUND DATA COLLECTED

Within the field area, at Pisgah lava field, five sites were selected such that a wide range of roughness and geologic surface types were included. For each site a 10 m by 10 m area, relatively uniform in roughness and soil and rock type, was chosen for further study. In these 10 m by 10 m areas, detailed ground measurements, including L-band dielectric constant and soil moisture, were made. In addition, stereo photographs of each site were taken by a helicopter carrying twin metric framing cameras. These stereo photographs were later reduced to profiles of the surface height, allowing us to estimate the power spectra of the surfaces.

For the purposes of this paper, we shall focus attention on three of the six sites chosen. These three sites are briefly discussed below:

Playa. This site was chosen on the dry lake surface of Lavic Lake, approximately 1 km from the edge of Pisgah lava flow. The surface composition was measured to be 98% clay, with 2% basalt gravel, driftwood and metal fragments. Using a dielectric constant probe, the L-band dielectric constant was measured to be 2.36 + i 0.2. Measurements of the moisture content of the clay ranged between 1.72% to 2.1% water by weight, with a mean of 1.88% . No stereo photography of this site were recorded, so profiles recorded the previous year were used to estimate the r.m.s. height of the surface as 0.63 cm. Since only two profiles were available, no estimate of the error on this measurement can be given.

Pavement. This site was also located on Lavic Lake approximately 200 m south of the edge of Pisgah lava flow. Approximately 50% of the playa surface is covered with angular basalt fragments ranging from gravel to cobble size. The dielectric constant of the basalt fragments was measured in the laboratory to be about 5.0 + i 0 at both L-band and C-band. The r.m.s. height of the surface as estimated from profiles derived from stereo photographs, was measured to be 0.95 ± 0.23 cm.

Pahoehoe. This site was located on an older pahoehoe flow within the Pisgah volcanic field, approximately 500 m from the lake bed. The surface is fairly level with low undulations (approximately 5 m wavelength). Vegetation is widely spaced (less than 5% cover), and is composed primarily of low bushes, shrubs and grasses. There are some patches of bedrock present (about 9% of the surface), but most of the surface is covered with pebbles and gravels (62%). Clay and silt are found in the intersicles and vesicles of the lava. Laboratory measurements put the dielectric constant of the rocks at 4.5 + i 0 for both L-band and C-band. The r.m.s. height of the surface was measured to be 2.97 ± 1.24 cm.

RADAR DATA CALIBRATION

A calibration technique developed by us was used to calibrate the three-frequency radar data set. This technique, described in detail elsewhere (van Zyl, 1989), is based on the theoretical result that the co-polarized response (HH or VV) and the cross-polarized response (HV) of rough surfaces are uncorrelated. This result is used to estimate and remove system cross-talk from the radar data. The co-channel imbalance is then calibrated using the responses measured from trihedral corner reflectors that were deployed in the area prior to imaging. Finally, an overall radiometric calibration is performed using the trihedral corner reflector responses. The result is a three-frequency radar data set with calibrated σ^o values. The size of the corner reflectors used, plus the presence of spurious signals from unknown sources on the ground, compromised the utility of the P-band data. The interference from the unknown sources is especially severe for low-backscatter targets, such as the smooth playa, making the P-band results for this surface type suspect.

MODEL INVERSION APPROACH

The first-order small perturbation model was used to invert the radar data to estimate the surface power spectrum. Since the small perturbation model is only valid for very smooth surfaces, the technique was only applied to the smoothest three classes of surfaces present in the radar images. According to the first-order small perturbation model, the predicted radar cross sections of a smooth surface are given by (Tsang, *et al.*, 1985)

$$\sigma^0_{xx} = 4k_0^4 \cos^4(\theta)|\alpha_{xx}|^2 W(2k_0 \sin(\theta)), \quad xx = hh, vv \qquad (1)$$

where

$$\alpha_{hh} = \frac{\epsilon - 1}{(\cos(\theta) + \sqrt{\epsilon - \sin^2(\theta)})^2} \qquad (2)$$

$$\alpha_{vv} = \frac{(\epsilon - 1)(\epsilon(1 + \sin^2(\theta)) - \sin^2(\theta))}{(\epsilon\cos(\theta) + \sqrt{\epsilon - \sin^2(\theta)})^2}, \qquad (3)$$

$W(K)$ is the power spectrum of the surface sampled at the spatial wavenumber K, and k_0 is the wavenumber of the radar wave. If the incidence angle, θ, and the dielectric constant ϵ of the surface is known, the value of the power spectrum for spatial wavenumber $K = 2k_0 \sin(\theta)$ can be calculated from (1). It should be pointed out that the small perturbation model is strictly only valid for very smooth surfaces, i.e. (Ulaby *et al.*, 1982)

$$h < \lambda/20, \qquad (4)$$

and

$$\sqrt{2}h/l < 0.3, \qquad (5)$$

where h is the surface r.m.s. height and l is the surface correlation length. We note from the measured values quoted in the previous section that all three surface classes satisfy (4) at P-band, only the playa and the pavement sites satisfy (4) at L-band, and all three surfaces are strictly speaking too rough to apply the small perturbation model at C-band. We shall now describe how we used (1) to estimate the power spectra of the three surfaces described in the previous section.

The approach is shown schematically in Figure 1. First, the radar images were segmented into six classes using an unsupervised tone and texture clustering algorithm. For a description of the algorithm and features used, the interested reader is referred to a companion paper elsewhere in these Proceedings (Burnette *et al.*, 1989). Using this segmentation information, a σ^o versus incidence angle curve was derived for each frequency from the original radar data for each of the three classes used in the rest of this study as shown in Figure 1(b) for the playa class. Curves were derived for both σ^o_{hh} and σ^o_{vv}. Using the information about the dielectric constants measured in the ground data collection in (1), we then calculated three sections of the power spectrum using the multiple incidence angle information at each frequency, as shown in Figure 1(c). To estimate the r.m.s. height and correlation length of the surfaces, we fitted the function

$$W(K) = \frac{2h^2l^2}{(1 + K^2l^2)^{3/2}} \qquad (5)$$

corresponding to the power spectrum of a surface with an exponential correlation function to the segments of the power spectrum calculated from the radar data, as shown in Figure 1(d).

RESULTS AND DISCUSSION

In Figure 2(a) - (c) we show comparisons of the power spectra inferred from the radar data and those calculated from the measured surface profiles of the three surface types described earlier. While in general we observe very good agreement between the measured power spectra and those inferred from the radar data, a few additional conclusions may be drawn from the results presented.

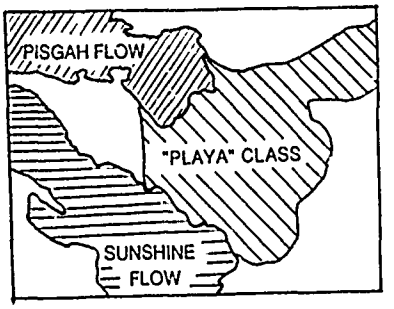
First, the results for the playa (Figure 2(a)) and the pahoehoe (Figure 2(c)) show that the functional form (5) assumed for the power spectrum is not correct for all surfaces. While this functional form appears to have the correct overall behavior, i.e. for small K, the power spectrum should be constant, while for large values of K the power spectrum should drop off as K^{-n}, $n = 3$ does not appear to be the correct value for the playa and the pahoehoe surfaces. Thus, a more general functional form, perhaps

$$W(K) = \frac{A}{(1 + K^2B^2)^{n/2}} \qquad (6)$$

should be used to fit the inferred segments of the power spectra.

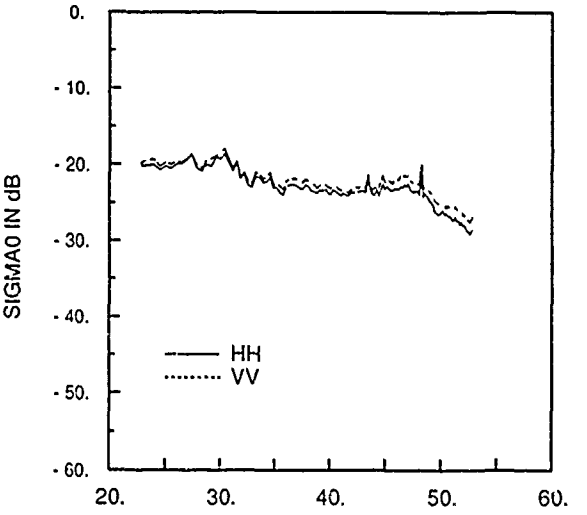
Secondly, we notice that the r.m.s. height of the playa was over-estimated by about a factor of 2.5 and the r.m.s. height of the pavement site was over-estimated by about 40%. While there may be numerous explanations for these results, we believe that, at least in the case of the playa surface, subsurface scattering may contribute to the over-estimation of the r.m.s. height. More work is needed, however, before any definite conclusions can be drawn.

In this paper, we have shown some results of estimating physical surface parameters, such as the power spectrum, from multifrequency radar data. While the present technique is only applicable to smooth surfaces, we have found good agreement between inferred parameters and those measured *in situ*. Future work will concentrate on extending the current technique to rougher surfaces.

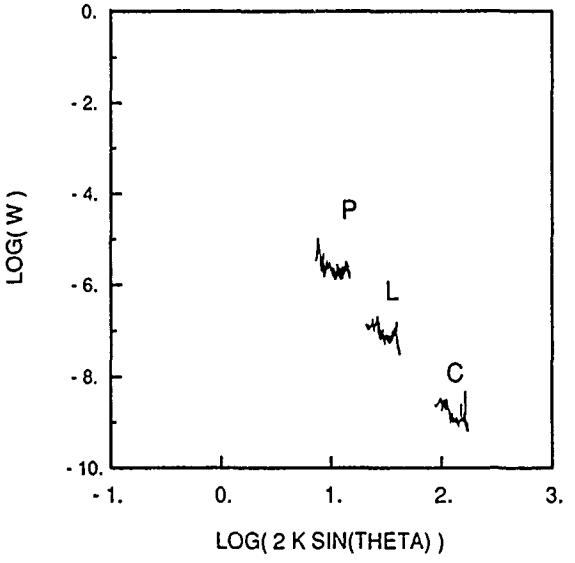


SEGMENTED IMAGE

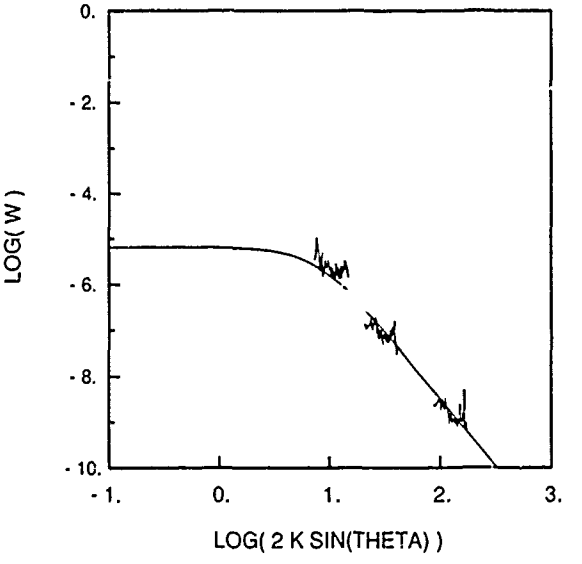
(a)



(b)

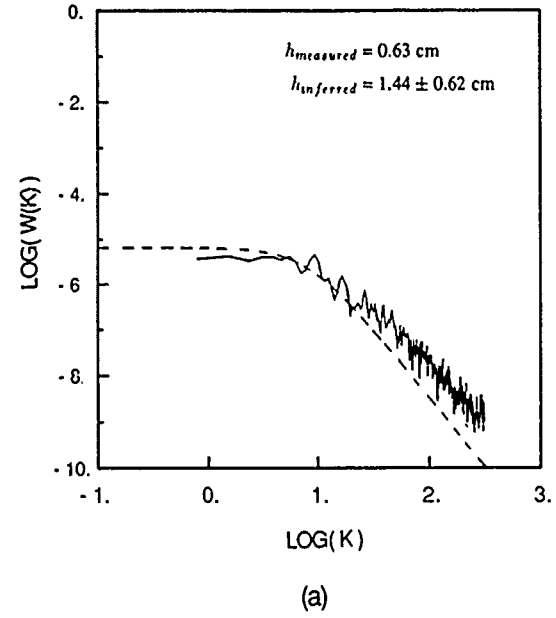


(c)

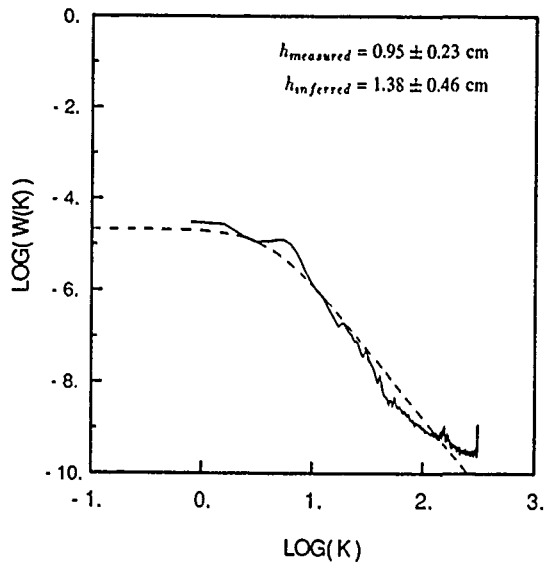


(d)

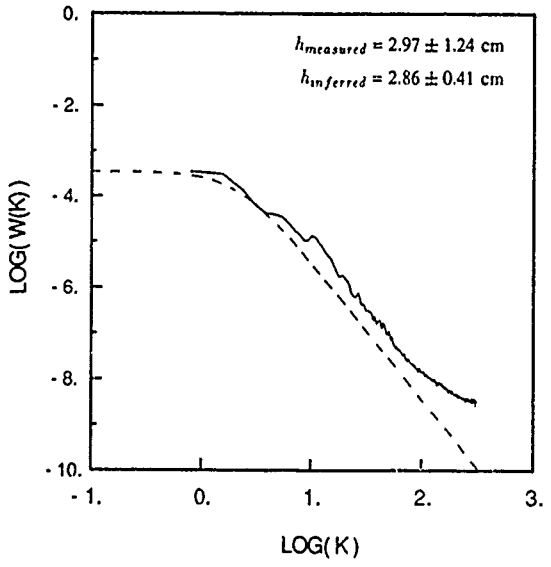
Figure 1. Schematic of inversion technique. First, the three-frequency radar images are segmented into six classes. Illustrated in (a) is the class with the smoothest surface type. The segmentation information is then used to construct a curve of σ^o as a function of incidence angle, at each frequency, as shown in (b) for the playa class. Three sections of the power spectrum, one corresponding to each frequency, is then calculated using the measured dielectric constant information. The result is shown (c). Finally, a function of the form given by (6) is fitted to the three segments of the power spectrum. The final result is shown in (d) for the playa class.



(a)



(b)



(c)

Figure 2. Comparison of the measured (full lines) and inferred (dashed lines) power spectra for the playa (a), the pavement (b) and the pahoehoe (c) classes in the Pisgah images. Also shown are the measured and inferred r.m.s. heights.

ACKNOWLEDGMENT

The research described in this paper was carried out by the Jet Propulsion Laboratory, California Institute of Technology, under a contract with the National Aeronautics and Space Administration.

REFERENCES

Burnette, C. F., P. C. Dubois and J. J. van Zyl, "Segmentation of multifrequency polarimetric radar images to facilitate the inference of geophysical parameters," IGARSS'89 Proceedings, 1989.

Tsang, L., J. A. Kong and R. T. Shin, *Theory of Microwave Remote Sensing*, Wiley & Sons, New York, 1985

Ulaby, F. T., R. K. Moore and A. K. Fung, *Microwave Remote Sensing, Active and Passive*, Volume II, Addison-Wesley, Reading, MA, 1982.

van Zyl, J. J., "A technique to calibrate polarimetric radar images using only image parameters and trihedral corner reflectors," Submitted to *IEEE Transactions on Geoscience and Remote Sensing*, 1989.

THE EFFECT OF AGRICULTURAL PARAMETERS ON RADAR CROSS-SECTION: A LINEAR REGRESSION MODEL

G.C. Hussey[1], G.J. Sofko[1], B. Brisco[2], J.A. Kochler[1], M.J. McKibben[1]

[1]*Institute of Space and Atmospheric Studies, University of Saskatchewan, Saskatoon, Saskatchewan, Canada*
[2]*Canada Centre for Remote Sensing, Ottawa, Ontario, Canada*

ABSTRACT

During the entire growing season of 1987, FM microwave scatterometer measurements in three frequency bands (L, C, and Ku), at ten incidence angles and at four polarization modes (HH, HV, VH, VV) were acquired in Saskatchewan for agricultural test plots of wheat. Concurrently with the radar measurements, both crop and soil data were collected. For identification of the relative importance of the agricultural parameters upon the radar signal, a linear correlation model between the differential scattering coefficient σ° and five agricultural parameters (topsoil and subsoil moisture, plant water content, dry green biomass, and leaf area index) has been developed. In this paper the regression model results are presented.

Keywords: scattering coefficient, agricultural parameters, linear regression, frequency band, polarization, row aspect angle.

INTRODUCTION

During the 1987 agricultural growing season radar measurements and corresponding crop and soil measurements on various agricultural crops common to Western Canada were gathered. This paper specifically is concerned with the measurements acquired on test plots of spring wheat. The variety of spring wheat grown was Katepawa.

MEASUREMENT PROCEDURE

Measurements for the differential scattering coefficient σ° were collected using microwave FM scatterometers operating at the three frequency bands L, C, and Ku. In each band, measurements were made in four polarization modes HH, HV, VH, and VV and at ten incidence angles ranging from 10° to 70°. In this paper only measurements of the like polarizations, HH and VV, are presented.

Concurrently with the radar measurements, agricultural measurements related to plant and soil characteristics were collected. Gravimetric soil moisture measurements were taken for depth ranges 0 to 5 centimetres and 5 to 15 centimetres and are denoted here as topsoil and subsoil, respectively. Canopy measurements included wheat canopy development interval, plant height, leaf area, and information required for biomass values.

The observations for the analysis presented here were taken on test Plot 13 from July 13, 1987 (Day Number 194) to August 26, 1987 (Day Number 238). The crop was 60 centimetres in height and the heads were starting to develop on July 13. The observations were gathered over the subsequent development stages - flowering, milk stage, soft dough - until August 26 when the crop was in the later stages of soft dough. It would have been interesting to divide this time period by about half, but there would have been insufficient observations for a meaningful multivariate linear regression analysis.

MICROWAVE DATA

In figure 1 a 3-dimensional graph of all the radar observations collected during the 1987 growing season for C-band HH polarization looking parallel to the furrows is presented. This graph is representative of the general shape of the graphs found for all the band-polarization combinations. Several key points can be observed from this graph: one, the radar signal varies over time, which indicates that the radar is sensitive to changes which occur in the scene, and two, more importantly, the lower incidence angles are more sensitive to changes which occur. Most of the temporal variation can be attributed to fluctuations in topsoil moisture content. At the lower incidence angles there are higher σ° values which appear to vary more when a change occurs. The observation that σ° is stronger at low incidence angles and that sensitivity to change is greater at low incidence angles are, to some degree, dependent on each other. This decrease in sensitivity as incidence angle increases has been observed by others (Ref. 1).

The model selection analysis, which will shortly be presented, tends to break down at the higher incidence angles. This may be explained by the fact that the radars are less sensitive to change at these angles. The radars may well be able to detect the changes, but the changes may be interpreted as random error due to the nature of linear regression.

REGRESSION ANALYSIS

Multivariate linear regression analysis and a model selection procedure were used to investigate how plant and soil characteristics influence the scattering coefficient σ°. The model selection process used is known as 'stepwise'. 'Stepwise' basically will only allow parameters into the model

which have met a preset significance level. The agricultural variables selected to be modelled with the radar measurements were: gravimetric topsoil and subsoil moisture content, dry green biomass, plant water content (expressed in absolute terms of kg/m^2), and leaf area index (LAI). These agricultural parameters were selected since they describe fundamental properties associated with scattering.

In figures 2a and 2b the results of the model selection procedure are presented. The radar measurements for each incidence angle have been modelled against the agricultural parameters and their contribution to the model, if any, is displayed in the bar graph. The predictor variables are those variables, in this case the agricultural variables, which influence the response variable, the scattering coefficient σ°. The graphs in figures 2 and 3 present the contributions of each individual predictor variable, which has been entered into a model, to the total multiple coefficient of determination, R^2. The multiple coefficient of determination indicates how well the predictor variables explain variation in σ° the response variable. An R^2 value of say, $R^2 = 0.88$, would indicate that the selected model is able to explain 88% of the variation in the response variable.

Radar measurements were collected looking both parallel (RAA = 0° were RAA stands for the Row Aspect Angle) to the furrows and perpendicular (RAA = 90°) to the furrows. Just as the human eye detects differences between observations made parallel and perpendicular to the furrows, it can be expected that differences in scattering will be detected at microwave frequencies. Consider a scene with a mature wheat canopy. When looking parallel to the furrows the human eye is able to detect soil between the furrows as well as the rows of plant material. This is especially true of the analysis presented here since, during the 1987 growing season a moderate drought was experienced. In contrast, when one looks perpendicular to the furrows, only plant material is detected. Of course, the radar signals penetrate the crops more effectively than visible light, but a difference between RAA 0° and RAA 90° is still expected.

DISCUSSION OF REGRESSION RESULTS

Some interesting effects can be seen by comparing the different row aspect angles. As well as differences in scattering between the frequency bands, there are differences between polarization states within a given frequency band.

For L-band HH polarization both row aspect angles indicate similar results; σ° is strongly influenced by topsoil moisture content over all incidence angles. The next predictor variable which shows a consistent contribution is subsoil moisture, although its contribution to the models is not as large as topsoil moisture. VV polarization shows the same type of result as HH with regard to topsoil moisture content. For VV, the contribution of plant characteristics is greater then for the HH polarization case. This is true especially at 60° and 70° incidence angle for RAA = 90°, where LAI becomes the primary contributor to the variation of σ°. This may be attributed to the fact that, at the high incidence angles, the incident radiation must transverse more crop, hence interacting more with the canopy before striking the soil. The fact that the wheat plants are vertically oriented may explain why this effect is present in VV but not HH polarization.

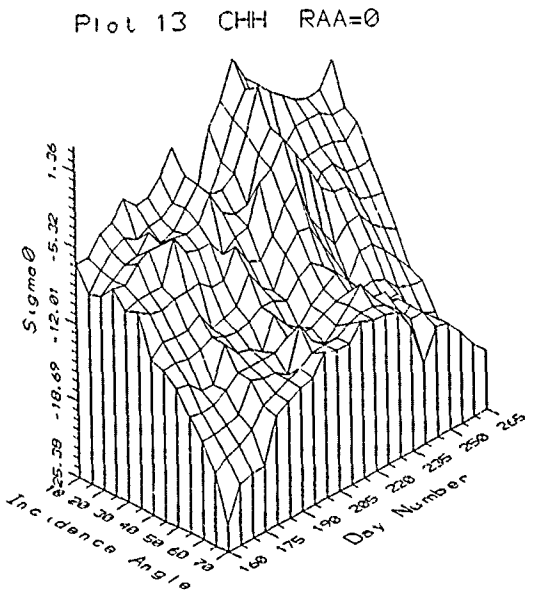


Figure 1: Three dimensional graph of σ° versus incidence angle and day number. This graph represents all the σ° measurements taken on wheat test Plot 13 at Row Aspect Angle 0° for C-band HH polarization for the 1987 growing season.

The fact that plant effects are indicated as being the major contributors to scattering for VV polarization at high incidence angles for RAA = 90° and not for RAA = 0° may be explained by the fact that at RAA = 0° the soil surface is only partially masked by the crop canopy, and hence the radar beam can penetrate to is able to interact with the soil.

Similarly to L-band, C-band HH polarization for both row aspect angles is primarily sensitive to topsoil moisture content except at the higher incidence angles of 60° and 70° where, unlike L-band, plant characteristics appear in the models. The graph of HH polarization at RAA = 90° indicates that most models fitted poorly. At the lower incidence angles for C-band VV polarization topsoil moisture content is the dominant contributing factor. For the high incidence angles at RAA = 0°, the linear model did not fit the data. This may result from one or more of the following: some parameters may just miss the significance level for entry into the linear model; a nonlinear model may be appropriate; the most important physical parameters are missing from the analysis. For RAA = 90° it may be observed that plant parameters become the major contributors to the models starting at lower incidence angles for VV polarization than for HH polarization. This is also the situation when comparing C-band to L-band for VV polarization at RAA = 90°.

An unusual effect in both the L-band and C-band modelling analysis is the change in the general modelling trend at incidence angles of 20° to 30°. At

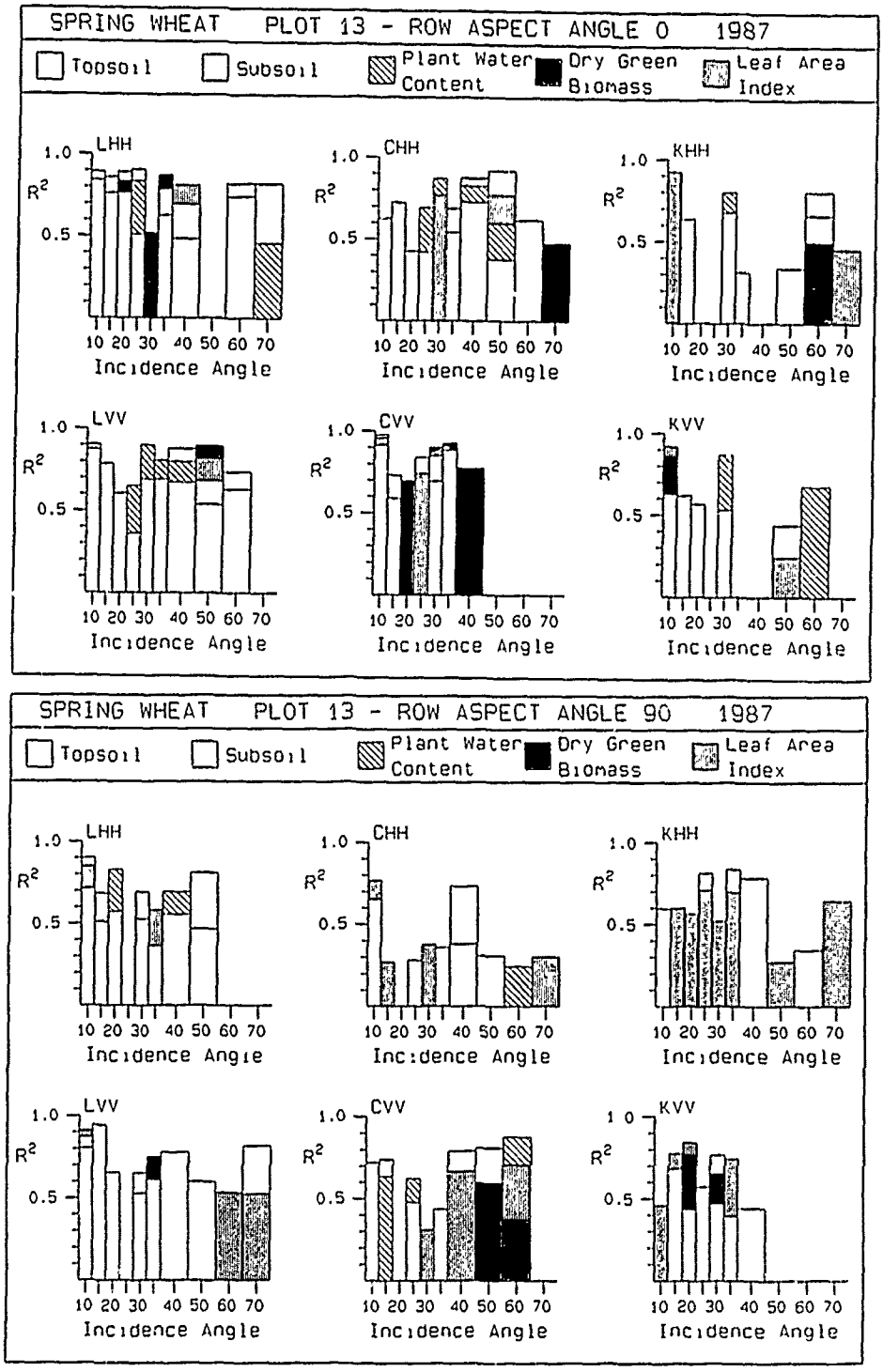


Figure 2: Bar graphs showing the contribution of the agricultural parameters, by use of the multiple coefficient of determination R^2, to the linear regression models. Figure 2a (top) represents measurements taken with the radars looking parallel to the furrows (RAA = 0°) and figure 2b (bottom) represents measurements taken when looking perpendicular to the furrows (RAA = 90°).

these angles, a plant characteristic is indicated as being the dominant contributor to the model or no fit to the model is indicated at all. This seems very unusual so, as a check the same analysis was performed on another test plot which was directly adjacent to Plot 13 and seeded on the same date, although at a lower density. Though not as pronounced, this trend to dominance by a plant characteristic was evident.

For the Ku-band radar, both HH and VV polarizations at a row aspect angle of 0° show incidence angles where no parameters fit the linear model. Where parameters do fit the models, the low incidence angles are soil dependent and the high incidence angles are plant dependent. For RAA = 90°, it is easily seen that HH polarization is much more sensitive to plant characteristics than is VV polarization. In comparing Ku-band HH polarization for row aspect angles 0° and 90°, it is obvious that at RAA = 90° the scattering is dominated by the plant characteristic LAI and at RAA = 0° this is obviously not the situation. The presence of topsoil moisture content at 10° incidence angle for RAA = 90° is obviously due to the ease of signal penetration to the ground at near vertical incidence. For VV polarization at RAA = 0° soil effects are dominant at the lower incidence angles and plant effects appear at the higher incidence angles. For RAA = 90° the lower incidence angles have both soil and plant characteristics appearing thus indicating scattering is due to both. No linear fit models were found for the higher incidence angles.

CONCLUSIONS

The linear regression analysis of agricultural parameters on the scattering coefficient σ° for a wheat canopy has, as have other papers [Ref. 1,5,6,7,8,9], shown the following:

- at L-band, soil parameters are dominant in the scattering process whereas at Ku-band plant parameters are dominant; at C-band, soil and plant parameter contributions are roughly comparable.

- the scattering process is affected by whether the radars are looking parallel or perpendicular to the furrows.

- for a given band, the polarization modes HH and VV indicate that scattering is orientation dependent; that is, scattering from a wheat canopy is anisotropic.

- for a row aspect angle of 90°, the linear regression models indicate that, for L-band and C-band, plant parameters are more important to the scattering process for VV polarization than for HH polarization, but in Ku-band this is reversed. A trend such as this is not obvious for a row aspect angle of 0°.

REFERENCES

1. Brakke, T.W., Kanemasu, E.T., Steiner, J.L., Ulaby, F.T., and Wilson, E.A., "Microwave Radar Response to Canopy Moisture, Leaf-Area Index, and Dry Weight of Wheat, Corn, and Sorghum", Remote Sens. Environ., 11, 1981, 207.

2. Danial, C. and Wood, F.S. Fitting Equations to Data 2nd Ed.. New York: John Wiley & Sons, 1980.

3. Draper, N.R. and Smith, H. Applied Regression Analysis 2nd Ed.. New York: John Wiley & Sons, 1981.

4. Hussey, G.C., Sofko, G.J., Koehler, J.A., Wacker, A.G., McKibben, M.J. and Brisco, B., "Agricultural Data Document for Ground Microwave Operations 1987 Campaign", Technical Report, University of Saskatchewan, Saskatoon, GMO-TR-003-88.

5. Le Toan, T., Lopes, A. and Huet, M., "On the Relationships Between Radar Backscattering Coefficient and Vegetation Canopy Characteristics", Proceedings of IGARSS'84 Symposium, Strasbourg, 27-30 August 1984, Ref. ESA SP-215.

6. Ulaby, F.T., Allen, C.T. and Eger III, G., "Relating the Microwave Backscattering Coefficient to Leaf Area Index", Remote Sens. Environ., 14 , 1984, 113.

7. Ulaby, F.T. and Batlivala, P.P., "Diurnal Variations of Radar Backscatter from a Vegetation Canopy", IEEE Trans. Antennas Propagat., 24:1, 1976, 11.

8. Ulaby, F.T., Moore, R.K. and Fung, A.K. Microwave Remote Sensing: Active and Passive Volume III: From Theory to Applications. Dedham: Artech House, 1986.

9. Ulaby, F.T. and Wilson, E.A., "Microwave Attenuation Properties of Vegetation Canopies", IEEE Trans. Geosci. Remote Sensing, 23:5, 1985, 746.

C-BAND BACKSCATTER SENSITIVITY TO MULTI-SCALE GEOMETRY AND SOIL MOISTURE VARIABILITY OF AGRICULTURAL SURFACES

A. Beaudoin, Q.H.J. Gwyn

CARTEL, Université de Sherbrooke
Sherbrooke, Québec, J1K 2R1, Canada
TEL: (819) 821-7180 FAX: (819) 821-7238

T. Le Toan

Centre d'Etude Spatiale des Rayonnements (CESR)
9, ave Colonel Roche, 31400 Toulouse, France
TEL: 61.55.66.71 FAX: 61.55.67.01

ABSTRACT/RESUME

This paper presents an experimental and theoretical study of the radar backscatter sensitivity to multi-scale geometry of agricultural fields. Theoretical modelling is validated using airborne C-band SAR data, and will aim at the derivation of simple soil moisture inversion models. The three geometric components of the surface have shown to induce significant backscatter variations: 1) at least 2 dB from the random roughness (soil clods); 2) up to 10 dB from periodic roughness (row crops and tillage) and 3) the large scale drainage pattern has induced periodic backscatter variations (up to 4 dB) due to both geometric effects and surface soil moisture profiles. The results show that soil moisture inversion from airborne SAR data is improbable. However, inversion using future spaceborne SARs appears highly probable because backscatter sensitivity to all surface components is minimized (2 to 4 dB) around 20° incidence angle, and the use of a simple inversion model should be possible.

KEYWORDS: airborne C-SAR data, backscatter modelling, agriculture, surface geometry, soil roughness, soil moisture.

I. INTRODUCTION

Estimates of soil moisture are of great importance in numerous environmental studies, including hydrology, meteorology and agriculture. Active microwave sensors have proven potential to provide soil moisture on a quantitative basis, in various operating circumstances [1]. Numerous previous scatterometer studies have shown the relative contribution of soil random roughness and surface moisture to the C-band backscatter, and have determined the best configuration for soil moisture detection: C-band, HH, incidence angle around 15° [2]-[4].

To prepare for future spaceborne C-band SARs, ERS-1 (1990) and RADARSAT (1994), for soil moisture estimation, one of the main objectives of Canadian airborne SAR agricultural experiments has been to go beyond scatterometers studies to a wider range of surface types and conditions encountered in the reality, in view of soil moisture inversion.

Following experimental results [5], this paper presents further investigations of C-band backscatter sensitivity to surface multi-scale geometry of bare fields/row crops at early growth stage. The approach is to 1) validate theoretical modelling using the SAR data set, and 2) extend observations of the limited C-SAR data set to a wider range of conditions (radar and scene), using the validated models.

2. EXPERIMENTAL SETTING

2.1 Site description

The test site is a 80 km² flat agricultural area, in the Québec St-Lawrence Lowlands, Canada. Land use consists of grain crops (corn, emergence stage; wheat and barley, end of tillerage stage), potatoes (emergence stage), pasture, grassland and wood stands. General azimuth of the rectangular fields is N40°E. Field surfaces show various combinations of three geometric components:

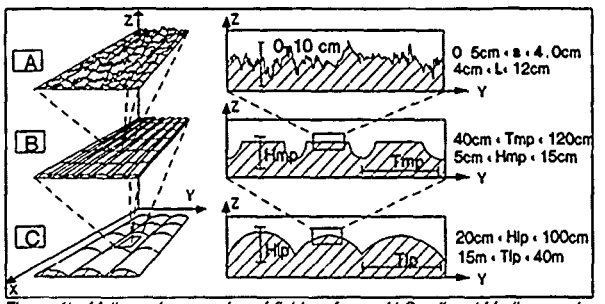


Figure 1) Multi-scale geometry of field surfaces. A) Small and Medium random roughness (soil clods), B) Medium Periodic roughness (row crops and tillage) C) Large Periodic roughness (drainage pattern).

1) Roughness of soil clods (Fig. 1a), due to natural aggregation, weathering and tillage, present on all fields at a small (mm) and medium (cm) scale. Mathematically, it is assimilated to a gaussian and isotropic random process (stationnary), described by the surface RMS height s and the correlation length l given by the correlation fonction $\rho(\zeta)$. We refer to this roughness as "Small and Medium Random (SR) roughness ".

2) Roughness of tillage pattern and row crop ridges, due to cultural practices (Fig. 1b). This medium-scale (cm) roughness is periodic in the y direction, and can be described by the height H_{mp} and spatial period T_{mp} of the pattern. We call it " Medium Periodic (MP) roughness ".

3) Roughness of large drainage pattern (Fig. 1c) is a large-scale (m) convex profiling of the surface, in the y direction. This permanent "swell and swale"

pattern is a common agricultural practices in the St-Lawrence Lowlands, where it is used to compensate for insufficient natural soil drainage. This pattern can be described by its height H_{lp} and spatial period T_{lp} and we will refer to it as " Large Periodic (LP) roughness ".

Observed roughness combinations are: SR / MP+SR / LP+MP+SR / LP+SR. The MP roughness is somewhat difficult to separate from the SR roughness (similar scale) For all combinations, along-field direction x is assumed to contain only the SR roughness. Many surface pattern variations can be found for a given land use, depending on cultural practices, weathering, soil type.

2.2 SAR data

The Canada Centre for Remote Sensing (CCRS) C-SAR Airborne Radar [6] was flown over the test site on June 11 1987. The images (amplitude 7-look) have been acquired in the HH mode with a high resolution mode (6m). To increase the SAR configurations, including the look angle ß (relative to field orientation) and incidence angle θ, three flight lines were used, N40°E (ß=90°), N130°E (ß=0°) and N160°E (ß=30°), with two acquisition modes : "Nadir" (20°<θ<74°) and "Narrow Swath" (45°<θ<76°), for a total of 6 images.

2.3 Ground data

Fifty (50) fields with varied land uses have been sampled for surface parameters. Soil gravimetric and volumetric moisture (M_g and M_v, top 5cm) was extracted along a diagonal transect. M_v show large variations (20% to 50%) due to heavy rains during the preceeding days. Usual canopy characteristics were acquired: crop type, growth stage, plant height and density, row direction and spacing.

Sampling of roughness parameters was done in 2 ways: 1) direct measurements of height and spatial period of MP and LP roughness, using a surveyor's level for the drainage pattern; 2) indirect measurements of RMS height s and correlation length l of SR roughness, from ground profiles acquired with an electronic device developed at CARTEL [7]. Paired profiles were acquired (3m long, 1cm step), one parallel and the other transverse to the periodic pattern, for 13 typical fields (tilled, harrowed, leveled, or row crops). s and l are derived in 2 fashions: 1) from parallel profiles (only SR roughness) or perpendicular ones without MP roughness; 2) from perpendicular profiles, spliting the roughness spectrum into its random and periodic parts.

3. METHODOLOGY

3.1 Extraction of relative backscattering coefficient

To retrieve the backscattering coefficient σ°_r from the images, a relative calibration within and between the images was necessary, as the C-SAR antenna gain pattern was not well compensated for by the STC gain. The calibration method is based on extended targets of known radar response [5], which consists of dense deciduous (maple) stands present in the scene. The deciduous angular backscatter behavior has been assessed using the Eom and Fung model [8], and can be approximated by a cosine law. A correction factor was derived from the comparison between the angular forest response retrieved from the images and the theoretical response.

3.2 Modelling of the backscattering coefficient

To model the backscattering coefficient σ° with respect to field types encountered, we integrated existing scattering models in a global model, which needs as input 1) the soil relative complex dielectric constant ϵ_r and 2) a set of parameters accounting for the surface geometry.

ϵ_r is simulated using the empirical model developed by Hallikainen et al. [9]. For the surface geometry, each roughness component has been characterized. The random roughness SR is accounted for in the modelling using s, l and a realistic correlation fonction $\rho(\zeta)$. Medium periodic roughness MP can be described by a function $Z_{mp}(y)$, depending upon the crop type. For example row crops (corn and potatoes) are seeded two rows at a time in the St-Lawrence Lowlands, resulting in a pattern of four ridges equally spaced, which can be approximated by a modified cosine function:

$$\begin{array}{l|l} \pi/2 < y_1 < 3\pi/2 & Z_{mp} = 0 \\ (4n+1)\pi/2-d < y_1 < (4n+3)\pi/2+d \quad n=1,2,3 & Z_{mp} = H_{mp}-H'_{mp} \\ \text{other } y_1 & Z_{mp} = H_{mp}\cos(y_1) \end{array}$$

with

$$y_1 = \text{mod}\,(yk_{mp},\ 8)$$
$$d = \pi/2 - \arccos(H_{mp}-H'_{mp})/H_{mp} \qquad (1)$$

where y_1 is the y position within one spatial period (8π), k_{mp} the ridge wavenumber and H'_{mp} and d extra parameters for inter-row ridges. Fig. 2 shows two simulations valid at survey time. This function can be used for many other row crops. Finally, the large periodic LP roughness Z_{lp} is approximated by:

$$Z_{lp} = H_{lp} \sin |(y\pi/T_{lp})| \qquad (2)$$

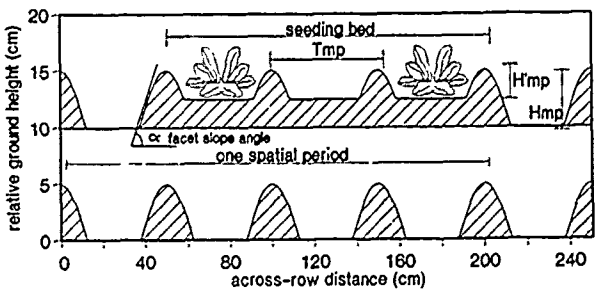


Figure 2) 2 Simulations of across-field periodic pattern of potato fields

For the soil and vegetation conditions encountered at survey time, the following assumptions are made. 1) volume scattering from sparse canopy and highly wet soil is negligible; 2) surface scattering is caused mainly by random roughness and 3) both periodic roughnesses (MP and LP) act as a local slope modulator of the surimposed random roughness [2]. Modelling is based on the three usual scattering models for random surfaces. small perturbations, physical optics and geometric optic models [10]. Depending on the different roughness combinations, three modelling approaches are used:

1) Fields with random roughness (SR): direct computation using the appropriate basic model;

2) Fields with medium periodic roughness (MP + SR): the model is based on Ulaby et al. [2] for perturbed periodic surfaces, using the deterministic function

Z_{mp}. As the SAR resolution cell contains many MP spatial periods, the simulation is done by computing the backscatter on a differential area of the surface and the total backscatter is obtained by incoherent integration along one spatial period of MP (shadowing effects also accounted for).

3) Fields with large periodic roughness (LP+SR, LP+MP+SR): the spatial period of the drainage pattern is greater than the resolution cell and approach # 2 is not valid. LP roughness induces a variation of the local incidence angle of the wave impiging on the overiding MP/SR roughness. The LP local backscatter can be derived using approach # 1 or # 2, coupled with the known variation of the local incidence angle.

4. RESULTS AND ANALYSIS

4.1 Sensitivity to Small and Medium Random (SR) roughness

Fig. 3a shows σ°_r as a function of the RMS height s, in two θ ranges (40°-50°, 60°-70°). We can point out the general increasing trend of σ°_r with s. For these two groups, the observed SR roughness range induces a $\Delta\sigma^\circ$ of 10 and 14 dB respectively. Ideal modelling was not possible as there was some missing parameters (M_v). However we can use the mean values of the available parameters, for example mean $\epsilon_r = 29\text{-}j7$ (M_v =33%). At these high incidence angles, the observed M_v range (20 to 50%) induces a $\Delta\sigma^\circ$ of about ± 1dB from the mean. Observed l has shown to have a slight relation with s, assumed to vary linearly from 4 to 8. This is an interesting point since realistic l values are often unavailable. The best correlation function compared to experimental curves has been found to be:

$$\rho(\varsigma) = [\,1 + (\varsigma/l)^2\,]^{-3/2} \qquad (3)$$

where ς is the displacement. Fig. 3a shows two simulations, using mean θ of 45° and 65°. Despite the data dispersion, there is a good agreement on backscatter behavior relative to s and θ; from a smooth surface (s=0.5) to a rough surface (s=3.5) $\Delta\sigma^\circ$ is from 10 dB to 3 dB.

Based on this validation, Fig 3b shows simulations for many incidence angles. For the entire RMS height range, SR roughness induces significant $\Delta\sigma^\circ$, for any incidence angles, and is never smaller than 2 dB. For θ>20°, σ° increases with s , for θ<20°, a reverse behavior is observed. SR roughness influence is minimized around θ=20°.

Most fields in early spring show a large RMS height range due to tillage practices (s=1.5 for harrowing, to s=3.5 for coarse tillage). Induced $\Delta\sigma^\circ$ at ERS-1 incidence angle (θ = 23°) and RADARSAT minimal incidence angle (20°) is 3 dB and up to 12 dB at maximum RADARSAT incidence angle (60°). As the season progresses, SR roughness will be smoothed by smaller tillage and weathering. Hence the range of s will drop under 1.0, with a narrow range but maximum backscatter sensitivity. For ERS-1, the induced $\Delta\sigma^\circ$ would be around 2dB and up to 6dB for maximum RADARSAT incidence angle.

4 2 Sensitivity to Medium Periodic (MP) roughness (row crops)

We have restricted our study to row crops including cereals, potatoes and corn. These two last are studied together (similar MP roughness). We define σ°_β as σ° at a look angle ß, and $\sigma^\circ_\perp$ and $\sigma^\circ_\parallel$ for the particular cases of perpendicular and parallel ß. A useful parameter here is the backscatter modulation function M(ß,θ) [2], defined as the σ°_β to $\sigma^\circ_\perp$ difference (dB) at a given incidence angle θ.

Corn and potato crops

Fig. 4 presents relative $\sigma^\circ_\perp$ and $\sigma^\circ_\parallel$ values for a) corn and b) potato. For the available θ range,

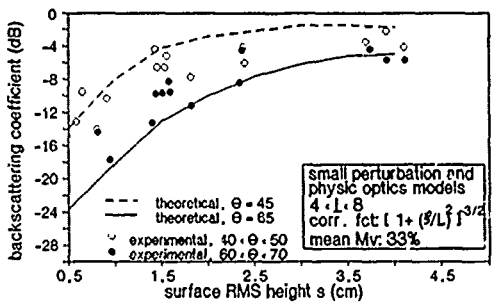


Figure 3a) Experimental and theoretical backscatter, function of RMS height for 2 incidence angle ranges (40-50 and 60-70)

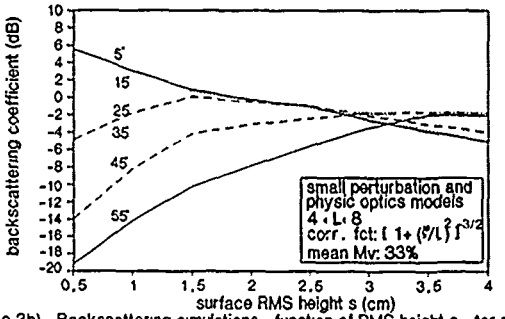


Figure 3b) Backscattering simulations, function of RMS height s, for many incidence angles

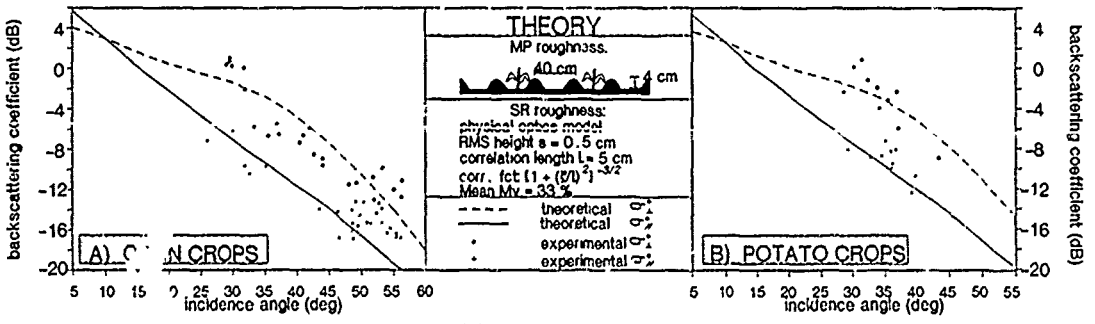


Figure 4) Experimental and theoretical backscattering coefficient, with parallel and perpendicular look angle, for A) corn and B) potato crops

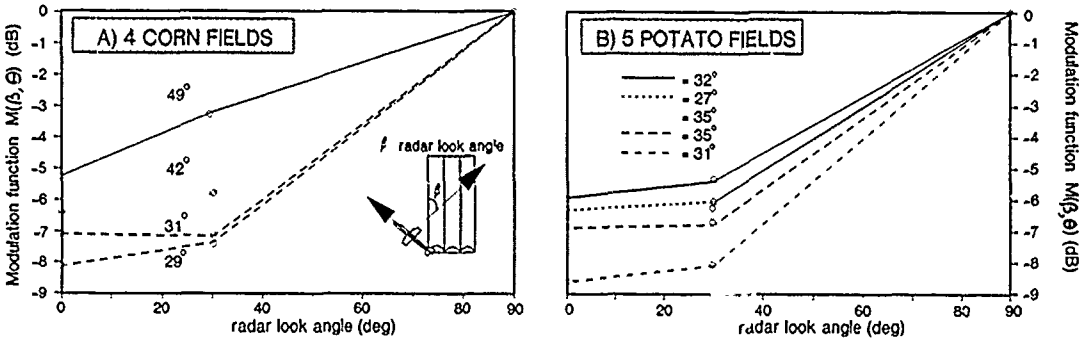


Figure 5) Experimental modulation function values M(β,θ), at three radar look angles (0, 30 and 90), for a) corn and b) potato fields

M(‖, θ) is maximum around θ=30° (8-9 dB), and decreases for higher θ, down to about 5 dB. Fig. 4 also presents $\sigma^\circ_\perp$ and $\sigma^\circ_\parallel$ simulation results which are in good agreement with the data set. Significant look angle effects are expected for any θ, except in the 5°-15° range. The maximum effect is found to be in a range of 25°-45°, related to the range of soil facet slopes α facing the radar (Fig. 2).

We have further analyse these effects using experimental M(ß,θ) values for a) corn and b) potato, for ß= 0°,30° and 90° (Fig. 5). In both cases, no significant differences are observed between M(‖,θ) and M(30°,θ), for θ< 40°. For corn, M(‖,θ) is greatest around θ = 30° (7-8 dB) and decreases for higher θ. For potato, the fields are viewed with similar θ (27°<θ<35°), but M(‖,θ) and M(30°,θ) show greater variations (6 to 9 dB). This is caused by the fluctuations of ridges height (H_{mp}) which is the main parameter influencing $\sigma^\circ_\perp$. At survey time, ridge heights ranged from 2 to 8 cm, inducing a M(‖,θ) fluctuation up to 6dB. However, as we progress through the season, MP roughness will suffer weathering, which will smoothen its periodic characteristics. Consequently the look angle effect will vanish.

Cereals crops

Fig. 6 shows the angular variations of relative $\sigma^\circ_\perp$ and $\sigma^\circ_\parallel$ values for barley fields. For comparison, the curves of experimental σ° of wheat fields (same growth stage), acquired with the RAMSES scatterometer [4], have been added. M(‖,θ) is again maximum around θ=35° (≈ 3 dB). M(ß,θ) behavior is similar to the one of corn and potato crops, except for a lower amplitude, caused by the smaller ridges of the wheat seeding bed. Therefore, no simulations have been done for this case.

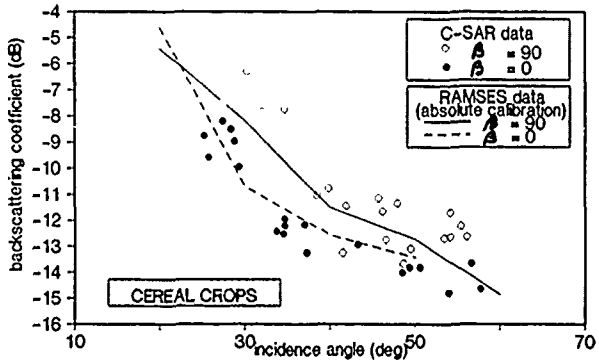


Figure 6) Backscattering coefficient of cereal crops, for parallel and perpendicular look angle

4 3 Sensitivity to the drainage pattern (LP roughness)

Fig. 7a shows a sub-scene including fields with drainage pattern. The most striking feature is the linear texture parallel to field orientation, appearing more clearly in Fig. 7b which has been texturally enhanced, using a mean filter and a vertical gradient detection. Fig. 8a presents 4 across-field profiles of the mean relative backscattering coefficient $\overline{\sigma^\circ}_r$, for the potato field pointed out in Fig. 7a. $\overline{\sigma^\circ}_r$ difference between the 4 radar configurations are caused by the MP roughness, as seen in section 4.2). Also $\overline{\sigma^\circ}_r$ is periodically modulated, with local minima and maxima approximately on the tops and bottoms of the drainage pattern ridges. Fig. 8b shows the power spectrum of profile # 3 in Fig. 8a. The peak in the 9th harmonic (i.e. 28m = LP roughness spatial period) indicates a

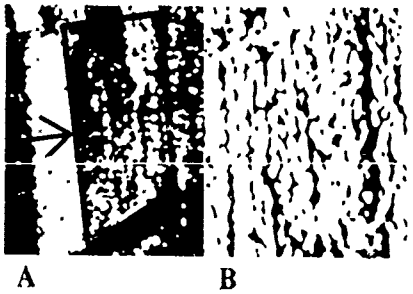


Figure 7) Sub-scene with drainage ridges

A) Original sub-scene
B) Textural enhancement

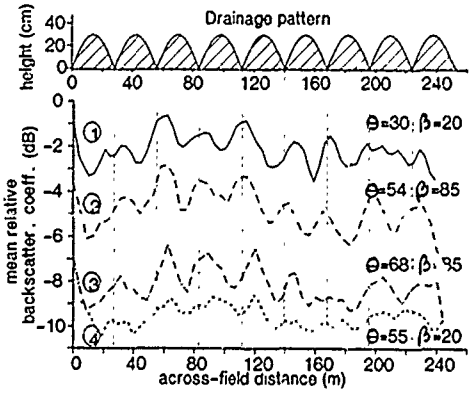


Figure 8a) Profiles of rel. mean backscatter of a field with drainage ridges (Fig. 7a)

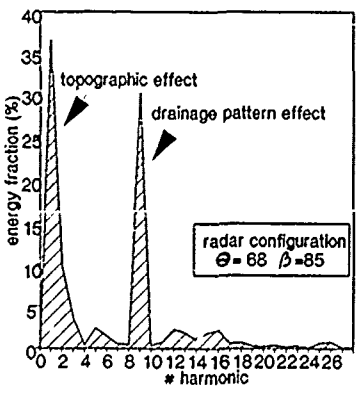


Figure 8b) Power spectrum of profile #3

strong relationship between drainage pattern and σ°. Notice also the low frequency peak, which reveals large-scale σ° variations, probably induced by the general field slope. σ° variations can be explained by two different effects: the drainage pattern geometry and the soil moisture variability, which should be related to a combined effect on σ°. Other parameters can be considered constant within a field.

The LP geometry causes a variation of the local incidence angle θ_l of the incoming signal : $\theta_l < \theta$, $\theta_l = \theta$ and $\theta_l > \theta$ on respectively the ridge front, top and back side. It can be shown that θ_l is given by:

$$\cos(\theta_l) = \frac{\pi\,\Omega\cos(y\pi/T_{lp})\,\sin\theta\sin\beta + \cos\theta}{\sqrt{1+[\pi\,\Omega\cos(y\pi/T_{lp})]^2}} \qquad (4)$$

where Ω is the LP ridges amplitude-to-period ratio. Thus deviation of θ_l from θ depends on 1) radar look angle β and 2) Ω ratio, ranging from 0.00 to 0.05, with a typical value of 0.020. For β perpendicular, θ_l is typically up to $\pm$ 3° away from θ. However the overall σ° local variation will depend on the σ° sensitivity of the MP/SR roughness overiding LP $[\delta(\sigma^\circ)/\delta(\theta_l)]$. Thus under the tilting effect, a rough (lambertian) surface would create no σ° variations, compared to a smooth surface (fast decay of σ° with θ_l) creating large σ° variations. All observed MP/SR surfaces fall between these two cases.

Looking back to Fig. 8a, we observe in effect significant $\Delta\sigma^\circ_r$ with β nearly perpendicular (profile #2 & 3). However, the profile #1 at β=20° (almost parallel) shows as strong variations as in #2-3, which cannot be explained by the LP geometric effect alone.

The drainage pattern induces lateral water run-off toward the hollows, which should result in an across-ridge soil moisture profile. The moisture sample strategy was modified during the 1988 C-SAR campaign (same test site), to account for the drainage pattern. For some selected fields, soil moisture transects have been acquired across many ridges, then were normalized and averaged. For example, Fig. 9 shows the mean across-ridge wetness profile of a corn field, approximated by a 2nd order polynomial given by:

$$M_{gl}(y) = \frac{(y_{lp}-T_{lp}/2)^2}{(T_{lp}/2)^2} * \Delta M_g + M_{gmin} \qquad (5)$$

where M_{gl} is the local gravimetric moisture, y_{lp} and T_{lp} respectively the LP distance and spatial period, ΔM_g the top/bottom M_g difference and M_{gmin} the relative minimum moisture (hollows). This approximation has shown to be valid in most cases encountered.

We can model the relative contribution of both drainage pattern geometry and soil moisture variability to the local backscattering coefficient σ°_l, assuming these two contributions are independent [4]. Thus σ°_l can be expressed in dB as:

$$\sigma^\circ_l = \sigma^\circ_{lm}(\theta_l, M_{vl}) + \sigma^\circ_{lg}[s,l,\theta,\beta,\Omega,\theta_l,(\delta\sigma^\circ/\theta_l)] \qquad (6)$$

where σ°_{lm} and σ°_{lg} are respectively the local soil moisture and geometric contribution to σ°_l. σ°_{lm} is directly proportional to the local Fresnel reflectivity, depending on the local dielectric constant ϵ_{rl}, derived using (5) in the Hallikainen model, and the local incidence angle θ_l; σ°_{lg} is obtained from the simulated σ° angular curves of the MP/SR roughness (as in Fig. 4) overiding LP roughness, and the θ_l deviation from θ.

Fig. 10 shows the experimental and simulated relative σ°_r of a) profile #1 and b) profile #3 of Fig. 8a. All the curves are relative as we are interested in the backscatter variations. In Fig. 10a the geometric contribution is negligible because the radar is viewing almost parallel to low ridges. Hence, in this case, σ°_l is pratically proportional to local soil moisture, which suggest that this SAR configuration (β parallel and low θ) should be used in applications related to soil moisture estimation and drainage management. This configuration eliminates the sensitivity to MP roughness, and only the sensitivity to SR roughness persists but is minimal (2dB) at low θ (20°).

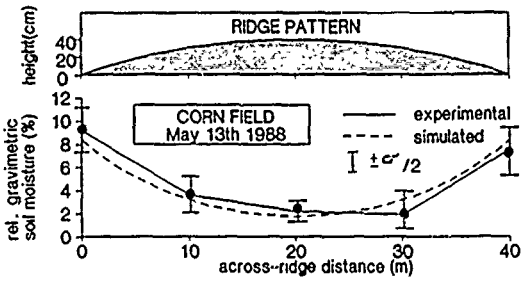


Figure 9) Relative mean across-ridge soil gravimetric moisture

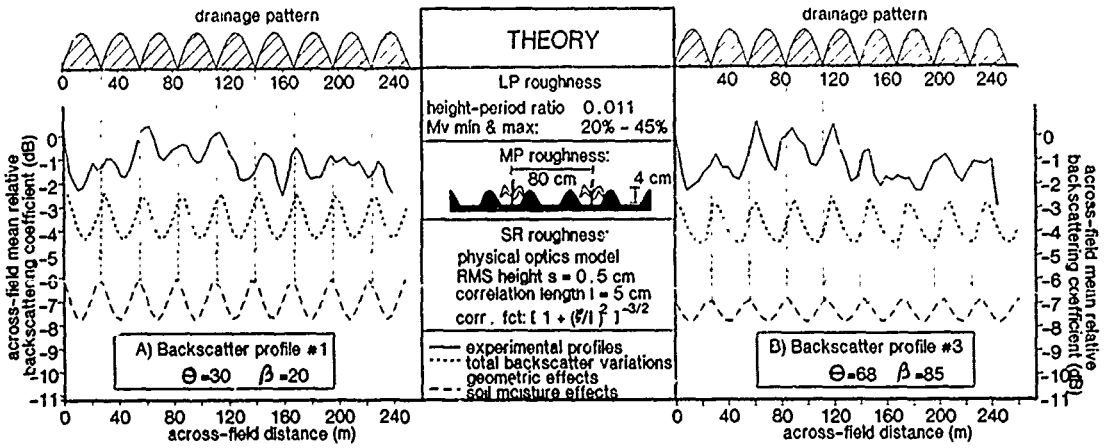


Figure 10) Simulations of backscatter variations due to drainage pattern geometry and related soil moisture pattern for A) profile #1 and B) profile #3 in Figure 8a)

In contrast,in Fig 10b the geometric contribution is clearly dominant, due to a perpendicular look angle. In addition the soil moisture contribution is lower, at higher incidence angle compared to the preceeding case. Hence local σ°_1 is proportional to the drainage pattern Such a SAR configuration (ß perpendicular and high θ) maximizes the sensitivity to all scales of roughness and should be used in applications requiring knowledge of local surface local topography for drainage management.

These two cases show that different parameter contributions can accidentally result in almost identical σ° variations, under two opposite radar observation configurations. Thus theoretical modelling is essential to avoid misunderstanding of the observed phenomena Globally, experimental and modelling results have shown that the local surface geometry and soil moisture variability induced by the drainage pattern (mean periodicity = 25m) is detectable in high resolution C-SAR imagery (= 6m).

5. CONCLUSION AND FUTURE WORK

This study has confirmed and demonstrated geometric effects of agricultural surfaces on the radar backscatter, using for the first time airborne C-SAR images. It has integrated and extended previous studies to a complete agricultural scene with a large variety of surface roughness states.

In particularly, the influence of random roughness has been confirmed, it is never smaller than 2dB, for any SAR configuration. Also, the influence of periodic roughness (row crops and tillage) can be as strong as 10 dB. To interprete the measurements, a global model is necessary with a complete set of terrain parameters which are difficult to obtain. Thus soil moisture mapping from single configuration airborne SAR data is highly improbable. However the use of multiple polarimetric/temporal airborne SAR images should be investigated.

Using ERS-1 and RADARSAT with an incidence angle around 20°, the expected maximum roughness effect (random and periodic) is 4 dB, which decreases around 2 dB if 1) the look angle is lower than 30°, or 2) fields are oriented in a single direction, which is common in Canada. Notice also that the 4dB perturbing effect is the worst possible case since the observed roughness range is maximum at the survey time. Thus soil moisture inversion, from ERS-1 and RADARSAT data, using simple model, appears highly possible.

The drainage pattern and induced soil moisture variability, typical of the St-Lawrence Lowlands, has been remotely detected for the first time, using active microwave sensors. Both geometric and wetness contributions to the C-band radar backscatter has been demonstrated, and it is possible to eliminate the geometric influence in order to detect the soil moisture profile. This drainage monitoring potential is interesting with respect to crop yield prediction. For example, in the St-Lawrence Lowlands, the drainage pattern enhances the overall drainage properties but causes saturation in the hollows and hence decreased crop yield. This problem is a concern to agronomists and farmers, and drainage monitoring using SAR data can help to study the best surface management with respect to crop yield.

To summarize, this study has shown the complex and multiple surface geometric effects of a typical agricultural scene, and modelling has been necessary to understand the radar backscatter behavior of such targets. In a next step, the validated models will be used to simulate radar backscatters for a larger range of both scene and SAR configurations, to assess precisely 1) the soil moisture inversion possibilities and 2) potentiality of multiple SAR observation/polarimetric/temporal configurations.

6. ACKNOWLEDGMENTS

This study as been possible with a NSERC fund (#4280, Radar appliqué aux ressources non-renouvelables) and a NSERC post-graduate award. We thank the Canada Centre for Remote Sensing (CCRS), who provided us with the C-SAR images and technical support. Special thanks to the microwave remote sensing group of the Centre d'Etude Spatiale des Rayonnements (CESR), where the modelling work reported in this paper has been realized.

7. REFERENCES

[1] Jet Propulsion Laboratory (JPL)," Shuttle Imaging Radar-C Science Plan ", *JPL Publication 86-29*, Passadena, CA, Sept. 1986.

[2] F.T. Ulaby, R.K. Moore and A.K. Fung, " Microwave Remote Sensing: Active and Passive", vol. III, ch. 21, *Norwood, MA: Artech House*, 1986.

[3] F.T. Ulaby, P.P. Batlivala and M.C. Dobson, " Microwave backscatter dependance on surface roughness, soil moisture, and soil texture, part I-Bare soil ", *IEEE Trans. Geosci. Electron.*, vol.GE-16,no. 4, pp.286-295, Oct. 1978.

[4] T. Le Toan, " Active microwave signatures of soil and crops: significant results of three years of experiments ", *IGARSS'82 Symp.*, Munich, FRG, June 1-4 1982.

[5] A. Beaudoin, T. Le Toan and Q.H.J. Gwyn, " Observations on the effect of geometric properties of agricultural soils on radar backscatter, using C-SAR data ", *IGARSS'88 Symp.*, Edinburgh, Scotland, 10-14 Sept. 1988.

[6] C.E. Livingstone et al, " CCRS C-SAR Airborne Radar- System description and test results ", *11th Canadian Symp. on remote sens.ng*, Waterloo, Ontario, June 22-25 1987.

[7] A. Beaudoin, Q.H.J. Gwyn, P. Cliche and G. Bordeleau," An economical electronic roughness sampler for radar land studies ",*IGARSS'89 Symp.*, Vancouver, 10-14 July 1989.

[8] H.J. Eom and A.K. Fung, " A scatter model for vegetation up to Ku-band ", *Remote Sensing of Environment*, vol. 15, pp. 185-200.

[9] M. Hallikainen, F.T. Ulaby, M.C. Dobson, M. El-Rayes and L.K. Wu, " Microwave dielectric behavior of wet soil- Part 1. empirical models and experimental observations ", *IEEE Trans.on Geosci. and Remote Sensing*, vol. GE-23, pp.25-34, Jan. 1985.

[10] F.T. Ulaby, R.K. Moore and A.K. Fung, "Microwave Remote Sensing: Active and Passive; ch 12: Introduction to random surface scattering and emission", vol. II, *Reading, MA: Addison-Wesley*, 1982.

EXPERIMENTAL INVESTIGATION OF THE DEPENDENCE OF RADAR BACKSCATTERING ON WIND SPEED, WIND STRESS AND WAVE HEIGHT

S. P. Gogineni and K. B. Katsaros

Radar Systems and Remote Sensing Laboratory
University of Kansas Center for Research, Inc.
2291 Irving Hill Road
Lawrence KS 66045-2969, USA
913/864-4835
FAX: 913/864-7789, TELEX: 706352, OMNET: KANSAS.U.RSL

Department of Atmospheric Sciences
University of Washington
AK-40
Seattle WA 98195
206/543-1203
FAX: 206/543-9295 (gen. univ.), OMNET: K.KATSAROS

ABSTRACT

During the summer of 1988, we performed radar measurements in conjunction with detailed environmental observations on Lake Washington at the University of Washington Sand Point field station. We collected radar data at 5.3 and 10 GHz for incidence angles between 30 and 60° with VV-polarization. The environmental measurements included wind speed and direction, large-wave heights, the high frequency portion of the wave spectrum, humidity, and air and water temperatures. We measured the small-scale wave spectrum using a resistance wire gauge.

The results show that backscatter increased with wind speed as expected. However, we did not observe much difference in the scattering coefficient for upwind and crosswind directions. The results also indicated an increase in the amplitude of small waves with friction velocity u^*.

In this paper we are presenting a brief description of the system used for radar experiments and preliminary results of the experiments conducted during the summer of 1988.

INTRODUCTION

During the summer of 1988, we conducted a pilot experiment to measure radar backscattering from water on Lake Washington. We conducted this experiment from the University of Washington Sand Point facility. We collected radar data at 5.3 and 10 GHz for upwind and crosswind directions over incidence angles form 30 to 60° with VV-polarization. We used a coherent FM radar for backscatter measurements. We collected environmental data that included mean wind speed, wind stress, temperature, humidity and wave height. The wind speed was measured at elevations of 2, 4 and 7 meters using a Gill propeller anemometers mounted on vanes. The wind stress was measured with the eddy-correlation technique utilizing a K-Gill propeller vane system [Atakturk and Katsaros, 1989]. The wave height was measured using a special-purpose resistance wire gauge.

The results show that backscatter increased with wind speed as expected. However, we did not observe much difference in the scattering coefficient for upwind and crosswind directions.

In this paper, we are providing a brief description of the radar used for the measurements and the preliminary results of the 1988 experiment.

SYSTEM

We used a dual-frequency FM radar for measurements during this experiment. The system is designed to measure the range to the target and the amplitude and phase of the backscattered signal. We used horn antennas for transmit and receive. The important specifications of the system are given Table 1, and a more detailed description is available in Gogineni et al., 1989.

Table 1

FM Radar Parameters	X Band	C Band
Radar Frequency	10 GHz	5.3 GHz
Modulating Waveform	Triang.	Triang.
Modulating Frequency	50 Hz	100 Hz
Sweep Bandwidth	1 GHz	500 MHz
Phase Accuracy (est.)	20°	20°
IF Weighting	Programmable	
Combined rf/antenna wt.	38 kg	

EXPERIMENT

We performed the pilot experiment over a period of two weeks during August 1988. We collected data at incidence angles between 30 and 60° with VV-polarization at 5.3 and 10 GHz. The antenna was pointed in the upwind or crosswind directions, and data were collected at each incidence angle for the two frequencies in five-minute segments. The system was calibrated before and after each data run by switching in a delay line of known length and loss in place of the antennas. Simultaneously with the radar measurements, all the relevant environmental variables including wind speed and direction, wave height, air and water temperature, etc. were recorded.

RESULTS

Figure 1 shows the results of wave height measurements made with the resistance wire gauge. The shift in the peak of the low-frequency part of the spectrum is related to the duration of the wind. There is little variation in the wind-generated high-frequency part of the wave spectrum. Figure 2 shows the wind speed and direction measured during a radar sampling period. As can be seen, wind direction and wind speed are fairly constant during the period over which wave height and radar measurements were made.

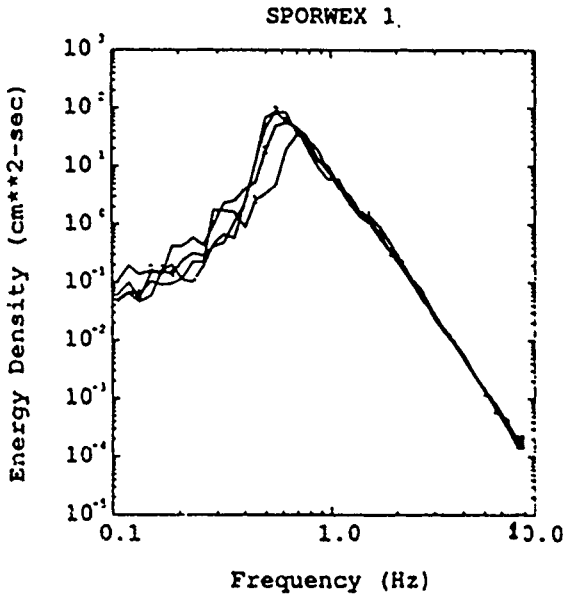


Fig. 1 Measured spectra using resistance wire gauge

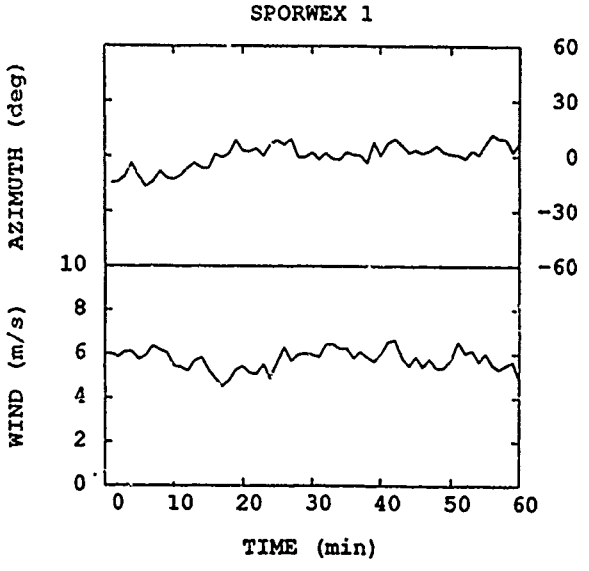


Fig. 2 Wind and direction during wave height measurements

Figures 3 and 4 show the angular response of the scattering coefficient at 5.3 GHz for upwind and crosswind directions, respectively. There is a linear decrease of σ° with speed between 30 and 60°. The results show that σ°s for upwind and crosswind directions are similar. This result is contrary to what has been observed in open ocean experiments. At this time, we do not know the reason for this discrepancy. However, it may be related to the difference in large-wave conditions between the open ocean and Lake Washington.

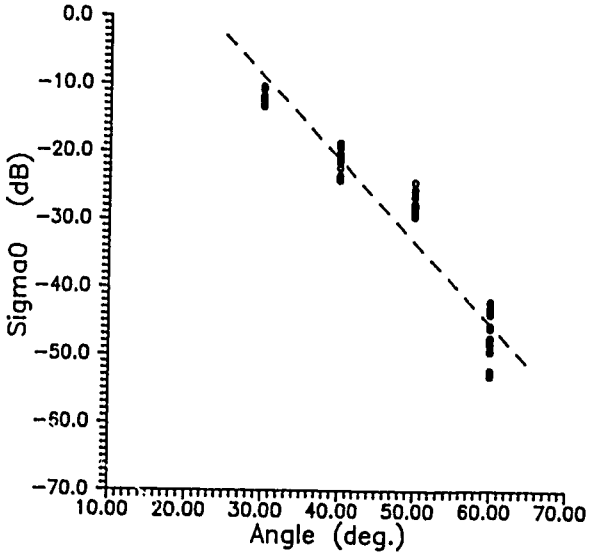


Fig. 3 σ° vs. incidence angle
C band, upwind look

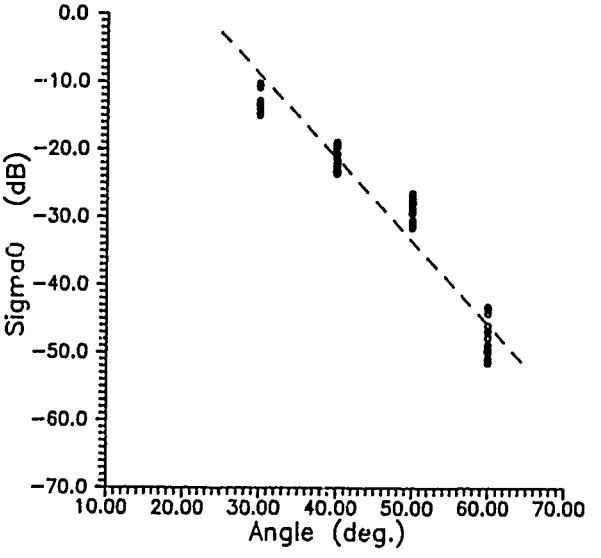


Fig. 4 σ° vs. incidence angle
C band, crosswind look

Figures 5 and 6 show the scattering coefficient dependence on upwind and crosswind directions. As expected, σ° increased, in most cases with wind speed.

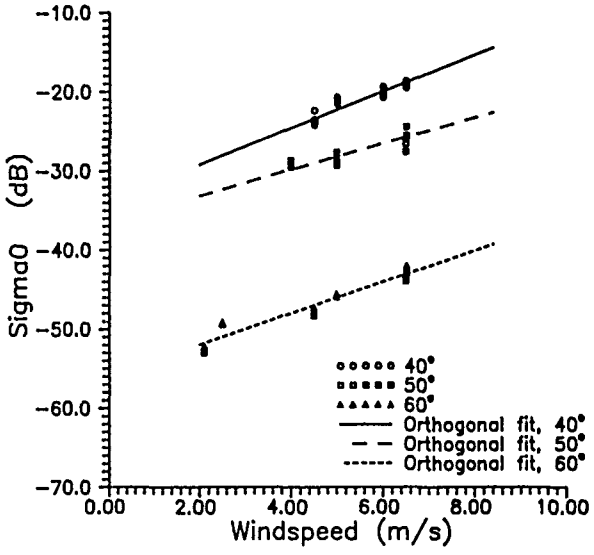


Fig. 5 σ° vs. wind speed
C band, upwind look

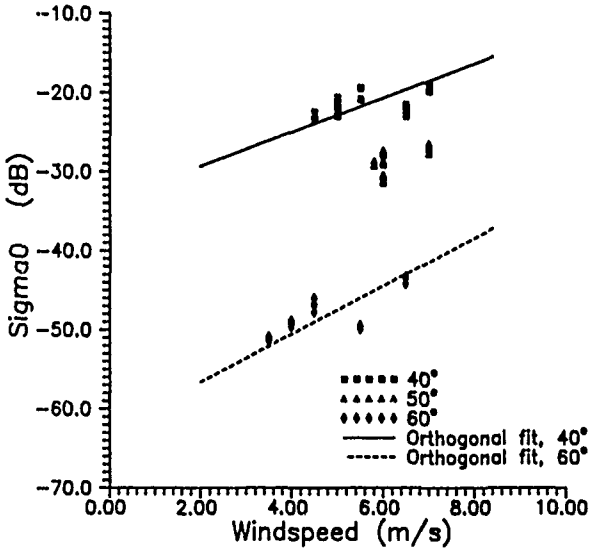


Fig. 6 σ° vs. wind speed
C band, crosswind look

We are currently investigating the effect of other environmental variables on wind speed vs. σ° relationship.

CONCLUSIONS

The pilot experiment verified that we can collect the data required to interpret radar signal dependence on environmental variables under semi-controlled conditions.

REFERENCES

Atakturk, S. S. & K. B. Katsaros, 1989, "K-Gill twin propeller-vane anemometer for measurements of atmospheric turbulence," J. Atmos. & Ocean Tech (in press).

Gogineni, S. P., J. W. Bredow & R. K. Moore, "An inexpensive polarimetric FM radar and polarimetric signatures of artificial sea ice," Proceedings of the 1989 IEEE National Radar Conference, IEEE 89CH2685-6, 29-30 March 1989, Dallas, Texas, pp. 188-191.

ACKNOWLEDGEMENT

This work is supported by grant NAGW-1278 from NASA-Headquarters to the University of Kansas and the University of Washington.

THE DEPENDENCE OF THE MICROWAVE RADAR CROSS SECTION ON OCEAN SURFACE SLOPE STATISTICS: COMPARISON OF MEASUREMENTS AND THEORY

David E. Weissman
Dept. of Engineering
Hofstra University
Hempstead, New York 11550

Tel: (516) 560-5546

This investigation is studying the ability of theoretical RCS models to predict the absolute magnitude of the ocean radar cross section under a wide variety of sea and atmospheric conditions. An extensive amount of experimental data is available from the FASINEX Experiment. This consists of RCS data from the NASA-JPL Ku-band (14.6 GHz) Scatterometer mounted on a C-130 aircraft (10 separate flights were conducted), a wide variety of atmospheric measurements (including stress) and sea conditions. Current theoretical models are being tested with this data, over a range of incidence angles (20 to 60 deg.) and azimuth angles. In situations where the Bragg scattering theory does not agree with the measured radar cross section (maximum magnitude and azimuth angle dependence), revisions to the theory are hypothesized and evaluated. For example, the V-pol theory works well in most cases studied while the H-pol theory is usually too low by about a factor of 2, at incidence angles of 50 and 60 degrees. Other workers have found a similar discrepancy. In response to this situation, the next phase involves studying the effect of surface slope statistics which deviate from the Gaussian statistics that are usually assumed in the theoretical models, and how the theoretical predictions would be affected if they used non-Gaussian surface slope statistics. Non-Gaussian statistics based on actual ocean measurements are being used with theoretical methods for the calculation of the RCS. These methods may deal with either closed form mathematical functions for the PDF or exact moment-method calculations based on a point-by-point spatial series of the surface elevation to calculate the backscattered field. A range of incidence and azimuth angles, and wind speeds for the radar cross section are being studied.

Speckle statistics in airborne SAR imaginary of the ocean surface.

by

Torbjørn Eltoft
University of Tromsø
Postbox 953
N-9001 Tromsø, Norway
Tlf. 47-83-86060
Telefax 47-83-89852

Harald Johnsen
FORUT
Postbox 3063, Guleng
N-9001 Tromsø, Norway
Tlf. 47-83-58622
Telefax 47-83-82420

During the NORCSEX-88 experiment on Haltenbanken the Canadian CV-580 airplane provided C- and X-band SAR imaginary of the ocean surface. Operated in the nadir looking narrow swath mode the footprint covered a ground range area of 41 km off nadir. Flying at an altitude of about 6100 meters this corresponds to a significant variation in incidence angle, integration time and resolution cell dimensions over the swath.

In the present paper the impact of these variations on the speckle statistics in the SAR images have been investigated. Since the plane was flown in a star pattern also the impact of the imaging geometry of the waves on the speckle statistics has been considered.
The objectives have been to relate the speckle properties to the possible scattering mechanisms involved in producing the micro wave echoes from the sea surface.

REMOTELY SENSED SURFACE TEMPERATURE OBSERVATIONS IN HAPEX

T. J. Schmugge

USDA-ARS Hydrology Laboratory
Beltsville, Maryland 20705, USA

J. P. Goutorbe

Centre Nationale de Recherche Meteorologique
42 Avenue Coriolis, Toulouse, France

ABSTRACT

Using data from the Thermal Infrared Multipectral Scanner (TIMS) on board the NASA-130 aircraft, surface temperature maps were created to study the representativeness of the surface flux measurements stations (SAMER). Data from high altitude flights (6 km) were used to obtain 7 by 10 km maps around 9 of the 12 SAMER stations. The histograms of the fields, about 2 or 3 hectares in size, containing the SAMER were compared with the histograms of the entire scene. The SAMER stations were typically located in cultivated fields which at the time (16 June to 2 July) or our aircraft overflights were mostly bare. The histograms of the total scenes ranged from temperatures of 25-30°C for the freely transpiring vegetated surfaces to greater than 40°C for the bare soils and the temperatures for most of the SAMER stations were in the latter category. While the proportions of vegetated and bare surfaces varied from scene to scene they typically were about evenly split. These results have implications for the comparison of airborne and surface flux measurements since the higher temperatures for the fields with the SAMER stations indicate that most of the exchange with the atmosphere is in the form of sensible heat, while for the whole region the exchange may be evenly split between latent and sensible heat fluxes.

Keywords: Brightness temperature, Evapotranspiration, Hydrological budget, Latent heat flux, Sensible heat flux, TIMS.

1. INTRODUCTION

The objective of the HAPEX-MOBILHY Hydrologic Atmospheric Pilot Experiment and Mod'lisation du Bilan Hydrique) program [1] and [2] is to improve the parameterizations of the land surface processes in atmospheric general circulation models (GCM's) which operate at the scale of 10^2 km. Therefore, a measurement program was designed to observe the relevant processes, e.g., the hydrological budget and evapotranspiration flux by a number of different and independent techniques over a 100 x 100 km^2 area. The parameters which control the water budget and ET flux have different time and space scales, e.g., surface hydrology and root zone soil moisture have typical horizontal scales on the order of 100 m, while their time scales are relatively long, i.e., several days. These parameters were observed for over a year at 10 day intervals with a number of point measurements at several locations (~12) over the 100 x 100 km^2. Conversely, surface properties such as air temperature, surface eddy fluxes and other atmospheric boundary layer parameters have rather large spatial scales, 1 to 10 km, but short time scales, about 1 h. The variation of these quantities has been monitored with a regularly space surface network and intervals of 1 hour or less.

The measurements were made over a 100 x 100 km^2 area in southwestern France roughly between Bordeaux and Toulouse and just north of the Pyrenees. The northeastern third of the area is relatively flat pine forest on poorly drained soils. The remaining area is mixed agriculture with rolling hills. The major instrumentation systems included: a) surface networks of the SAMER [3] flux measurement stations and micro-met stations, b) acoustic and radio sounding of the boundary layer, c) airborne measurements of the fluxes and d) satellite and aircraft remote sensing of the surface radiative properties. The aircraft remote sensing was performed by the NASA C-130 which carried sensors operating at microwave, thermal infrared and visible wavelengths. This paper will present results from the Thermal Infrared Multispectral Scanner (TIMS).

2. TIMS

TIMS is a six channel NASA aircraft scanner operating in the thermal infrared (8-12 microns) region of the electromagnetic spectrum. The response functions for the individual channels are shown in Fig. 1 along with the Blackbody radiation for a 300 K body. The 6 channels cover the peak of the emitted radiation curve. Note that channels 5 and 6 roughly correspond to the AVHRR channels 4 and 5.

The instantaneous field of view is 2.5 mrad, while the detector signals are sampled every 2.08 mrad along the scan [4]. The digitized field of view contains 638 pixels and covers 76.6°. Thus, from an altitude of 6000 m the pixel size is 15 m, the cross track separation is 12.5 m at nadir, and the swath width is 9.5 km. For these high altitude flights the scan rate was typically set at 8.7/sec, which at an aircraft speed of 250-260 knots (125 m/sec) provided a separation of 15 m between scans and thus contiguous coverage.

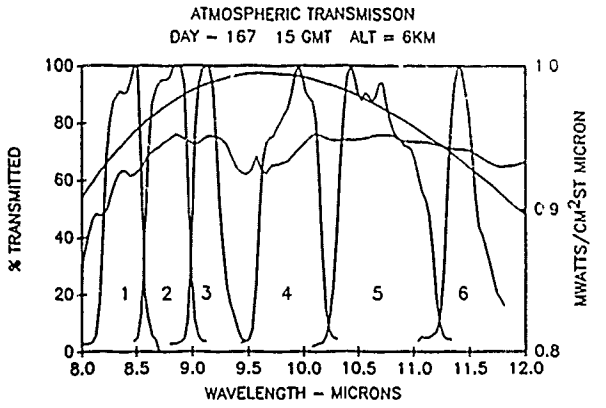


Fig. 1. The TIMS response functions in the 8-12 micron range. The 300 K blackbody emission curve is the smooth curve and the atmospheric transmission up to a 6 km aircraft altitude for day 167 is the third curve.

For calibration the system is equipped with cold and warm reference sources or blackbodies, approximately covering the temperature range of interest. For most of HAPEX the temperature separation between the two references was set at 30°C. All pixels are assigned a digital count value between 0 and 255 (DN). Reference source 1 is scanned at the beginning, and the second at the end of a line. The reference temperatures are known to better than 0.5 K [4].

TIMS responds to the incident radiance (W/m^2 sr cm^{-1}) and not to the temperature directly. The brightness temperature of the observed surface is related to the incident radiance via the Planck equation for blackbody radiation. However, if the observed temperatures are close to the calibration temperatures, say within 10 or 5°C, the error arising from using a linear temperature calibration is less than 1°C and that approximation will be used here [5].

The temperatures given by the sensor represent the detected radiance at the aircraft altitude. In order to convert this result to the actual surface temperature, atmospheric effects must be taken into account. These include the absorption and emission by the atmospheric gases, primarily water vapor for this portion of the spectrum. Since the relationship between radiance and temperature is considered linear, Eq. 1 can be used to correct for these atmospheric effects.

$$T_{ir} = \tau * T_s + (1 - \tau) * T_{av} \qquad (1)$$

T_{ir} is the observed temperature (K), T_s the surface radiation temperature, T_{av} the average air temperature for the atmospheric layer between the surface and the aircraft and τ the transmittance of this layer for a specific channel. Radiosoundings released at the central site within one hour of the pass are used to determine the transmittance and the air temperature. For calculating the average air temperature (T_{av}) the following equation was used:

$$T_{av} = \Sigma T_i * W_i * z_i / \Sigma W_i * z_i \qquad (2)$$

in which T_i, W_i, and z_i are the temperature, water vapor content in kg/m^3, and layer thicknesses of the atmosphere obtained from the radiosoundings.

This equation assumes that the principal absorption is by water vapor and so the temperatures are weighted by the vapor content of the layer. LOWTRAN-6, an atmospheric path radiance model developed by the Air Force Geophysics Laboratory [6], is used to calculate the transmittance for the different channels. An example of the resulting transmittance for the 6 km altitude is given in Fig. 1. The product of the transmittance and the response function for the individual channels yields the transmission factor for each channel. On June 16 the values of τ ranged from 0.60 for channel 1 to 0.74 for channel 5. The low value for band 1 indicates the increased water vapor absorption at the 8 micron edge of the window. These results are for the 6 km altitude with T_{av} = 284.1 K and an integrated water vapor content of 2.5 gm/cm^2.

Using these TIMS data, surface temperature maps were created to study the representativeness of the surface flux measurements stations (SAMER) for possible comparison with the remote flux measurements made from the King-Air. Data from high altitude flights (6 km) for band 5 were used to obtain a 9.4 by 15 km map around the central site and forest tower. The histograms of the surface temperatures for several representative surface conditions were obtained and are presented in Fig. 2 for 16 June 1986 (day 167).

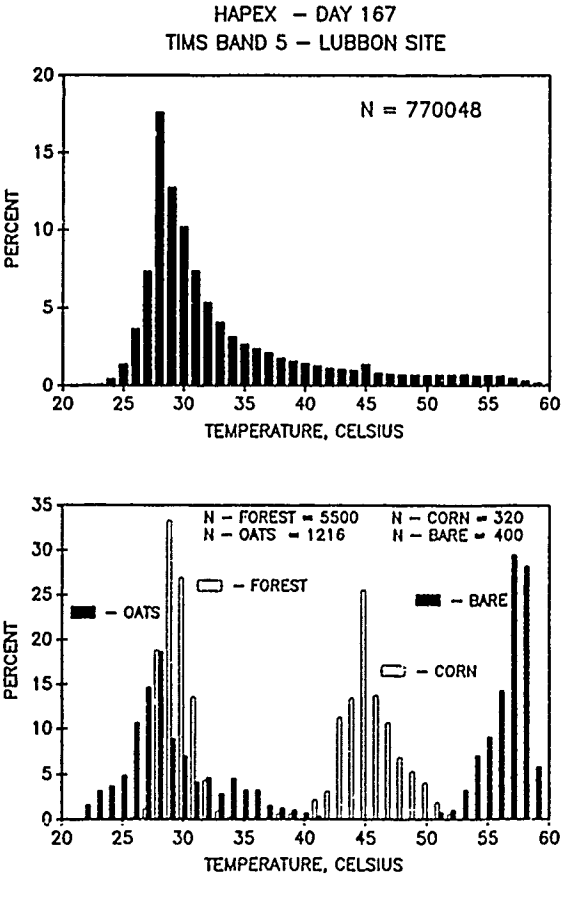


Fig. 2. Histograms of aircraft observations of surface temperature from an altitude of 6 km for TIMS band 5 on day 167. The upper plot is for all the pixels in a 9.4 by 15 km scene and the lower one is for specific surfaces within that scene.

The upper histogram gives the distribution for the entire scene showing a temperature range from 20-60°C with the majority being less than 35°C. The air temperature near the ground at the time of the flight was 29°C. The lower portion of the figure gives the histogram of several individual fields to indicate the sources of this large range in temperature. They ranged from a mature oat field to a bare field with average temperatures of 29.4 and 6.9°C, respectively. The corn field had plants which were 50-70 cm high with a leaf area index (LAI) of 1.8. Because of sufficient root zone soil moisture, the corn was freely transpiring but there was a considerable amount of bare soil visible to the radiometer accounting for its high temperature, T_{ir} = 45.8°C, which is not indicative of its evaporative flux. The temperature for the corn field changed as the crop matures as seen in Fig. 3 which presents the results for 27 June. The overall temperature distribution is about the same but the temperature of the corn field is now cooler than the oat field which has mostly passed into senescence. The average values for these fields are summarized in Table 1.

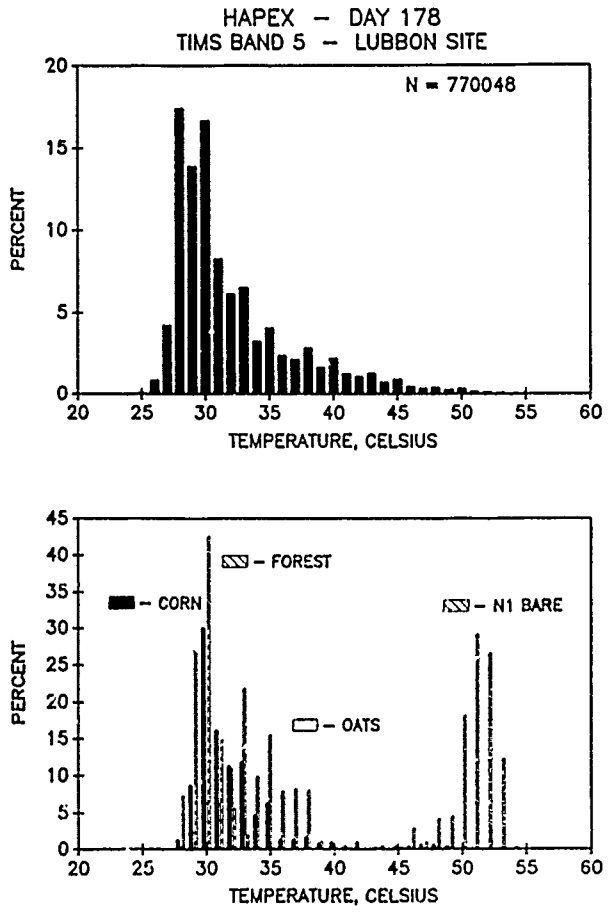


Fig. 3. Histograms of aircraft observations of surface temperature from an altitude of 6 km for TIMS band 5 on day 178. The upper plot is for all the pixels in a 9.4 by 15 km scene and the lower one is for specific surfaces within that scene.

Table 1. Surface temperature statistics in Celsius

Field	Day 167 Mean	Day 167 S.D.	Day 178 Mean	Day 178 S.D.
Bare	56.9	1.6	51.3	1.6
Corn	45.9	2.4	37.6	3.5
Oats	29.4	3.9	34.3	2.4
Forest	30.0	1.2	30.3	1.1
All	33.4	7.4	32.4	4.8

The SAMER flux measurement stations were typically located in cultivated fields which at the time (29 May to 2 July) of our aircraft overflights were mostly bare. The histograms of the total scenes ranged from temperatures of 25-30°C for the freely transpiring vegetated surfaces to greater than 40°C for fields which were bare or partially vegetated. The temperatures for most of the SAMER stations were in this range. While the proportions of vegetated and bare surfaces varied from scene to scene they typically were about evenly split. These results have implications for the comparison of airborne and surface flux measurements. For example, in this situation the majority of the area had a surface temperature typical of the forest, the flux measured by the King-Air would be close to that measured by the forest tower.

3. REFERENCES

1. Andre, J. C., J. P. Goutorbe, & A. Perrier, "HAPEX-MOBILHY: A hydrologic atmospheric experiment for the study of water budget and evaporation flux at the climatic scale", Bull. Amer. Met. Soc., 67, 138-144, 1986.

2. Andre, J. C. & 30 others, "Evaporation over land surfaces: First results from HAPEX-MOBILHY special observing period," Annales Geophysicae, 6, 477-492, 1988.

3. Riou, C., "Une expression analytique du flux de chaleur sensible en conditions sur adiabatiques a partir de mesures du vent et de la temperature a deux niveaux," J. Rech. Atmos., 16, 15-22, 1982.

4. Palluconi, F., & G. R. Meeks, "Thermal infrared multispectral scanner (TIMS): An investigator's guide to TIMS data," NASA, JPL, Pasadena, June 1985.

5. Schmugge, T. & L. Janssen, "Aircraft remote sensing in HAPEX," Proc. 4th Int'l Coll. on Spectral Signatures in Remote Sensing, Aussois, France, 18-22 January 1988 (ESA SP-287), 463-467, 1988.

6. Kniezys, F. X., E. P. Shettle, W. O. Gallery, J. H. Chetwynd, L. W. Abreu, J.E.A. Selby, S. A. Clough & R. W. Fenn, "Atmospheric Transmittance/Radiance Computer Code LOWTRAN 6," Air Force Geophysics Laboratory Report No. AFGL-TR-83-0187, Hanscom AFB, Massachusetts, 01731, 1983.

REMOTELY SENSED MEASUREMENTS OF SURFACE TEMPERATURES OVER SPARSE CANOPIES

C. S. T. Daughtry
USDA-ARS Remote Sensing Research Laboratory
Beltsville, MD 20705 U.S.A.
301-344-3015
FAX: 301-344-1048

W. P. Kustas
USDA-ARS Hydrology Laboratory
Beltsville, MD 20705 U.S.A.
301-344-2498

M. S. Moran and R. D. Jackson
USDA-ARS Water Conservation Laboratory
Phoenix, AZ 85040 U.S.A.
602-261-4356

Regional scale evaluation of the energy balance requires the use of remotely sensed data to provide timely information over the area of interest. Surface temperature is an important variable in energy calculations. However evaluating surface temperature with remote sensing is difficult, particularly when the surface is partially vegetated. The temperature measured by a radiometer is affected by the proportions of vegetation and soil and represents a composite temperature. When the soil is not completely covered by vegetation, its temperature may be considerably warmer than transpiring vegetation. Thus the composite temperature may not be representative of the evaporating surface and may cause significant errors in model estimates of evapotranspiration.

The objectives of this study were to (1) compare ground-based estimates of the composite surface temperature from measurements sunlit and shaded soil and vegetation to measurements from an aircraft, (2) develop a method of estimating soil heat flux an net radiation from remotely sensed data and (3) examine the spatial distribution of surface temperature, soil heat flux, and net radiation over irrigated cotton fields.

Measurements of soil heat flux, soil temperatures, and net radiation of large bare soil, alfalfa, and cotton fields were recorded continuously on selected dates in 1987 and 1988 at the Maricopa Agricultural Center near Phoenix, AZ. Spectral reflectance factors were acquired at selected times with a radiometer mounted in a airplane flying at 150 m and with a radiometer mounted on backpack-type frame at 1.7 m. Radiant temperatures of vegetation and soil were measured on east- and west-facing sides of the rows with a hand-held infrared thermometer.

Composite surface temperatures were modeled and compared to temperatures measured from the aircraft along transects of the fields. Modeled and measured surface temperatures agreed well. The ratio of soil heat flux and net radiation were estimated with spectral vegetation indices. With this technique, available energy (net radiation minus soil heat flux) could be estimated within about 3%. If this technique can be generalized for use over different vegetative surfaces, the result could be more accurate estimates of surface energy balance than is possible with meteorological measurements alone.

ON INTERCONNECTION BETWEEN MICROWAVE AND INFRARED RADIATION FIELDS AND NATURAL OBJECT CONDITION

E.A.Reutov

Institute of Radio Engineering & Electronics USSR Academy of Sciences, GSP-3, 103907, Marx avenue,18, Moscow K-9, U.S.S.R.

The management of natural objects over large areas need the near-real-time and reliable information about its condition. Apparently, this information may be collected over large areas only by means of remote sensors.

This problem may be decided using two ways. First one is usual and it relies the determination of separate soil and canopy parameters, such as moisture,temperature,salinity,biomass etc. using remote sensors. It's believed that potential users of remotely sensed data (for one in agriculture) using above mentioned separate parameters will estimate the total condition of natural objects and make a correspondent technical decision. This way is however very complex and not quite reliable as it needs the multichannel remote sensors and solution of multiparametric inverse problems.

The second possible way is to delineate the condition of these objects using the parameters which could be immediatly determined with remote sensors (for example brightness temperature at microwave and infrared ranges). This approache permits to avoid the amplification of model errors which considerably grow under solution of inverse problems.

Let consider the agricultural geosystem to simplify the problem. In such case energy balanc equation of this geosystem will be

$$R = H + L E_t - G$$

where R is the radiation balanc, G is the energy consumption for the soil heating, H is the energy consumption for the air heating, L is the latent evaporation heat, E_t is the evapotranspiration.

The optimum condition for crop development and to obtain the top yield is

$$R - L E_t = 0 \quad \text{or} \quad R / L E_t = 1$$

when all energy is consumed for evapotranspiration. In such case the difference between current condition and optimum one may be estimated as

$$\Delta J = R / L E_t - 1 = (H - G) / L E_t \text{ , or in other units}$$

$$\Delta J = h_s \frac{c_s}{L} \cdot \frac{d T_s}{dW + \Delta W} - h_a \cdot \frac{c_a}{L} \cdot \frac{d T_a}{dW + \Delta W}$$

where dW, dT_a, dT_s are the changes of soil moisture,air temperature and soil temperature, respectively, per time dt, ΔW is the watering or precipitation between two measurements, c_s, c_a are the specific heats of soil and air, h_s, h_a are the thicknesses of soil and air layers taking part in the heat exchange.

The dW value may be estimated using microwave radiometry, the dT_s -with infrared radiometry, dT_a -using meteorological data. Then, we can obtain for soil/canopy system with error about 10% (if G=0, as the G is about 3-4% of R)

$$\Delta J \simeq \frac{h_a}{L} \cdot \frac{c_a \cdot b_m \cdot b_{if} \cdot \beta \cdot dT_a}{b_{if} \cdot dT_m^b + (1-\beta)\,(T_{if}^b - A_{if}) - \Delta W \cdot b_m \cdot b_{if} \cdot \beta}$$

where dT_m^b, dT_{if}^b are the changes os soil/canopy system brightness temperature at microwave and infrared ranges, respectively, β is the transfer function of canopy, b_m , b_{if} , A_{if} are the coefficients depending on the wavelength at microwave and infrared ranges.

The ΔJ value permits to estimate the difference between the current condition of agricultural geosystem and its optimum (from the point of view of the maximum yield) condition. Also, the linkage of ΔJ value and the yield losses is quite strong.

AVHRR-derived cloudiness and surface temperature patterns over the Los Angeles area and their relationships to land use

Bénédicte DOUSSET

Graduate School of Architecture and Urban Planning
University of California Los Angeles[1]

Abstract. Eighty-four co-registered NOAA AVHRR satellite images of the Los Angeles basin were analyzed for the months of August 1984 and 1985, with the objectives of quantifying the differences between surface temperatures and co-located air temperatures, and of studying micro-climates using surface temperature and cloudiness statistics. Assuming a surface emissivity of 0.97 and ignoring atmospheric absorption, satellite temperatures were in good agreement with air temperatures at night, with a standard deviation of 1.6°C. During the day, they were biased 1.1°C warm in the morning and the late afternoon, and 5.4°C warm at 15:30 local time, in qualitative agreement with models of the planetary boundary layer. Pixels were classified as cloudless, covered with low (warm) or high (cold) clouds. At 03:40, low clouds occur over the ocean and cover the basin up to the edge of the San Gabriel mountains. They have dissipated by 14:40. Around 14:40, high clouds are found over the mountains and the desert, but not over the basin. Surface temperature is correlated with altitude. At 15:10, the 1800-m high San Gabriel mountains are ~17°C colder than the basin. Valleys are seen as warmer features crossing the mountain range. At 03:40, the south flanks of the mountains are ~3°C warmer than the basin. Downtown Los Angeles corresponds to a heat-island which follows contours of building density. Despite the 1-km resolution of the AVHRR, narrow features such as low thermal inertia freeways or high thermal inertia irrigated fields in the desert are detected (satellite infrared images, micro-climates).

1. Introduction

Existing meteorological observations, recorded over a sparse network of stations (often at suburban airports), do not have details at micro-climatic scales. Surface temperature and cloudiness data are most efficiently collected today by infrared scanners on board polar meteorological satellites, which provide surface temperature over the entire earth twice daily at a resolution of 1 km. The frequency and global extent of their coverage create new capabilities and provide accurate data directly applicable to the analysis of micro-climates.

In this paper, we analyze a series of co-registered satellite images of the Los Angeles basin for the months of August 1984 and 1985, with the following objectives: (1) understand and quantify the differences between satellite-derived surface temperatures and co-located air temperatures; 2) demonstrate that satellite-derived temperature and cloudiness statistics can yield useful informations on urban micro-climates.

The data and the processing steps are described in section 2. Air and satellite-derived temperatures are compared in section 3 (some results of this section have previously been discussed by Dousset, 1986). Statistics of cloudiness and surface temperature are presented in sections 4 and 5. Applications and future improvements of the method are discussed in the conclusion.

2. Data, processing and corrections

The images are from the Advanced Very High Resolution Radiometer (AVHRR) on board the NOAA-6, -7, -9 and -10 satellites. The AVHRR is a scanning radiometer with five spectral bands, centered at 0.62 μm (band 1), 0.91 μm (band 2), 3.74 μm (band 3), 10.8 μm (band 4) and 12.0 μm (band 5, NOAA-7 and -9 only). The measurement noise is ~0.1°C in bands 4 and 5, but can reach 1°C in band 3. The swath width on the earth is 3000 km and the ground resolution (pixel size) varies from 1.1×1.1 km^2 below the satellite to 1.5×4.0 km^2 at the swath edges. The satellites are launched at an altitude of about 800 km into near-polar sun-synchronous orbits and pass in view of any point on the earth twice daily, around 07:00 and 19:00 local time for NOAA-6 and 8, and 03:00 and 15:00 for NOAA-7 and 9.

Eighty-four images were processed at the Scripps Satellite Oceanography Facility. Their time distribution is shown in Fig. 1. They were geometrically corrected for the earth rotation and curvature and interactively registered to a common Mercator grid, the coverage of which is shown in Fig. 2. Small lakes such as Lake Perris (at [114,86] km on Fig. 7, 8 and 9) were used as reference points to obtain absolute positions accurate to ~2 pixels rms.

[1]corresponding address: Department of Meteorology, University of Hawaii at Manoa, Honolulu HI 96822.

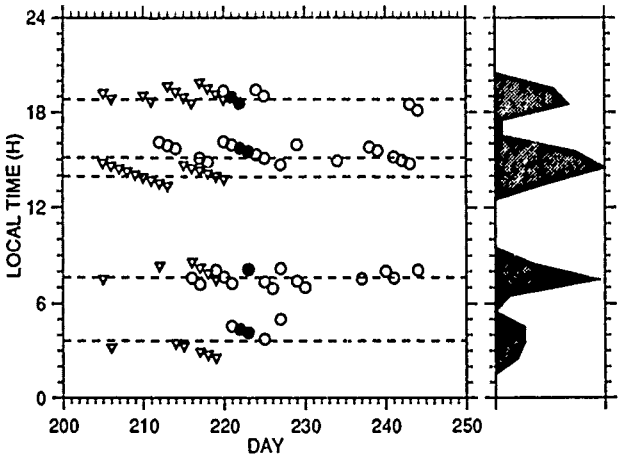


Fig. 1. Time distribution of the images. X-axis: day of the year, Y-axis: local solar time. The 1984 images are shown by circles O and the 1985 images by triangles ∇. The 1984 images for which in situ data are available are shown by filled circles •. The histogram of the distribution is shown in the right panel.

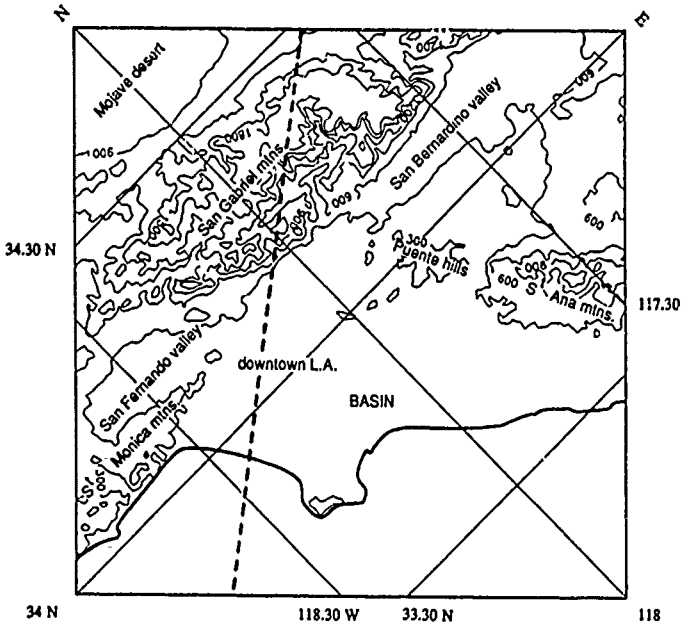


Fig. 2. Topography over the area of the images (128 km on a side). The contour lines are labeled in m above sea level. North is at the upper left of the frame. The section shown in Fig. 10 is indicated by a dashed line. Copying this figure on a transparent overlay now will facilitate the interpretation of Fig. 5 to 9.

The radiant energy in each thermal infrared band was then converted to brightness temperatures using Planck's law. In the 8 μm to 14 μm band, most natural and man-made materials found in urban areas have an emissivity smaller than the ocean (0.99) (Buetner and Kern, 1965): dry sand (0.914), wet sand (0.938), vegetation (0.975), glass (0.865), asphalt (0.958), plywood (0.962), concrete (0.966). A residential block in Los Angeles, consisting mainly of wood shingle roofs, asphalt, concrete, vegetation and pools, has an average emissivity of 0.965 to 0.97. The brightness temperature T_B is thus lower than the true surface temperature T, and, in bands 4 and 5, can be approximated by a formula derived from Planck's law:

$$T_B = T - (1-\varepsilon)\lambda T^2/C$$

where ε is the emissivity, λ is the wavelength band in m, C = $1.44 \cdot 10^{-2}$ m°K, and T and T_B are in °K. For T = 20°C, the correction is 0.6°C for each % of emissivity below 100%. Lacking detailed information on surface emissivities, we applied this correction to all the images, assuming ε=0.97.

A surface that is not a perfect black body also reflects solar radiation (Kirchoff's law), resulting in a brightness temperature error. For a reflectance α=1−ε=0.03 in the thermal infrared, this error can reach 10°C at 3.74 μm, but is negligible in bands 4 and 5 around 11 μm.

The infrared radiation emitted by the surface is partially absorbed in the atmosphere, mainly by water vapor. This absorption lowers the brightness temperatures measured by the satellite. The spectral bands of the AVHRR correspond to "windows" at which the radiation is least absorbed. In tropical regions, where the moisture content of the atmosphere reaches 60 kg/m², the errors due to absorption are still about 10°C in band 5, 7°C in band 4 and 3°C in band 3. Surface temperatures are then estimated using empirical multispectral corrections (McClain *et al.*, 1985). Since moisture content in dry subtropical regions such as Southern California is typically ~10 kg/m², corresponding to a brightness temperature error less than 1°C for band 4, this band was chosen and the water vapor correction was neglected in this preliminary study.

3. Comparison with air temperatures

Seven images correspond to an intensive experiment during which hourly data from 44 meteorological stations and 4-hourly data from 11 sounding stations were recorded (Project BASIN, August 8 to 10, 1984; Wakimoto and Wurtele, 1984). Surface temperature co-located with the stations were extracted from the images. Cloudy pixels were rejected using the radiance difference between 3.7 μm and 10.8 μm for night images, and the near-infrared reflectance for daytime images. Stations at an altitude higher than 400 m such as Mt. Wilson were also rejected to avoid the problem of relating pixel-average temperatures with point data taken in rough terrain.

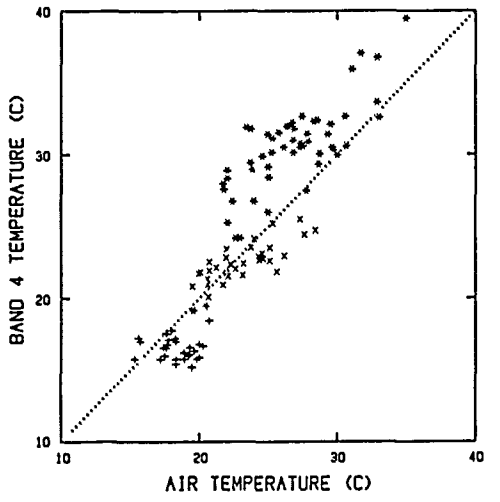


Fig. 3. Scatter plot of satellite temperature versus air temperature for NOAA-6 at 08.15 and 19:00 (+) and NOAA-7 at 04:00 (×) and 16:00 (*).

Table 1. Mean and standard deviation σ of the difference between satellite and air temperatures at different times of the day.

time	symbol	mean	σ	# points	# images
04:30	+	0.0°C	1.6°C	29	3
07:00,19:00	×	1.1°C	1.5°C	34	2
15:30	*	5.4°C	2.3°C	54	2

Fig. 3 is a scatter plot of satellite temperature versus surface air temperature for all the satellite images; the corresponding statistics are shown in Table 1. At night, the satellite temperatures are in good agreement with air temperatures; the standard deviation of the difference is 1.6°C (to provide a reference, comparisons between corrected satellite temperatures and ocean surface temperatures have a typical scatter of ~0.6°C; Bernstein, 1982). In the afternoon, satellite temperatures are about 5.4°C warmer than the air. The standard deviation is 2.3°C.

The standard deviations are remarkably small given the fundamental difference between pixel-averaged surface temperatures and point measurements of air temperature. The diurnal cycle of the mean difference is in qualitative agreement with simple planetary boundary layer models, which predict the largest difference in the early afternoon when the incoming solar radiation is most intense. The larger standard deviation at 15.30 reflects more turbulent conditions in the afternoon, as evidenced by peak-to-peak fluctuations of 1.5°C over less than 10 minutes observed at a meteorological station in downtown L.A.

4. Cloud statistics

During the day, cloudy pixels were flagged when band 2 exceeded a threshold based on the histograms of cloud-free images. At night, cloudy pixels were flagged when the difference between band 3 and band 4 exceeded a threshold based on the histograms of cloud-free images (the emissivity of clouds is 0.70 at 3.7 μm versus 0.99 at 11 μm). Examples of histograms of cloud-free and cloud-covered images are shown in Fig. 4. The brightness temperatures in band 4 was used to differentiate between low (stratus) clouds and high (convective) clouds. Maps of cloud probability were then constructed for each set of images at the average times of satellite passage shown in Fig. 1.

At ~03:40 average local time (Fig. 5), stratus clouds cover the ocean and the coast in 80% of the images and in 40% to 60% penetrate over the low-altitude Los Angeles basin to the edge of the San Gabriel mountains and to San Bernardino valley. These low clouds do not occur over the mountains and in the Mojave desert. Similar patterns are seen at ~07:40 (not shown), when stratus clouds cover the ocean and the coast in 30 to 40% of the images and in 10 to 20% extend over the basin. At ~15:10 (Fig. 6), low clouds have dissipated and high clouds are found in 15% to 20% of the images over the mountains and the desert, but in less than 3% over the basin. At ~18:50 (not shown), low clouds have formed again over the ocean in 10 to 25% of the images and in 5 to 15% extend over the basin.

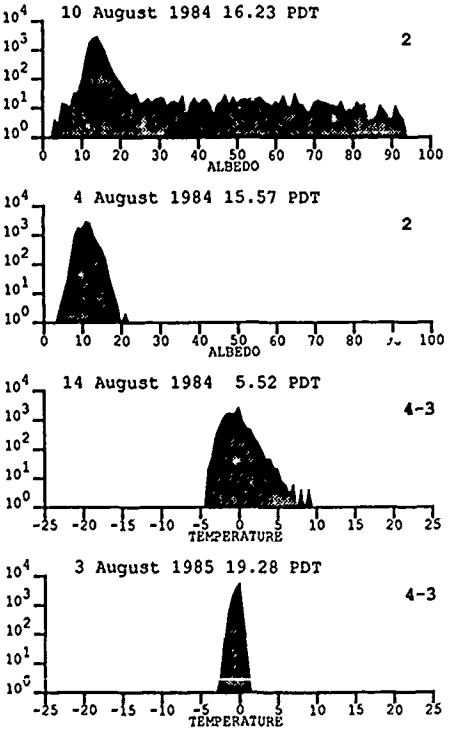


Fig. 4. Histograms used for cloud classification. (a) albedo of a partially cloudy daytime image, (b) albedo of a cloudless daytime image, (c) band 4 - band 3 of a partially cloudy night image, (d) band 4 - band 3 of a cloudless night image.

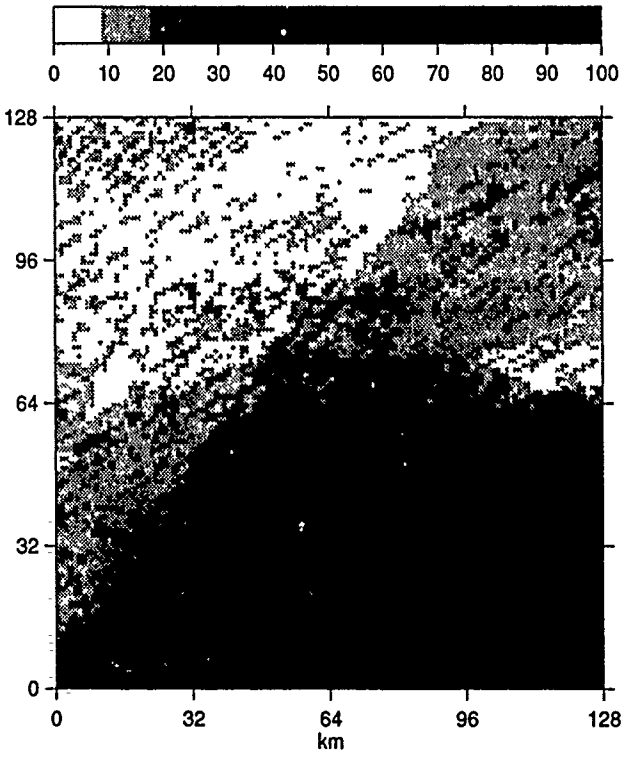


Fig. 5. Probability of cloud occurrence at ~03:40 local time.

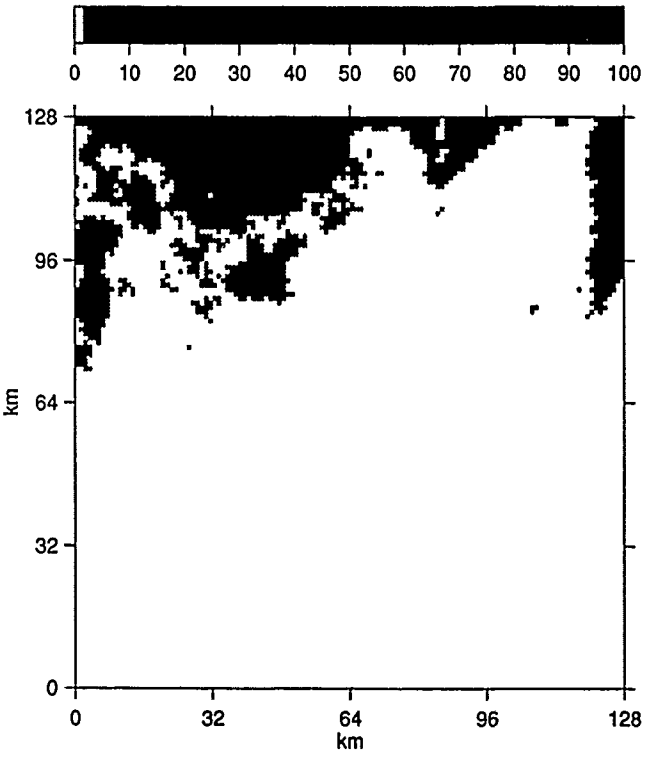


Fig. 6. Probability of high cloud (T < 0°C) occurrence at ~14:40 local time.

These cloud patterns corresponds to the mesoscale climate of the region. A deck of marine stratus form over the ocean when warm moist marine air masses are advected over cold coastal water, and fills the entire low-altitude basin. The maximum elevation reached by the stratus deck in the 03:40 and 07:40 images is about 600 m, consistent with the reported 300 to 600-m thickness of the marine layer (Elford, 1970).

5. Temperature statistics

Maps of average surface temperature were constructed for each set of images at the different times of satellite passage (Fig. 1). Because the rate of change of surface temperature is large in the afternoon, separate averages were computed for 1984 (~15:10 local time) and 1985 (~13.55), but the two years were combined for 03.40, 07.40 and 18:50. Pixels flagged as cloud in the previous section were omitted from the average, and average pixels with a cloud probability larger than 75% were discarted (they appear as a white mask over the ocean in Fig. 8).

At 15:10 (Fig. 7), the ~1800-m high San Gabriel mountains are 12°C colder than the low-altitude basin. Valleys are seen as warmer features crossing the mountain range (e.g. near [64,96] km). Coastal relief such as the Santa Monica and Santa Ana mountains are also colder than the basin. Surface temperatures exceeds 33°C in San Fernando valley, San Bernardino valley and the Mojave desert.

The San Andreas fault can be seen as a cold anomaly crossing the desert near [0,96] km. Other cold spots in the desert correspond to irrigated fields, which were identified in a multispectral Landsat image. Higher temperatures are found downtown L.A. near [45,45] km, and in the industrial area of Long Beach near [64,32] km.

At ~03:40 (Fig. 8), the south flanks of the Santa Ana and San Gabriel mountains are in general ~3°C *warmer* than San Fernando and San Bernardino valleys (e.g. near [50,75] km, or [120,70] km), despite their higher altitude. The causes of this effect are not understood. Warmer temperatures are also found in downtown L.A. and around Long Beach, presumably reflecting nocturnal heat islands, the contour of which closely follows contours of building density obtained from a Landsat image. The San Andreas fault appears as a warm anomaly. Small details such Interstate 10 running parallel to the San Gabriel mountains (from [69,78] km to [105,114] km) are distinctly warmer than the background.

To further illustrate these temperature variations, a profile of average temperatures from LAX Airport to the Mojave desert through downtown L.A. is shown in Fig. 9. The correlation with topography is clear. Heat islands in downtown L.A. (at 20 km from the coast) and South Pasadena (at 40 km) are noticeable. The transition from ocean to land is spread over ~2 km, as should be expected given the ~2-km rms geolocation error.

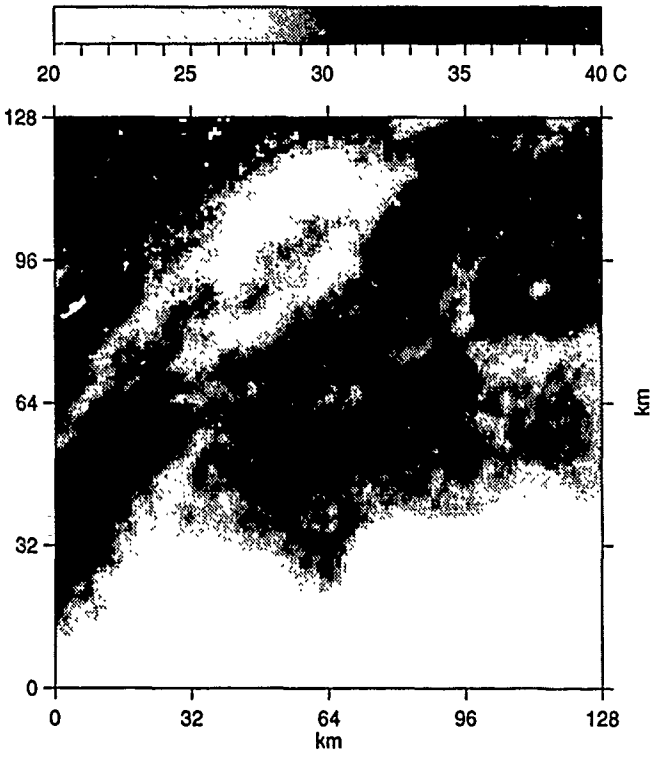


Fig. 7. Average temperature of cloudless pixels from 18 images measured around 15:10 in August 1984.

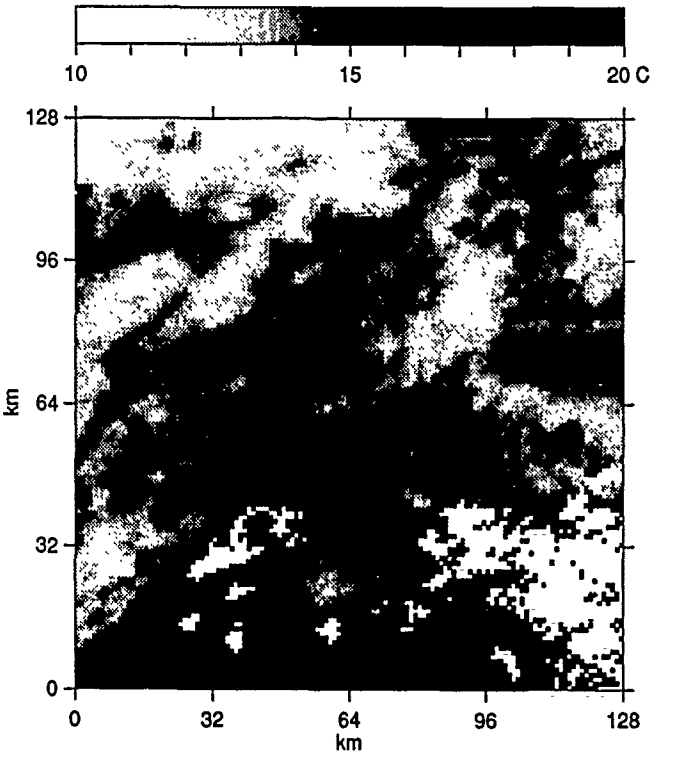


Fig. 8. Average temperature of cloudless pixels from 11 images measured around 03:40 in August 1984 and 1985.

The difference between the 15:10 and 03:40 average temperature maps is shown in Fig. 10. The amplitude is less than 5°C over the ocean and Perris lake, ~10°C over the mountains, ~15°C over the basin, and reaches 25° in the west of San Fernando Valley, the east of San Bernardino valley, and in the Mojave desert.

6. Conclusion

We have demonstrated that statistics of series of accurately co-registered satellite infrared images yield a wealth of informations on micro-climatic processes at a scale of a few kilometers. Although for this preliminary study we chose a very simplified treatment of atmospheric absorption, surface emissivity and cloud detection, the qualitative aspects of the temperature and cloudiness maps, such as amplitude of the diurnal cycle, positions of urban heat islands and distribution of cloud types should not be affected.

Future work will improve the treatment along the following lines:

(i) the error due to the absorption of infrared radiation by water vapor in the atmosphere will be computed directly from radio-sonde data using an atmospheric transmittance model such as LOWTRAN (Kneizys *et al.*, 1980); this will provide an estimate of the emissivity difference between bands 4 and 5;

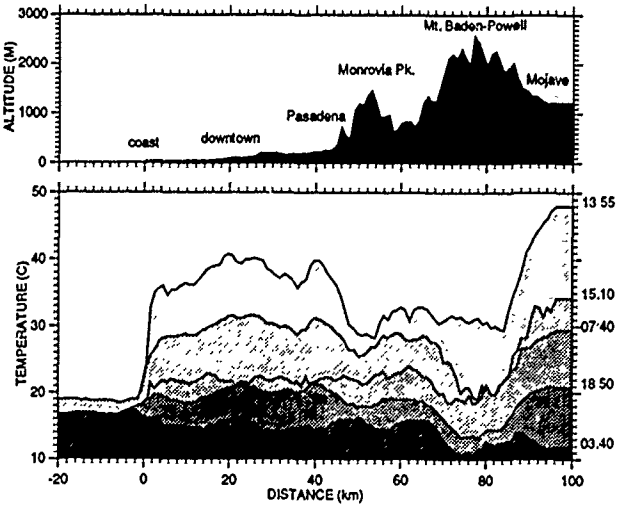


Fig. 9. Bottom: average profiles of surface temperature across the Los Angeles basin from 33°45N 118°41W to 34°27N 117°36W, at ~03:40 (11 images), ~07:40 (21), ~13:55 (15), ~15:10 (18) and ~18:50 (19). Top: corresponding digital topography. Salient geographic features are indicated.

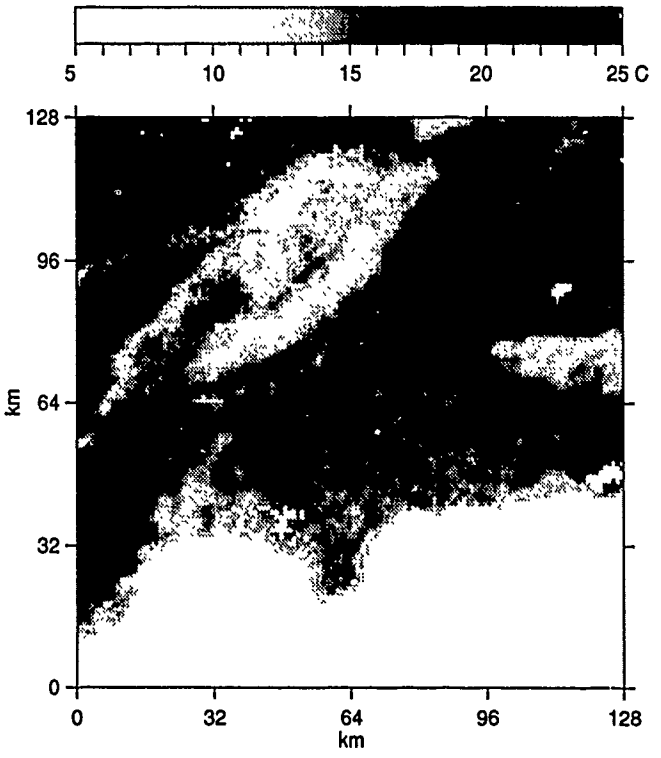


Fig. 10. Difference between the average temperatures shown in Fig. 7 and 8.

(ii) the cloud-detection algorithm will be improved by comparing the albedo of each pixel with the average clear albedo, allowing a better differentiation between white sands and thin stratus or sub-pixel cumulus clouds, which may be missed by the threshold technique used here;

(iii) the observed difference between air and surface temperature will be compared with the predictions of a planetary boundary layer model forced by the meteorological data available from project BASIN.

Microclimatic statistics such as those presented here have direct applications to urban planning and building design. First, a better understanding of urban heat island effects may help control them when planning new neighborhoods, by selecting the albedo of roofs, the clustering of buildings and the distribution of green areas. Second, these statistics provide high resolution climatic data needed when designing energy-efficient constructions.

References

Bernstein, R.L., "Sea surface temperature estimation using the NOAA-6 satellite AVHRR," *J. Geoph. Res.*, vol. 87, pp. 9455-65 (1982).

Buettner, K.J.K. and C.D. Kern, "The determination of infrared emissivities of terrestrial surfaces," *J. Geoph. Res.*, vol. 70, pp. 1329-37 (1965).

Dousset, B., "Some applications of satellite thermal infrared images to urban climatology," in *Proc. 10th Congress of the International Council for Building Research*, pp. 3150-3159, National Bureau of Standards, Gaithersburg (1986).

Elford, C.R., "Climates Of The States: California," *Climatography of the United States*, vol. 60-4, pp. 86-94 (1959).

Kneizys, F.X., E.P. Shettle, W.O. Gallery, J.H. Chetwynd, L.W. Abreu, J.E. Selby, R.W. Fenn, and R.A. McClatchey, "Atmospheric transmittance radiance: computer code Lowtran-5," TR-80-0067, p. 233, Air Force Geophysics Laboratory, Cambridge, 1980.

McClain, E.P., W.G. Pichel, and C.C. Walton, "Comparative performance of AVHRR-based MCSST," *J. Geoph. Res.*, vol. 90, pp. 11587-601 (1985).

Wakimoto, R.M. and M.G. Wurtele, "Project BASIN," *Bull. Am. Met. Soc.*, vol. 65, pp. 1210-11 (1984).

Acknowledgements. The authors thanks B. Givoni, M. Wurtele, and P. Flament for helpful discussions during the course of this research. R. Bernstein supplied some of the images. D. Kelley provided an outstanding plotting package. This work was funded by the California Space Institute through grant CS-78-86. Additional support was provided by the International Federation of University Women in Geneva and by a B.A. Berkus Systems Building and Housing Research fellowship.

Estimating Evapotranspiration Over Large Areas

John C. Price

Agricultural Research Service, USDA, Beltsville, MD 20705

Abstract

Several methods may be used to estimate regional scale evapotranspiration (ET) from satellite thermal measurements. These procedures assume knowledge of surface roughness, albedo, vegetation properties, etc. In many areas these parameters are not accurately known due to rapidity of change of vegetation, lack of a ground data base, and low spatial resolution of satellite data which results in composite picture elements. In this paper, ET is estimated by relating surface temperature to vegetation index. The method requires estimates of ET for completely vegetated area and non-vegetated areas. A property is identified for discrimination of cirrus clouds from areas of varying soil moisture. A data bank of surface properties is needed for description of large scale surface processes. (Keywords-Evapotranspiration, AVHRR, crops)

Introduction

Several procedures exist using satellite data to estimate evapotranspiration (ET) over large areas. Numerical models (Carlson, et al. 1981, Taconet et. al., 1986) and field experiments have produced accurate methods for describing various surfaces. These methods may be applied over the entire earth, yielding data for numerical simulation models. Although the land surface is non-uniform at the 50-200 km scale of such models, a method for aggregating results can yield parameterized values (Price, 1982a). However present methods of analysis assume that values are known for surface roughness, albedo, and emissivity, plus canopy parameters for vegetated areas. This assumption is unsuitable when surface conditions are not in a natural state, but change according to agricultural and industrial practices, clearing and burning, reforestation, etc. Also the Advanced Very High Resolution Radiometer (AVHRR) lacks spatial resolution to assure uniform surface conditions within a field of view. This is not a problem, if detailed information exists regarding surface properties for each area. Analysis at small scales, followed by aggregation of results would provide a solution.

This approach is not applicable, as Landsat and SPOT do not provide coverage frequency. Figures 1 and 2, showing surface temperature and vegetation index, illustrate typical variability. Considering cloud cover, the interval between Landsat observations is too long. Aggregation of Landsat data would result in multiple surface types for most 1 km AVHRR data points.

The problem of describing ET accurately for each data point without knowledge of the surface parameters is insoluble. We address this problem, relying on the variability typical of most areas. Section 2 discusses elements of the surface energy and moisture balance. A vegetation index can discriminate vegetation from soils, which have different moisture transfer properties. Section 3 addresses variability within each AVHRR measurement and the procedure for inferring regional scale properties. The need for a statistical data base of surface properties is discussed in section 4.

2. Surface Energy and Moisture Balance

Fluxes of energy, momentum, and moisture at the earth's surface are affected by meteorological conditions and by the physical state of the surface. Some variables may be inferred from remote sensing.

A. Surface Temperature and Emitted Radiation

Surface temperature affects the partitioning of energy at the earth's surface through its effect on latent and sensible heat fluxes, sensible heat flux into the ground, and radiation to the atmosphere. Land surface temperatures may be estimated by (Price, 1984)

$$T_{BB} = T_4 + 3.3 \cdot (T_4 - T_5) \qquad (1)$$

Here T_{BB} is the radiance temperature and T_4 and T_5 are temperatures derived from AVHRR channels 4 and 5. This formula must be adjusted for emissivity. We correct a sign error in the previous evaluation

$$T_s = [\, T_4 + 3.3 \cdot (T_4 - T_5) \,] \cdot (\, 5.5 - \epsilon_4)/4.5 - 0.75 \cdot T_4 \cdot (\, \epsilon_4 - \epsilon_5) \qquad (2)$$

where ϵ_4 and ϵ_5 are emissivities and T_s is the surface temperature. However information concerning spectral emissivity is generally unavailable.

B. Albedo

Albedo may be estimated from AVHRR channels 1 and 2. Complicating factors exist, e.g. these spectral intervals do not represent the full solar spectrum, angular variations occur in the satellite data, and the effect of the earth's atmosphere. Gutman (1988)

Figure 1. Vegetation index derived from Landsat data, with higher values bright, lower values dark. This area corresponds to 170 AVHRR measurements.

Figure 2 Temperature values corresponding to figure 1, with high values bright, cooler areas dark.

C. Vegetation Index

The presence or absence of vegetation affects surface energy and moisture fluxes. Vegetation can extract water from significant depths, while bare soil is to a first approximation either wet or dry. Vegetation also increases surface roughness, while bare soil rapidly erodes to a very smooth condition.

There is no standard formula for deriving vegetation cover from the satellite data. From a number of formulas we select the normalized difference vegetation index V (Perry and Lautenschlager, 1984, Gutman, 1987)

$$V = \frac{(a_2 - a_1)}{(a_2 + a_1)} \qquad (3)$$

where a_1 and a_2 are radiance values measured by AVHRR channels 1 and 2. In AVHRR data V values generally range from 0.05 to approximately .65. Note that V is an index, and does relate directly to leaf area index or other measurable quantities. V does not satisfy the associative property for area measurements. Thus, in general,

$$A \cdot V(A) + B \cdot V(B) \neq (A+B) \cdot V(A+B) \qquad (4)$$

where A and B are contiguous areas. The discrepancy increases with the spatial inhomogeneity of the area. Thus in general

$$f \cdot \frac{(a_{2v} - a_{1v})}{(a_{2v} + a_{1v})} + (1 - f) \cdot \frac{(a_{2g} - a_{1g})}{(a_{2g} + a_{1g})} \neq \frac{[\, f \cdot (a_{2v} - a_{1v}) + (1 - f) \cdot (a_{2g} - a_{1g})\,]}{[\, f \cdot (a_{2v} + a_{1v}) + (1 - f) \cdot (a_{2g} + a_{1g})\,]} \qquad (5)$$

as illustrated in figure 3 for a mixture of totally vegetated (v) and non-vegetated (g) areas with fractions f and 1-f. Evidently the straight line relationship between V and vegetation cover is not reproduced by a hypothetical sensor which observes an area of mixed cover, where the difference can be close to 30%.

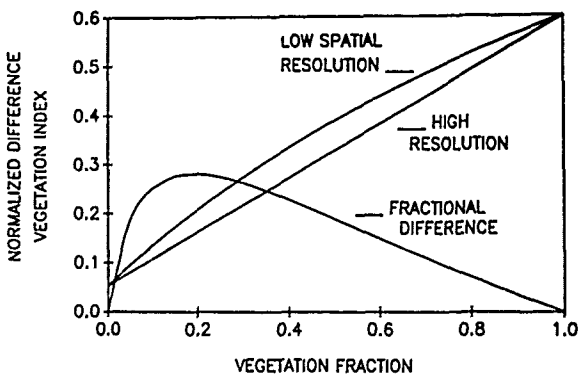


Figure 3. Normalized vegetation index does not satisfy the associative property, as illustrated by the difference between hypothetical high resolution measurements and low resolution (average) values.

Evidently additional information is required to distinguish a uniform vegetation cover of low density, as a crop in early growth, or a drought distressed area, from an area with midseason vegetation and bare or fallow areas.

3. Evapotranspiration from Heterogeneous Areas

Because the interaction of the earth's surface and atmosphere is extremely complex, we address only estimation of the moisture flux. The problem has two aspects, availability of surface parameters, and the method of analysis, given these data.

Firstly, a regional description of vegetation and soil characteristics is not available. One could not go to a general location and find values for surface

roughness, moisture characteristics, etc. Neither conventional nor satellite based work to date emphasizes the values for these quantities at regional scale. Such data could also be obtained from a sampling of ground sites.

Secondly, we do not know the fraction of different surface types for each AVHRR pixel. In most regions farming, forest clearing, irrigation projects, and changing habitation density produce nonuniformity of the earth's surface. If a detailed description of the land surface is not required, then the procedure developed here applies.

A. Procedure

We allow for variable surface conditions and mixed surface types. Each area is comprised of a mixture of a vegetated area and an area of bare soil and/or dead vegetation. We compute values for the "mixed pixels", or heterogeneous areas. For such measurements the satellite measured radiance is $\Sigma a = a_v + a_g$ and the vegetation index at p is

$$V_p = f \cdot V_v + (1-f) \cdot V_g \qquad (6)$$

Using Eq. 2 for surface temperature, and for a 20-40C temperature range,

$$T_p = f \cdot T_v + (1 - f) \cdot T_g \qquad (7)$$

where T_p is the weighted sum of values from soil and vegetation. Error sources in section 2 tend to destroy the relationship between temperature and vegetation index. However many data sets show a linear relationship between T_s and V (Fig. 4).

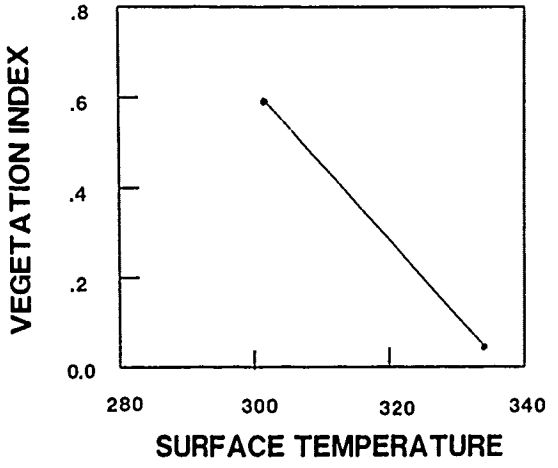


Figure 4. Typical relationship of vegetation index and temperature.

Thus we may use the ratio,

$$R = \frac{f \cdot V_v + (1-f) \cdot V_g}{f \cdot T_v + (1-f) \cdot T_g} \qquad (8)$$

or a two dimensional histogram of the data to estimate temperatures for vegetation and bare soil. Given these temperatures we derive values for ET of the two surface types from models which use surface meteorological data plus assumed surface parameters. The proper ET yields a temperature equal to that seen by the satellite. Regional evapotranspiration is obtained by summing weighted values from bare soil and vegetation, where vegetation index gives the proportions.

B. Application to a satellite data set

The data discussed represents a large relatively cloud free area at a time of local variability in moisture conditions. Figures 5 and 6 illustrate V and T_s in the central plains.

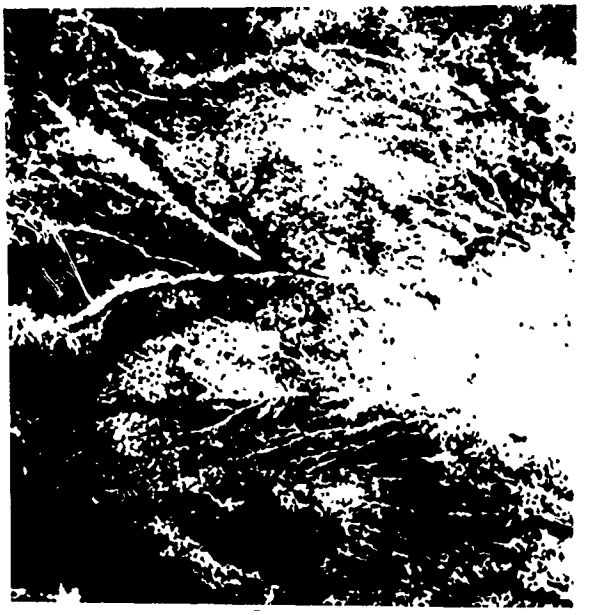


Figure 5. Vegetation index on July 20, 1981 for an area in western Nebraska. At the lower right is an irrigated region on the North Platte river.

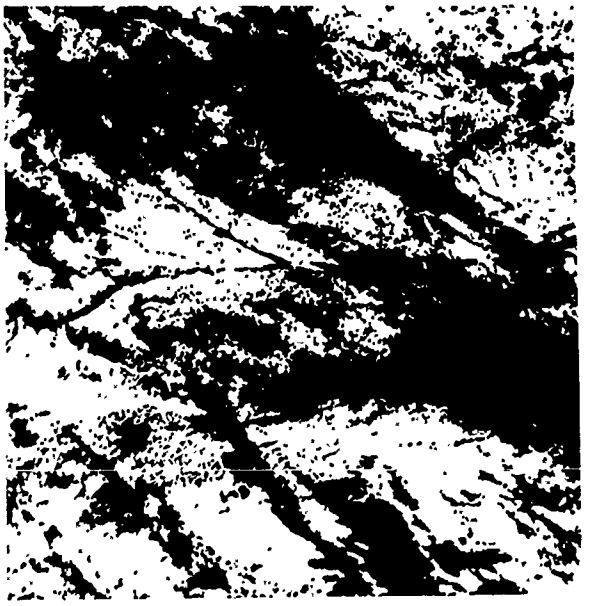


Figure 6. Surface temperatures for area of figure 5.

Comparison shows the association of high temperature with low vegetation index. However a substantial portion of the total non-vegetated area is cool (dark) relative to other areas which are also not vegetated. We relate cooler patterns to increased surface moisture in the soil and stubble remaining from harvested winter wheat. The population of T-V values (Fig 7) corresponds to mixtures of vegetation, dry soil, and wet soil.

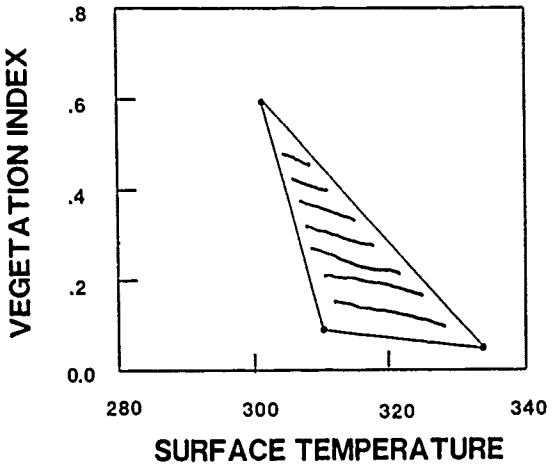


Figure 7. Form of the population of values from figures 5 and 6.

We treat soil of intermediate moisture as a weighted sum of wet and dry soil. Then letting subscript 1 correspond to vegetation, 2 to dry soil, and 3 to wet soil, we have

$$T_p = f_1 T_1 + f_2 T_2 + f_3 T_3$$
$$V_p = f_1 V_1 + f_2 V_2 + f_3 V_3 \qquad (10)$$

where $f_1 + f_2 + f_3 = 1$. The solution for f_i is

$$f_i = \frac{T_p \cdot (V_j - V_k) + V_p \cdot (T_k - T_j) + T_j V_k - T_k V_j}{T_i \cdot (V_j - V_k) + V_i \cdot (T_k - T_j) + T_j V_k - T_k V_j} \qquad i \neq j \neq k = 1,2,3 \qquad (11)$$

The moisture flux given T_p, V_p is expressed as

$$LE(T_p, V_p) = f_1 \cdot LE(\text{dry soil}) + f_2 \cdot LE(\text{moist soil}) + f_3 \cdot LE(\text{vegetation}) \qquad (12)$$

where the LE_i are provided by numerical simulation (Price, 1989) using surface weather observations from North Platte. If we normalize to unity the two dimensional probability function P then total area ET may be expressed as area averages. Thus

$$\langle LE \rangle = \sum_{i \neq j \neq k} LE_i \cdot \{ \langle T \rangle \cdot (V_j - V_k) - \langle V \rangle \cdot (T_j - T_k) + (T_j V_k - T_k V_j) \} / N \qquad (13)$$

where $N = \sum_{i \neq j \neq k} \{T_i \cdot (V_j - V_k)\}$, i, j and k are in cyclic order, and the brackets < > represent area averages.

The utility of this approach depends on using figure 7 to locate (T_i, V_i) for pure vegetation cover, for dry soil, and for moist soil. These values permit estimation of the LE_i values, which in turn permit estimation of LE for mixed surface types.

Figures 5 and 6 include small water bodies as well as scattered cumulus clouds. While clouds should be eliminated, water bodies provide additional context through their temperature and reflectance values. Such information could be used to correct satellite data for atmospheric effects.

Discrimination of Moist Soil from Cirrus Clouds

The darker areas in figure 6 resemble cirrus clouds. We can eliminate this possibility. Thin cirrus lowers temperatures of all surface features, including vegetation. Thus we expect to see a temperature difference between vegetation in cirrus areas, versus that in uncontaminated areas. In the example vegetation temperatures are the same in both areas. Thus the blotchy appearance is due to surface moisture variations

4. Conclusion and recommendations

A method has been developed which relates contextual information in AVHRR data to regional ET. This procedure permits estimation of ET in areas where spatial variability prevents knowledge of detailed surface descriptions. Instead only regional values of vegetation and soil characteristics are required. An atlas or data base for surface characteristics is needed for such a description.

References

Carlson, T. N., J.K Dodd, S.G. Benjamin, and J.N. Cooper, 1981, Satellite Estimation of Surface Energy Balance, Moisture Availability and Thermal Inertia, J. Applied Meteor., 20: pp 67 - 87.

Gutman, G., 1987, The derivation of vegetation indices from AVHRR data, Int. J. Remote Sensing, 8, 1235-1243

Gutman, G. 1988, A Simple Method for Estimating Monthly Mean Albedo of Land Surfaces from AVHRR Data, J. Appl. Met., 27, pp 973-988

Perry, C.R., and Lautenschlager, L.R., 1984, Functional Equivalence of spectral vegetation indices, Remote Sens of Environ., 14, pp 169-182

Price, J. C., 1982a, On the Use of Satellite Data to Infer Surface Fluxes at Meteorological Scales, J. Appl. Met 21, pp 1111-1122.

Price, J.C., 1984, Land Surface Measurements from the Split Window Channels of the NOAA-7 Advanced Very High Resolution Radiometer, J. Geophys. Res., 89, pp 7231-7237.

Price, J.C., 1989, Submitted to Journal of Applied Meteorology.

Taconet, O. R. Bernard, and D. Vidal-Madjar, 1986, Evapotranspiration over an Agricultural Region Using a Surface Flux/Temperature Model Based on NOAA- AVHRR Data, J Clim and Appl. Met, 25, pp284-307

A ONE- AND TWO-LAYER MODEL FOR ESTIMATING EVAPOTRANSPIRATION WITH REMOTELY SENSED SURFACE TEMPERATURE AND GROUND-BASED METEOROLOGICAL DATA OVER PARTIAL CANOPY COVER

William P. Kustas
USDA-ARS Hydrology Laboratory
Beltsville, MD 20705

Bhaskar J. Choudhury
NASA-Goddard Space Flight Center
Greenbelt, MD 20771

Kenneth E. Kunkel
Illinois State Water Survey
2204 Griffith Drive
Champagne, IL 61820-7495

ABSTRACT

Surface-air temperature differences are commonly used in a bulk resistance equation for estimating sensible heat flux, H. This estimate of H is inserted in the one-dimensional energy balance equation to solve for the latent heat flux, LE, as a residual. Serious discrepancies between estimated and measured LE have been observed for partial canopy cover conditions, which are mainly attributed to inappropriate estimates of H. In an attempt to improve the estimates of H over sparse canopies and as a result compute more accurate values of LE, one- and two-layer resistance models are developed to account for some of the factors causing poor agreement. The utility of these two approaches for estimating LE at the field scale is tested with remotely sensed and micrometeorological data collected in an arid environment from a furrowed cotton field with 20 percent cover and a dry soil surface. It is found that the one-layer model performed better than the two-layer model, but only when a theoretical bluff-body correction for heat transfer instead of an empirical adjustment, thought to be applicable to a fairly wide range of conditions, is used. This is attributed to the significant size of the furrows relative to the height of the vegetation.

Keywords: Energy balance, latent heat flux, resistance models, sensible heat flux, thermal IR

1. INTRODUCTION

One of the applications of remote sensing data in the thermal infrared (IR) spectrum is to determine the skin temperature of the earth's surface. This information has been incorporated in various models for computing the evapotranspiration (ET), which is usually the second largest quantity (precipitation being first) in the hydrologic cycle.

This paper discusses a method whereby ET is solved as a residual in the one-dimensional surface energy balance equation [1], viz.,

$$LE = Rn - G - H \qquad (1)$$

In Eq. (1) LE, Rn, G, and H are the latent heat flux, net radiation, soil heat flux, and sensible heat flux, respectively. Energy associated with photosynthesis, stored in the vegetation and in the canopy air space is neglected.

Estimates of Rn and G come from micrometeorological measurements or from a combination of meteorological and remote sensing data [2]. Attempts to provide a simplified parameterization for estimating sensible heat flux have not been very successful because of the relatively strong dependence of the transfer coefficient on various surface and meteorological conditions [3]. Consequently, computing reliable values of H represents the most formidable obstacle in the residual method (i.e., Eq. (1)). It is the objective of this paper to consider one- and two-layer parameterizations of the surface energy balance over partial canopy cover to assess the complexity required to obtain accurate estimates of ET. This will mainly involve the comparison of various schemes for estimating the one-dimensional resistance to heat transfer across the substrate-canopy-atmosphere interface.

2. ONE- AND TWO-LAYER MODEL FORMULATIONS

With mean wind speed and air temperature gradients several meters above the surface, the eddy diffusivity concept has been a common method for computing H. The integrated form of this equation has typically been expressed as a resistance type formula, analogous to Ohm's Law [4]. With a composite radiometric temperature of the surface, the one-layer approach combines the substrate and vegetation into a single idealized surface, yielding bulk resistance equation, i.e.,

$$H = \rho C_p (T_s - T_a)/R_{ah} \qquad (2)$$

where ρC_p is the volumetric heat capacity (J m^{-3} K^{-1}), T_a (K) air temperature at height z (several meters above the canopy), T_s (K) the representative temperature of the idealized one-layer surface and R_{ah} (s m^{-1}) is the bulk resistance to heat transfer across the surface-atmosphere interface. Recent studies have shown a consistent inequality between thermometric measurements of T_s and T_s evaluated by Eq. (2) [5], [6]. Clearly, view angle of the thermal radiometer as well as canopy geometry and canopy temperature profile will also significantly affect the value of T_s employed in the one-layer approach [5].

The formula for R_{ah} is derived by invoking surface layer similarity theory for scalars (e.g., temperature) which yields an expression of the form

$$R_{ah} = \left[\left(\ln ((z-d_o)/z_{om}) + \ln (z_{om}/z_{oh}) - \psi_h\right)\left(\ln ((z-d_o)/z_{om}) - \psi_m\right)\right]/k^2 u \quad (3)$$

The symbols ψ_h and ψ_m are stability correction functions of heat and momentum, d_o and z_{om} the displacement height and roughness length for momentum, z_{oh} the roughness length for heat, k (≈0.4) Von Karman's constant and u the wind speed at level z. The distinction made in roughness lengths for heat and momentum is to account for differences in transfer processes when close to the surface [7]. Historically the relationship between z_{oh} and z_{om}, has been expressed in a dimensionless form, viz., $\ln (z_{om}/z_{oh})$ or kB^{-1} [8]. For vegetative surfaces, data and theoretical results suggest a constant $kB^{-1} \sim 2$; but, for bluff-rough elements (e.g., a furrowed or plowed field) theoretical results show that kB^{-1} can vary from 2 to around 10. Moreover, a recent analysis with (2) over sparse canopy cover required that kB^{-1} be made a function of the thermometric surface temperature observed from a nadir viewing thermal IR sensor to obtain satisfactory results [9]. Since the present analysis considers sparse cover, the equation for kB^{-1} developed by Kustas et al. [9] was adopted, i.e.,

$$kB^{-1} = 0.17\ u\ (T_s - T_a) \quad (4a)$$

However, because the field contained relatively large furrows compared to the height of the vegetation (see below) the theoretical equation for bluff-rough obstacles was also employed [10]

$$kB^{-1} = 2.46(z_{o+})^{1/4} - 2 \quad (4b)$$

The symbol z_{o+} is a roughness Reynolds number ($= u_* z_{om}/\nu$) where ν is the kinematic viscosity and u_* the shear velocity.

Partial cover conditions have forced others to consider the energy balance of the substrate and vegetation separately to account for large differences in the source/sink heights of momentum, sensible and latent heat (e.g., [11], [12]). In this paper only the partitioning of net radiation, soil and sensible heat flux between the vegetation and the soil surface is explored since the utility of the residual approach (i.e., Eq. (1)) is to be tested. Figure 1 illustrates the energy partitioning between the vegetation and the soil substrate, and the resistance network for heat transfer in a two-layer model.

The mean canopy resistance to heat transfer r_{ac} was evaluated by first considering the leaf boundary layer resistance, r_b. An equation from Jones [13] was utilized,

$$r_b = A\ (w/u)^{1/2} \quad (5)$$

where A is a constant of order 90 $s^{1/2}\ m^{-1}$ [4], w a characteristic length scale taken as the average leaf width (~0.05 m), and u is the local wind speed at some level inside the canopy. In order to calculate a mean resistance per unit leaf area, $\bar{r}_b$, the wind profile inside the canopy had to be estimated.

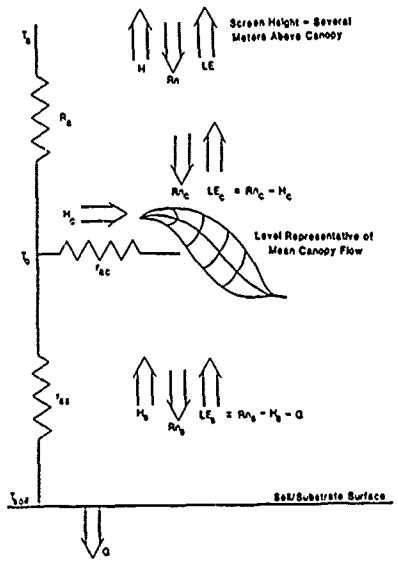


Fig. 1. Schematic diagram of the two-layer model illustrating the resistance network and the partitioning of the energy fluxes between the soil/substrate and vegetation. The representative temperature within the canopy air space, T_o, has been equated to the aerodynamic temperature [5]. Adapted from Shuttleworth and Wallace [11].

An exponentially decreasing u with height was employed, which previous studies have shown to be a reasonable approximation [10],

$$u = u_h \exp [-\beta(1 - z/h)] \quad (6)$$

The symbol h is the canopy height (~ 0.3 m), u_h the windspeed at h, and β is the extinction factor evaluated with the equations of Goudriaan [14]. Values of u_h were calculated with the wind profile data and roughness parameters from Kustas et al. [15], and assuming near-neutral conditions due to roughness sublayer effects [16]. With the leaf area index, LAI (~0.4), assumed uniform over the canopy height, Eqs. (5) and (6) were numerically integrated to obtain $\bar{r}_b$. The resistance per unit ground area r_{ac} was obtained by using the theoretical results of Shuttleworth [17], i.e.,

$$\bar{r}_{ac} = \bar{r}_b/LAI \quad (7)$$

Values of Rn_s and Rn_c were determined using Beer's Law, assuming Rn is proportional to the global shortwave radiation flux and neglecting longwave radiation in the canopy caused by vertical temperature gradients. For Rn_s this yields

$$Rn_s = Rn \exp (-\alpha\ LAI) \quad (8)$$

and from Fig. 1 this gives for Rn_c

$$Rn_c = Rn\ (1-\exp (-\alpha\ LAI)) \quad (9)$$

The attenuation coefficient, α, is primarily a function of the leaf angle distribution and solar elevation [4]. Adjustment to α for diffuse radiation

was neglected. The value of α was estimated by taking a spherical leaf distribution exposed at a sunangle θ, yielding $\alpha = 0.5 \csc \theta$.

In order to avoid estimating the representative temperature inside the canopy air space, T_o (see Fig. 1), expressions for H_c and H_s had to be developed, which were not explicitly dependent on T_o. An equation for H_c was derived from algebraic manipulation of the resistances in Fig. 1 [18]:

$$H_c = [\rho C_p (T_c - T_a) - H_s R_a]/(R_a + r_{ac}) \qquad (10)$$

The resistance R_a defined as the bulk aerodynamic resistance, was evaluated with the semi-empirical formula from Mahrt and Ek [19] for unstable conditions. A value of H_s was computed indirectly by approximating LE_s with the following expression

$$(Rn_s - G)/(1 + B_{os}) \qquad (11)$$

where B_{os}, the Bowen ratio for the substrate, was assigned a constant value of 6 from the simulation results of a drying soil by Choudhury and Monteith [12]. Thus, the total latent heat flux of the substrate and canopy was estimated by combining the energy balance equations in Fig. 1 and (10)-(11), yielding,

$$LE = Rn_c - \left\{\rho C_p (T_c - T_a) - 6 (Rn_s - G) R_a/7\right\} /(R_a + r_{ac}) + (Rn_s - G)/7 \qquad (12)$$

with Rn_s and Rn_c evaluated by (8) and (9), respectively.

In summary, LE was solved with the one-layer model using Eqs. (1)-(4a), $LE_{1L}(a)$, and (1)-(4b), $LE_{1L}(b)$. The two-layer model estimated LE with Eq. (12), LE_{2L}.

3. THE DATA

The data for this study were collected over a furrowed cotton field 1500 m east-west x 300 m north-south located in Maricopa Farms in central Arizona (33.08°N, 111.98°W) from June 10, 1987 Day of Year (DOY) 161 to June 16, 1987 DOY 165. A detailed discussion of the micrometeorological instrumentation and agronomic measurements made in the cotton field is given by Kustas et al. [15]. Briefly, the agromonic measurements gave a crop height of 0.3 m spaced 1 m apart and a furrow depth of about 0.2 m with rows running north-south. Profiles of wind speed and temperature were determined at 5 levels above the furrow bottom in the center of the field. The fluxes LE and H were measured by the Bowen ratio/energy balance and eddy correlation method with instruments located within 150 m of each other and the profile mast. Differences in the half-hourly values of H and LE given by the two methods were analyzed by Kunkel et al. [20]. The eddy correlation measurements of H and LE solved by (1) were used for comparison with model results.

A thorough discussion and analysis of the ground- and airborne-based thermometric observations (nadir-looking) is given in Kustas et al. [21]. One of the findings significant to this study was that the composite surface temperature given by the measurement scheme for the ground-based observations of the sunlit and shaded soil temperatures and the canopy temperature were comparable to values obtained for one of the days by an aircraft flying at an altitude of 150 m. Consequently, the larger quantity of the ground-based radiometric data collected for field 28 between DOY 161-165 were used.

There were 19 thermometric observations made between DOY 161-165 that were used in this analysis. A weighted average of the 1/2 hour means of the terms in (1) and the meteorological data, u and T_a, were determined using the mean time of the thermal IR measurements.

4. RESULTS AND CONCLUSIONS

Estimates of LE with the one- and two-layer models are illustrated in Fig. 2. For the one-layer model, $LE_{1L}(a)$-values are significantly smaller than measured LE with appreciable scatter while $LE_{1L}(b)$-values are only slightly larger on average than the measured with much less variability (see Table 1). Hence, it appears that with the nadir thermometric observations providing a composite surface temperature, the theoretical bluff-rough formulation for kB^{-1} (i.e., Eq. (4b)) is more applicable. This conclusion seems physically plausible given the fact that the furrows were of relatively large size (i.e., ~1/2 the height of the vegetation) that they significantly augmented momentum transfer relative to heat transfer resulting in a large kB^{-1} required in (3). The parameterizations used in the two-layer model resulted in LE_{2L}-values underestimating measured LE, on average. But there was less scatter compared to $LE_{1L}(a)$-values.

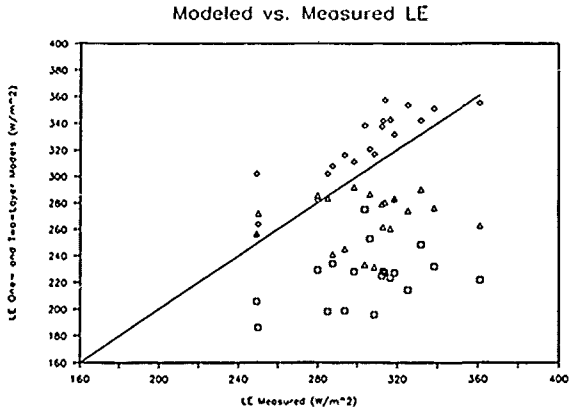


Fig. 2. Comparison of measured latent heat flux versus modeled estimates with $LE_{1L}(a)$ □, $LE_{1L}(b)$ ◇, and LE_{2L} △. The line represents perfect agreement with measured values.

Table 1 summarizes the results of the one- and two-layer models as the root mean square error (RMSE) between computed and observed values of latent heat flux. The results listed in the table suggest that the one-layer model with Eq. (4b) provides the most accurate estimates of LE. However, changes in view angle of a thermal IR sensor may result in significant changes in T_s and consequently in the adjustment to k_{ah}, which cannot be accounted for in the model. Moreover, the expression developed by Kustas et al. [9] was not applicable to this situation. Thus, the ability to adjust k_{ah} with simple formulations like Eq. (4a) may not be feasible.

Estimates from the two-layer model are an improvement over the $LE_{1L}(a)$-values, but are relatively poor in comparison to $LE_{1L}(b)$-values. However, only the

resistances to heat transfer in the model formulation were considered. When Penman-Monteith equations for the water vapor transport from the vegetation and the soil are included, model sensitivity in computing LE to errors in r_{ac} and R_a are relatively small [11]. Yet, this necessitates the estimation of soil and stomatal resistances to water vapor transfer which may be difficult to specify, especially under partial canopy conditions, without extensive ground-based measurements [18].

Table 1. Values of the root mean square error (RMSE) between model and predicted values of LE.

Model	RMSE ($W\ m^{-2}$)
LE_{1L}(a)	85
LE_{1L}(b)	24
LE_{2L}	48

6. REFERENCES

1. Jackson, R. D., "Evaluating evapotranspiration at local and regional scales", Proc. IEEE 73:1086-1095, 1985.

2. Jackson, R. D., S. M. Moran, L. W. Gay and L. H. Raymond, "Evaluation evaporation from field crops using airborne radiometry and ground-based meteorological data", Irrig. Sci. 8:81-90, 1987.

3. Seguin, B. and B. Itier, "Using midday surface temperature to estimate daily evaporation from satellite thermal IR data", Int. J. Remote Sens. 4:371-383, 1983.

4. Monteith, J. L., "Principles of Environmental Physics", Arnold, London, 1973.

5. Huband, N. D. S. and J. L. Monteith, "Radiative surface temperature and energy balance of a wheat canopy, I. Comparison of radiative and aerodynamic canopy temperature", Bound. Layer Meteorol. 36:1-17, 1986a.

6. Huband, N. D. S. and J. L. Monteith, "Radiative surface temperature and energy balance of a wheat canopy, II. Estimating fluxes of sensible and latent heat", Bound. Layer Meteorol. 36:107-116, 1986b.

7. Thom, A. S., "Momentum, mass and heat exchange of vegetation", Q. J. R. Meteorol. Soc. 98:124-134, 1972.

8. Chamberlain, A. C., "Transport of gases to and from surfaces with bluff and wave-like roughness elements", Q. J. R. Meteorol. Soc. 94:318-332, 1968.

9. Kustas, W. P., B. J. Choudhury, M. S. Moran, R. J. Reginato, R. D. Jackson, L. W. Gay and H. L. Weaver, "Determination of sensible heat flux over sparse canopy using thermal infrared data", Agric. For. Meteorol. 44:197-216, 1989a.

10. Brutsaert, W., "Evaporation into the Atmosphere, Theory, History and Applications", Reidel, Dordrecht, Holland, 1982.

11. Shuttleworth, W. J. and J. S. Wallace, "Evaporation from sparse crops - an energy combination theory", Q. J. R. Meteorol. Soc. 111:839-855, 1985.

12. Choudhury, B. J. and J. L. Monteith, "A four-layer model for the heat budget of homogeneous land surfaces", Q. J. R. Meteorol. Soc. 114:373-398, 1988.

13. Jones, H. G., "Plants and Microclimate", Cambridge University Press, New York, 1983.

14. Goudriaan, J., "Crop Micrometeorology: A Simulation Study Center for Agricultural Publishing and Documentation", Wageningen, The Netherlands, 1977.

15. Kustas, W. P., B. J. Choudhury, K. E. Kunkel, and L. W. Gay, "Estimate of the aerodynamic roughness parameters over an incomplete canopy cover of cotton", Agric. For. Meteorol., 1989b (in press).

16. Garratt, J. R., "Surface influence upon vertical profiles in the atmospheric near-surface layer", Q. J. R. Meteorol. Soc. 106:803-819, 1980.

17. Shuttleworth, W. J., "A one-dimensional theoretical description of the vegetation -atmosphere interaction", Bound. Layer Meteorol. 10:273-302, 1976.

18. Smith, R. C. G., H. D. Barrs and R. A. Fischer, "Inferring stomatal resistance of sparse crops from infrared measurements of foliage temperature", Agric. For. Meteorol. 47:183-189, 1988.

19. Mahrt and Ek, "The influence of atmospheric stability on potential evaporation," J. Clim. Appl. Meteorol. 23:222-228, 1984.

20. Kunkel, K. E., L. W. Gay and W. P. Kustas, "Comparison of eddy correlation and Bowen ratio measurements of the energy budget of a partially vegetated cotton field," 1989 (in preparation).

21. Kustas, W. P., B. J. Choudhury, Y. Inoue, P. J. Pinter, Jr., M. S. Moran, R. D. Jackson and R. J. Reginato, "Ground and aircraft infrared observations over a partially vegetated area", Int. J. Remote Sens., 1989c (in press).

CONTROL OF IRRIGATION BY FOLLOWING THE WATER BALANCE FROM NOAA-AVHRR THERMAL IR DATA

Alain VIDAL - Laboratoire Commun de Télédétection CEMAGREF-ENGREF BP 5095 - 34033 MONTPELLLIER CEDEX 1 - FRANCE

Alain PERRIER - INRA Bioclimatologie - 78850 THIVERVAL-GRIGNON - FRANCE

ABSTRACT

Satellite data give now localized and frequent information used in agriculture. Thermal IR data from NOAA AVHRR are used to map surface temperature of pixels of 1 Km^2. Physical relationships give evapotranspiration from surface temperature and other meteorological parameters usually measured.

Maps of evapotranspiration are then made for areas where sugar cane is dominant, and are used to follow the water balance and to control irrigation on this crop.

Key-words : NOAA AVHRR, Surface temperature, Evapotranspiration, Irrigation.

1 - INTRODUCTION

Recent results obtained by CEMAGREF (France), ORMVAG and IAV Hassan II (Morocco) in the irrigated area of Gharb (Morocco) show that remote sensing can be applied to the management and the control of irrigation. The study area is a plain of 250 000 ha, where 100 000 ha are irrigated, in North-West Morocco.

The estimation of evapotranspiration and the monitoring of water balance from satellite thermal infrared data have been widely studied in the past few years (SEGUIN, 1984). Recent studies carried out at the Station de Bioclimatologie INRA, Avignon, France (VIDAL et al., 1986) have resulted in a first mapping of daily evapotranspiration. It is based on the possibility of estimating evapotranspiration from an instantaneous value of the Ts-Ta difference near midday, according to a statistically established and theoretically justified linear relation (JACKSON et al., 1977 ; SEGUIN & ITIER, 1983 ; NIEUWENHUIS et al., 1985 ; RAMBAL et al., 1985) :

$$ETR - Rn = A - B\,(Ts - Ta) \qquad (1)$$

In all cases, ETR and Rn represent daily evapotranspiration and daily net radiation expressed in mm of water. Ts and Ta (respectively surface and air temperature) are obtained at various times (satellite pass for Ts, simultaneous or maximum value for Ta). Rn and Ta are obtained from the meteorological network.

This paper presents the method used to obtain the surface temperature from thermal IR data, then to map evapotranspiration of sugar cane from surface temperature. A new method of irrigation control using thermal IR data from NOAA satellite is then presented.

2 - LAND SURFACE TEMPERATURE OBTAINED FROM THERMAL IR DATA OF NOAA-AVHRR

Estimation of land surface temperatures from thermal infrared bands of NOAA satellites is nowadays usual. Radiances obtained in 2 atmospheric windows, where molecular absorption is minimum, give, by inversion of Planck's function, 2 black body temperatures. These temperatures, combined together by a method named "Split-window" (PRABHAKARA et al., 1975), give finally a surface temperature (DESCHAMPS & PHULPIN, 1980). Considering NOAA bands 4 (10.3 - 11.3 μm) and 5 (11.5 - 12.5 μm), the split-window equation can be written :

$$T_{BB} = a\,T_4 + b\,T_5 + c \qquad (2)$$

where T_{BB} is the equivalent black body surface temperature,
T_i is the black body temperature obtained from band i.

Recent results obtained by CEMAGREF (VIDAL, 1989 (1)) show that such an equation may be obtained directly from NOAA data, as far as, for certain days, the difference $T_4 - T_5$ is due to atmospheric effects. For the area of Gharb, this equation becomes :

$$T_{BB} = T_4 + 2.78\,(T_4 - T_5) \qquad (3)$$

Another problem is to obtain the actual land surface temperature from the black body equivalent surface temperature. One has then to take into account the emissivity of the land surface observed. This emissivity may vary between bands 4 and 5 of NOAA AVHRR : SALISBURY (1986) showed that there is a difference between land surface emissivities in bands 4 and 5 of NOAA AVHRR (respectively noticed ϵ_4 ϵ_5), and BECKER (1987) showed that :

$$Ts - T_{BB} = 50\,\frac{1 - \bar{\epsilon}}{\bar{\epsilon}} + 300\,\frac{\epsilon 4 - \epsilon 5}{\bar{\epsilon}} \qquad (4)$$

where Ts is the actual surface temperature,
$\bar{\epsilon}$ is the mean emissivity in bands 4 and 5.

If a mean emissivity of 0.96 is assumed for sugar cane, VIDAL (1989 (1)) showed that $\epsilon_4 - \epsilon_5$ may be

obtained also directly from NOAA data. One obtains:

$$\epsilon_4 - \epsilon_5 = 0.003 \qquad (5)$$

With such methods using only NOAA AVHRR data, it seems possible to obtain the actual land surface temperature with a global error of ± 3°K.

3 - ESTIMATION OF EVAPOTRANSPIRATION FROM SURFACE TEMPERATURE

Equation (1) varies according to the crop which is considered. The theoretical justification of this relation has been described by SEGUIN & ITIER (1983) for a thin surface. From Equation (6) of the energy balance of a thin surface, where H is the sensible heat flux, "d" means "daily" (fluxes in equivalent mm of water), "i" means "instantaneous" (fluxes in $W.m^{-2}$) if the ground heat flux is negligible, we have :

$$ETR_d = Rn_d + H_d \qquad (6)$$

An analysis of the terms A and B of (1), which depend on the crop, was made by VIDAL & PERRIER (1988).

3.1. Analysis of the term A

If we write :

$$D_i = H_d / Rn_d - H_i / Rn_i \qquad (7)$$

Then

$$A = \overline{D_i \cdot Rn_d} \qquad (8)$$

The agrometeorological model, initially developped by CHOISNEL (1977) and modified by VIDAL et al. (1986), simulates the energy and water balance of a crop on an hourly and a daily basis from meteorological data and simultaneously computes ETR and the mean Ts of the crop.

This model was used to test the sensitivity of parameters D_i and A. It appeared that their values were rather stable with time. It seemed then possible to compute mean A from (8), and to apply its value to a given area. Sensitivity to crop height was rather low and we could assume, for an irrigated sugar cane between 1 and 3 m high (under conditions of Gharb, Morocco) :

$$A = 0.00 \pm 0.02 \text{ mm} \qquad (9)$$

3.2. Analysis of the term B

We consider a developped crop, whose height is z_h (Figure 1). We suppose that all the surfaces of leaves are at the same surface temperature Ts, equal to the radiative temperature of the crop, which can be measured by IR radiometer, or by satellite (see also TACONET et al., 1986).

Convective exchanges between leaves and the atmosphere are controlled by r_a, r_c and $1/h_i$, described on figure 1.

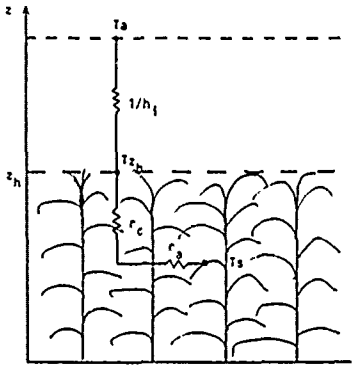


Figure 1 : The one-vegetation-layer, with various resistances to convective exchanges, and temperatures (From VIDAL & PERRIER, 1988) :
- r_a, aerodynamic resistance of the leaves,
- r_c, structure resistance due to stratification of leaves in the crop,
- $1/h_i$, aerodynamic resistance between the top of the crop and the reference level,
- Ts, surface temperature of the leaves, equal to radiative temperature,
- Tz_h, air temperature at level z_h, height of the crop,
- Ta, air temperature at reference level.

Then, H_i can be written :

$$H_i = -\rho C_p h_i (Tz_h - Ta) = -\rho C_p \frac{(Ts - Tz_h)}{r_a + r_c} =$$

$$-\rho C_p h_i \frac{(Ts - Ta)}{h_i (r_a + r_c) + 1} \qquad (10)$$

we notice :

$$r_o = r_a + r_c \qquad (11)$$

as the global resistance in the crop.

Finally, we can write :

$$B = \overline{(Rn_d / Rn_i) \rho C_p h_i / (1 + h_i r_o)} \qquad (12)$$

The sensitivity of B to the crop height for a wheat crop and a sugar cane crop was simulated by VIDAL & PERRIER (1988). The effect of z_h on B is attenuated by the effect of r_o, as it can be seen in equation (12). Finally, we could write :

Irrigated wheat : $B = 0.30 \pm 0.05$
Irrigated sugar cane :
$B = 0.25 \pm 0.05$ (13)

4 - CONTROL OF IRRIGATIONS BY USING MAPS OF EVAPOTRANSPIRATION

NOAA AVHRR data give thus maps of surface temperature as shown in figure 2.
From such a map, one may compute a map of actual evapotranspiration, as in figure 3, by applying equation (1) with A=0 and B=0.25. The global error on ETR does not exceed 1 mm per day.

It is possible to set a method to control irrigations, presented by VIDAL & BAQRI (1988).

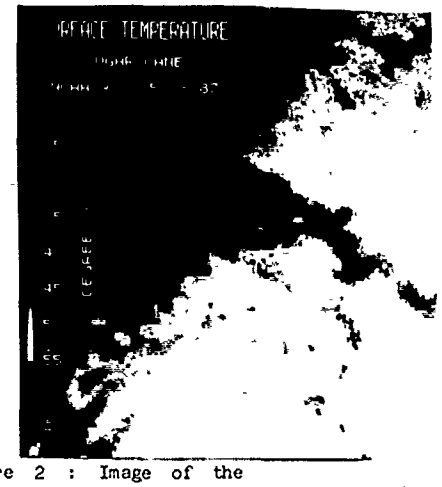


Figure 2 : Image of the surface temperature in the area of Gharb (Morocco). The area displayed is approximatively (100 Km)^2. From VIDAL (1989 (1)).

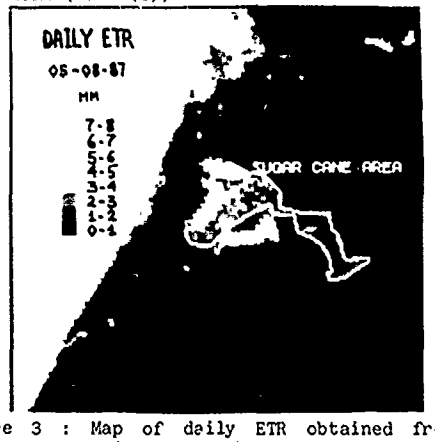


Figure 3 : Map of daily ETR obtained from the previous image (see fig. 2). The surrounded area is the area where sugar cane is dominant.

4.1. Hypothesises

If we consider that NOAA pixels (1 km x 1 km) are generally heterogeneous, we limit us to areas dominated by sugar cane, where it is possible to map evapotranspiration, as the surrounded area on figure 3. This choice implicitly assumes that there is a global stress on such areas, due to low efficiencies in the irrigation network. This last point concerns essentially secondary channels feeding an irrigated area of 500 to 1000 ha, which represents 5 to 10 NOAA pixels.

Thus the control of irrigations by remote sensing cannot be made at the scale of the plot, but only on the most disadvantaged areas, at the scale of a NOAA pixel.

4.2. Method of control

We present here the method applied to a homogeneous plot, whose whole surface is simultaneous irrigated.

Real time starting and management of irrigation are made from a reference plot, giving the maximal evapotranspiration (METR) of the crop.

The relationship between the ratio ETR/METR and the available water capacity (AWC) of the soil is presented in figure 4. CM and Cm may be estimated from ground measurements for a soil-crop couple.

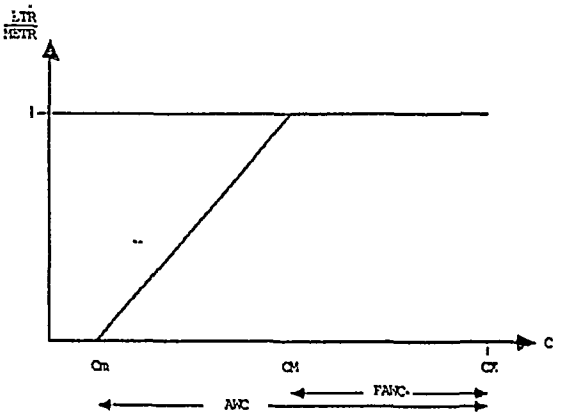


Figure 4 : Relationship between the ratio ETR/METR and the soil water capacity C.
- AWC = Available Water Capacity = CX - Cm,
- FAWC = Freely Available Water Capacity = CX - CM,
- ETR/METR = 1 for $C \geqslant CM$,
- ETR/METR = $\frac{C - Cm}{CM - Cm}$ for $Cm \leqslant C \leqslant CM$,
- ETR/METR = 0 for $C \leqslant Cm$.

We suppose that the crop is irrigated on day I, and that it was observed by satellite on day $J \leqslant I$ (see figure 5) ; ETR on day J is only known on day J+7, considering the period between acquisition and processing of the data. We also suppose that this crop is disadvantaged compared to the reference plot, so that :

$$C(J) < CM \text{ and } ETR(J) < METR \qquad (14)$$

Remote sensed ETR gives then the deficit D(J) of the observed crop, compared to CM :

$$D(J) = CM - C(J) = \left(1 - \frac{ETR}{METR}\right)(CM - Cm) \qquad (15)$$

If we obtain an image on day I (day of irrigation), we know then D(I). The crop receives an amount of water equal to FAWC = CX - CM, and, on day J+7, the deficit D(J+7) = D(I) is conserved. There are then 2 solutions to make up this deficit at the next irrigation, described on figure 6.

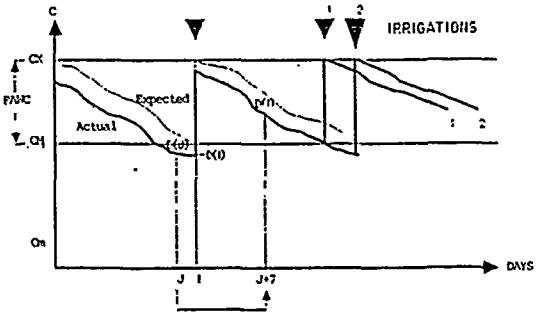


Figure 6 : Evolution of the soil water capacity C vs time (notations : see fig. 5). Remote sensing gives on day J+7 the deficit D(J)=CM-C(J) or, better, D(I). This deficit gives the difference between the actual level of C (thick line) and the expected one (thin line). It may be maked up by an earlier irrigation with the usual amount of water (1), or by

an irrigation on the expected day with a arger amount of water (2).

The same method may be applied at the NOAA pixel scale. From the previous map of daily ETR, one can compute a map of daily ETR, as shown on Figure 7, in order to detect the mean deficit by pixel. Such a map indicates the areas to be favoured at the next irrigation, using one of the two presented solutions.

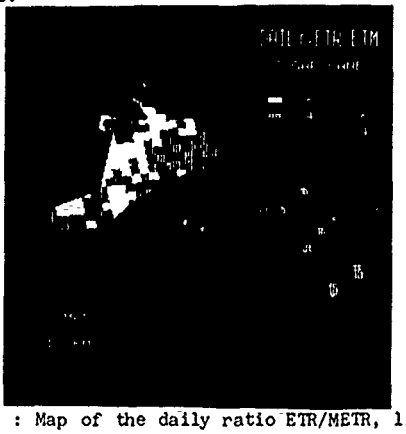


Figure 7 : Map of the daily ratio ETR/METR, limited to the area where sugar cane is dominant. This map shows the areas to be favoured at the next irrigation.

5 - CONCLUSION

It seems thus possible to control irrigations at the scale of NOAA pixels (1 km^2), by using data of thermal IR bands. This control may be made during the irrigation period, and the results obtained may then be precised by using high resolution remote sensing (Spot or Landsat data), whose information will be useful analysed before the next year (VIDAL, 1989 (2)).

ACKNOWLEDGEMENTS

The research is being conducted within the framework of the co-operative program between CEMAGREF ORMVAG and IAV Hassan II (Morocco), entitled "Application of Remote Sensing to Management and Control of Large Irrigated Areas".

REFERENCES

BECKER F., 1987. The impact of spectral emissivity on the measurement of land surface temperature from a satellite. Int. J. Rem. Sens., 8, p. 1509-1522.

CHOISNEL E., 1977. Le bilan d'énergie et le bilan hydrique du sol. La Météorolgie, n° spécial "Evapotranspiration", VI, 11, p.103-133.

DESCHAMPS P.Y & PHULPIN T., 1980. Atmospheric correction of infrared measurements of sea surface temperature using channels at 3.7, 11 and 12 m. Boundary-Layer Meteorol., 18 (1980), p.131-143.

JACKSON R.D., REGINATO R.J., IDSO S.B., 1977. Wheat canopy temperature : a practical tool of evaluating water requirements. Water Resour. Res., 13 (3), p.651-656.

NIEUWENHUIS G.J.A., SMIDT E.H., THUNISSEN H.A.M., 1985. Estimation of regional evapotranspiration of arable crops from thermal infrared images. Int. J. Rem. Sens., 6(8), p.1319-1334.

PRABHAKARA C., DALU G., KUNDE V.G., 1975. Estimation of sea surface temperature from remote sensing in the 11 and 13 m window region. J. Geophys. Res., 79, p.5039-5044.

RAMBAL S., LACAZE B., MAZUREK K., DEBUSSCHE G., 1985. Comparison of hydrologically simulated and remotely sensed actual evapotranspiration from some Mediterranean vegetation formations. Int. J. Rem. Sens., 6 (8), p.1475-1481.

SALISBURY J.W., 1986. Preliminary measurements of leaf spectral reflectance in the 8-14 μm region. Int. J. Remote Sensing, 7 (12), p.1879-1886.

SEGUIN B., 1984. Estimation de l'évapotranspiration à partir de l'infrarouge thermique. IIème Coll. Int. Signatures spectrales d'objets en télédétection. Bordeaux, 12-16 sept. 1983. Ed. INRA Publ., 1984 (Les colloques de l'INRA, n°23), p.428-446.

SEGUIN B., ITIER B., 1983. Using midday surface temperature to estimate daily evaporation from satellite thermal IR data. Int. J. Remote Sens. 1983, 4, 2, p.371-383.

TACONET O., BERNARD R., VIDAL-MADJAR D., 1986. Evapotranspiration over an agricultural region using a surface flux/temperature model based on NOAA/AVHRR data. J. Clim. Appl. Meteor., 25, p.284-307.

VIDAL A., KERR Y., LAGOUARDE J.P., SEGUIN B., 1986. Remote sensing and water balance : combined use of an agrometeorological model and of NOAA-AVHRR satellite thermal IR datas. Agric. For. Meteor., 1987, 39(2-3), p.155-175.

VIDAL A. & BAQRI A., 1988. Télédétection et contrôle de l'irrigation. Perspectives d'avenir. Génie Rural, 11, Nov. 1988, p.32-39.

VIDAL A. & PERRIER A., 1988. Analysis of a simplified relation used to estimate daily evapotranspiration from satellite thermal IR data. To be published in Int. J. Rem. Sens.

VIDAL A., 1989. Atmospheric and emissivity correction of land surface temperature measured from satellite using ground measurements or satellite data. Submitted to Int. J. Rem. Sens.

VIDAL A., 1989. Remote sensing and management of large irrigated areas. Proceeding of the Round Table Conference "Ressources Management from Space" held in New Delhi (India) on 6-7 February 1989.

REMOTELY SENSED SOIL MOISTURE INPUT TO A HYDROLOGIC MODEL

E T Engman and W P Kustas

USDA Hydrology Laboratory
Beltsville Agricultural Research Center
Beltsville, MD 20705

J R Wang

NASA-Goddard Space Flight Center
Greenbelt, MD 20771

ABSTRACT

The water balance of a small drainage basin has been simulated using a simple storage model. Aircraft microwave measurements of soil moisture have been used to construct two-dimensional maps of the spatial distribution of the soil moisture. Using data from flights on different dates also provided the temporal changes resulting from soil drainage and evapotranspiration. A time series of soil moisture appears to be a valuable new data form for verifying model performance and for updating and correcting simulated streamflow.

Keywords: Evapotranspiration, Microwave, Storage Model, Streamflow, Water Balance

1. INTRODUCTION

This study is part of the First International Satellite Land Surface Climatology Project Field Experiment (FIFE) that was conducted at and around the Konza Prairie Long Term Ecological Research site near Manhattan, Kansas during the summer of 1987. The objectives of FIFE are to understand better the role of biology in controlling the interactions between the atmosphere and the "vegetated" land surface and to investigate the use of satellite observations to infer climatologically significant land-surface parameters [1]. One phase of FIFE is an investigation of soil moisture and its role in the biology and climatology of the region. Although there are no operational satellite sensors suitable for measuring soil moisture, an aircraft-borne microwave radiometer was used to make remotely sensed measurements of soil moisture.

One of the FIFE objectives was to establish the relationship between areal soil moisture and the other water balance variables (i.e., rainfall, evapotranspiration, streamflow, and groundwater recharge). During the first FIFE experimental campaign we were able to collect spatial soil moisture data for several days during a dry-down period following a heavy rain. These data gave us, for the first time, the ability to temporarily map soil moisture over a small drainage basin and relate these data to the hydrology of the basin, specifically the base flow over a seven-day period [2]. This paper examines the modeling results from this previous work and discusses how remotely sensed soil moisture might be used as input data to a hydrologic model. The idea being pursued is that if detailed spatial soil moisture maps could be produced on a frequent basis, say every day or two, how might these data be used in a new input variable to a runoff model and what complementary data would also be required. The possibility of obtaining such data is very real for the EOS era.

2. STUDY SITE AND DATA COLLECTION

The drainage basin selected for study is a 37.7 ha area totally in natural grasses. The area is drained by a stream and flow system that flows only during the winter and late spring when precipitation is frequent. During the summer and fall the stream is dry except for heavy rainstorms and then the streamflow is totally storm runoff. Streamflow was measured with a 120° sharp crested weir and a clock driven analog recorder. Low flow calibration was accomplished with timed volumetric measurements. Figure 1 is a detailed map of watershed 1D that shows the stream network, approximate elevation contours and instrument locations.

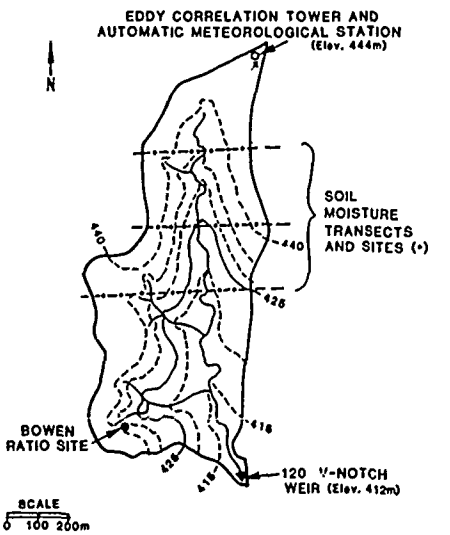


Fig. 1. Detailed map of watershed 1D showing approximate elevation contours, instrument locations and soil moisture sampling transects.

Evapotranspiration (ET) and rainfall were measured as part of the overall FIFE instrument network. ET was measured by several techniques including the Bowen ratio and eddy correlation methods. General weather conditions were measured by several automatic meteorological stations. Near surface soil moisture was measured with a push broom microwave radiometer (PBMR) mounted in a NASA C-130 aircraft. The radiometer operates at the 21-cm wavelength which has been shown to be most effective for measuring soil moisture [3]. It operates with four beams which sweep a path with a width of about 1.2 times the aircraft's altitude above the ground. The resolution of each beam is about 0.3 times the altitude. The PBMR measures the electromagnetic emission from the soil surface as a brightness temperature. Generally, a calibration curve is developed from the ground measurements of soil moisture and the measured brightness temperature normalized by the actual temperature of the top 5 cm soil layer. Volumetric soil moisture data from the transects were used to develop the calibration curve. The PBMR provides a quasi two-dimensional data set. Each of the four beams traces out a swath about 90 m wide so that the total swath is about 350 m wide for an aircraft altitude of 300 m.

Ground measurements were taken to support the remotely sensed data. Three transects were established under the aircraft flight lines for taking gravimetric soil moisture samples. The transects and sample sites are also shown in Fig. 1. The bulk density of the soil at each site was measured by a water displacement method. Full or partial transect data were taken on all except one day (June 2) of the study period. Aircraft PBMR data were taken on four of the days.

As stated above, the period of study coinciding with FIFE began with a soaking rain on May 27 and was followed by an essentially rain-free period or drydown for the next eight days. This provided the opportunity to trace the changes in soil moisture in response to a high ET demand and gravity drainage to the stream. Table 1 summarizes the daily climatic conditions during this period.

Table 1. Summary of climatology for the period May 27 to June 4, 1987. Dates of microwave flights and soil moisture are also shown.

Date	Rain (mm)	ET (mm)	Streamflow (mm)
May 27	61.72	0.00	0.54
May 28	0.08	3.63*	0.16
May 29	0.68	2.37*	0.10
May 30	1.69	5.69	0.08
May 31	0.00	6.22	0.09
Jun 1	0.00	5.70	0.08
Jun 2	0.25	1.17*	0.10
Jun 3	0.00	5.91	0.07
Jun 4	0.00	6.40	0.06

*Full day's data not available; total daily value estimated using data from other Bowen ratio sites.

3. SPATIAL DISTRIBUTION OF SOIL MOISTURE

Figure 2 shows the results of remotely sensed soil moisture in a two-dimensional map format for the days of May 28 and 30, and June 4. The numbers associated with the dashed contours indicate the volumetric soil moisture values in percent. The north and south ends of the basin are missing soil moisture data for two days because the aircraft flight lines were displaced slightly. Examination of soil moisture maps for the three dates points up both the temporal changes and certain locations where the changes in soil moisture are occurring at different rates. The June 4 map shows rather uniformly low soil moisture over most of the basin except along the southern boundary. This relatively high area of soil moisture is probably the result of down gradient redistribution of the soil moisture. This area also would represent the more dynamic region of the basin, in partial area or source area terms.

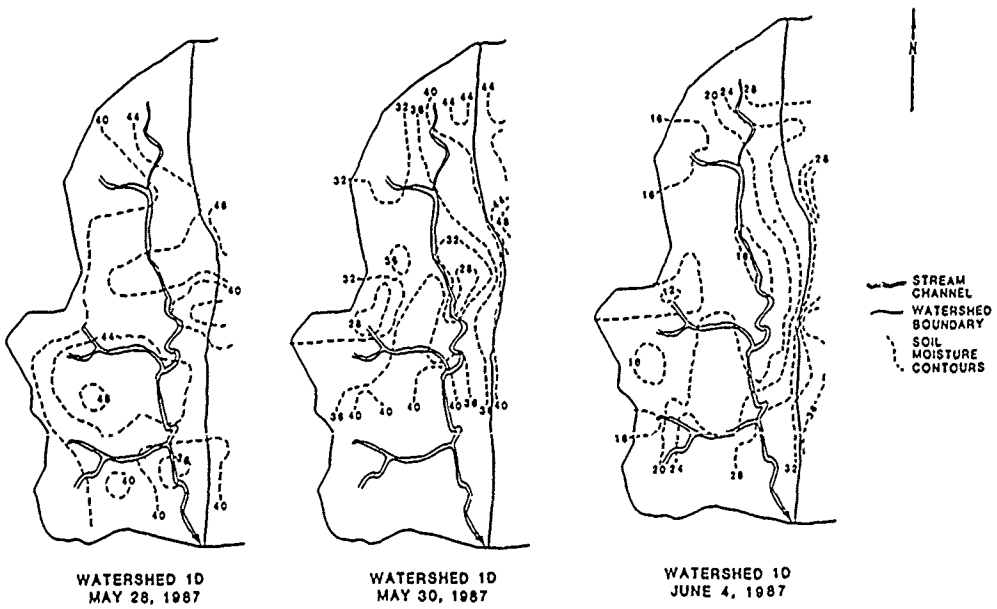


Fig. 2. Soil moisture map developed from PBMR data for three dates: May 28, May 30 and June 4 [4].

4. MODEL SELECTION

The availability of spatial and temporal soil moisture data provided a unique opportunity to model the hydrology of the basin where the soil moisture aspects could be emphasized. Our search for a model that has a reasonably sound theoretical base and focuses on the between storm base flow led us to several models that are all based on Hewlett and Hibbert [5] experiments. This class of models, which will be referred to as sloping slab models, was developed to simulate subsurface stormflow from relatively steep and primarily forested watersheds. Our search led us to a simple storage-discharge model developed by Sloan and Moore [6]. Figure 3 is a schematic of the Sloan and Moore [6] model. This model is based on a water balance of a hillside segment that acts as the control volume for the continuity of mass equation. The water balance aspect of the model is particularly important to this study because the change in storage may be a measurable quantity with remote sensing.

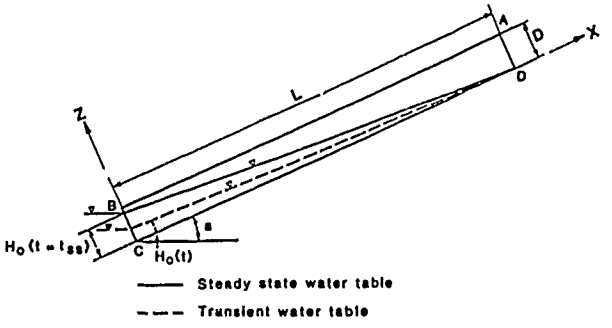


Fig. 3. Schematic of sloping slab model.

The Sloan and Moore [6] model assumes an idealized slope segment of length L, slope angle α, and an impermeable boundary at depth D as shown in Fig. 3. The mass balance is written as

$$\frac{S_2 - S_1}{t_2 - t_1} = (r+i-ET)L - \frac{q_1 + q_2}{2} \qquad (1)$$

where q is the discharge out of the slab, t is time, r is the rainfall input, ET is the evapotranspiration, i is the vertical flux from the unsaturated zone to the saturated zone, S is the drainable volume of water in the saturated zone.

$$q = H_o K_s \sin \alpha \qquad (2)$$

where K_s is the saturated conductivity. H_o, the saturated thickness normal to the hill slope, is computed by

$$H_{o_2} = \frac{\left[H_{o_1}(L\theta_d - K_s \sin\alpha\, \Delta t) + 2L(r+i-ET)\Delta t\right]}{(L_d + K_s \sin\alpha\, \Delta t)} \qquad (3)$$

where θ_d is the drainable porosity of the soil. The saturated region is recharged from the overlying unsaturated zone by

$$i = K(\theta) \qquad (4)$$

where K(θ) is the unsaturated hydraulic conductivity. The soil moisture θ then can be calculated by a mass balance of the unsaturated zone from

$$\theta_2 = \theta_1 U_u + L\Delta t\,(r-ET-i)/U_u \qquad (5)$$

where U_u is the total volume of the unsaturated zone.

5. MODEL SIMULATIONS

The Sloan and Moore [6] sloping slab model was used to simulate the eight day recession flow from watershed 1D. A detailed description of how the model was set up and how the various parameters chosen can be found in Engman et al. [2]. As initial conditions, the hydraulic head, H, was assumed equal to the soil depth, D. The initial unsaturated volumetric soil moisture content was taken to be 43 percent which was determined from the microwave measurements. The saturated volume was assumed to be 50 percent of the total slab volume. The time step Δt was taken as one hour and the data for ET were taken from Bowen ratio measurements.

The model simulation results are given for a unit width of slab, in this case, 1 mm. In order to evaluate the results, a total length of channel must be determined so that the simulated discharge can be compared to the discharge measured at the weir. The results shown in Fig. 4 are for a channel length of 600 m. This result also assumes that most of the drainage comes from the west side of the channel. This is based on the asymmetrical location of the channel in the basin.

Figure 4 shows the changes in soil moisture developed by the model and compares them to two points in the southern most ground sampling transect for soil moisture. It can be seen that the model results are quite good, being more or less in between the two gravimetric measurements. It can also be seen that the model responded to the low ET day (see June 2 data in Table 1) in the same way that the ground samples responded.

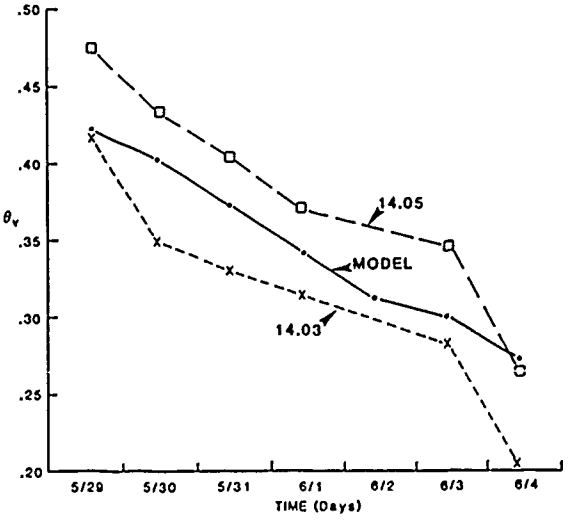


Fig. 4. Comparison of model simulated soil moisture and measured soil moisture at two locations on transect 14 (southern most transect).

6. SOIL MOISTURE AS MODEL INPUT

After examining the remotely sensed soil moisture maps and the good model simulations of soil moisture for the study period, one is tempted to speculate on how soil moisture could be used as direct input to the model to predict the base flow. To do this we first must examine the model and in particular Eqs. 1 and 5.

Equation 1 shows that the change in streamflow is related to the change in storage represented by the net flux to the saturated zone, which in turn is made up of the rainfall and recharge from the unsaturated zone and the ET loss. The rate of recharge, i, depends on the water balance of the unsaturated zone (Eq. 5). It is in Eq. 5 that one would like to use soil moisture as data. Rewriting this equation to solve for i gives

$$iL\Delta t = (r - ET)L\Delta t + (\theta_1 - \theta_2)U_u \qquad (6)$$

In this form it can be seen that i depends upon the change in soil moisture and the net balance between rainfall and ET. Thus, i cannot be estimated from soil moisture data alone but must also have an estimate of the ET for a rain-free period.

It was pointed out above that the model and the microwave measurements of the soil moisture both reflected the June 2 ET which was considerably lower than the other days' ET (Table 1). This being the case, one may be able to use estimates of ET based on simplified potential ET considerations [7] rather than local short term measurements requiring significant micrometeorological instrumentation. The consequences of this would be that for any one day (such as June 2), the ET may be overestimated and the model calculation of soil moisture would be lower than the actual. But subsequent measurements of soil moisture would show this discrepancy and allow correction so that the model predicted soil moisture would agree with the measured value. Thus, by using the remotely sensed soil moisture as feedback to the model, one may be able to run water balance models calibrated for a basin without the need of detailed on-site ET measurements.

7. DISCUSSION

Several researchers have suggested that evaporation or ET could be estimated from microwave measurements of soil moisture [8] and [9]. However, these studies have been limited to bare soil evaporation and to idealized lower boundary conditions of either a constant water table or a constant flux. These conditions cannot be met in this study.

The implications of this approach are that this type of model structure could be used to verify an intermediate step in the simulation process; that is, the unsaturated soil moisture balance during drainage and ET. Thus, a time series of soil moisture would provide an important check on the model performance and possibly provide feedback data so that the model could be corrected and updated if the simulations were under or over estimating the streamflow. This research shows a possible application of temporal soil moisture maps to a hydrologic problem.

8. REFERENCES

1. Sellers, P. J., F. G. Hall, G. Asrar, D. E. Strebel and R. E. Murphy, "The first ISLSCP field experiment (FIFE)," Bull. Am. Met. Soc., 69(1), 22-27, 1988.

2. Engman, E. T., G. Angus and W. P. Kustas, "Relationships between the hydrological balance of a small watershed and remotely sensed soil moisture," Proc. IASH Third Int. Assembly, Baltimore, MD, May 1989, IAHS Publ. No. 186, 1989.

3. Schmugge, T. J., P. E. O'Neill and J. R. Wang, "Passive microwave soil moisture research," IEEE Trans. Geosci. and Remote Sens., GE-24(1), 12-22, 1986.

4. Wang, J. R., J. C. Shiue, T. J. Schmugge and E. T. Engman, "Mapping surface soil moisture with L-band radiometric measurements. Remote Sens. Environ., 1989 (Accepted for publication)

5. Hewlett, J. D. and A. R. Hibbert, "Moisture and energy considerations within a sloping soil mass during drainage," J. Geophys. Res., 68(4), 1081-1087, 1963.

6. Sloan, P. G., and I. D. Moore, "Modeling subsurface stormflow on steeply sloping forested watersheds. Water Resour. Res., 20(12), 1815-1822, 1984.

7. Priestly, C. H. B. and R. J. Taylor, "On the assessment of surface heat flux and evaporation using large-scale parameters", Monthly Weath. Rev. 100, 81-82, 1983.

8. Bernard, R., M. Vauclin and D. Vidal-Majar, "Possible use of active microwave remote sensing data for prediction of regional evapotranspiration by numerical simulation of soil water movement in the unsaturated zone," Water Resources Research 17(6):1603-1610, 1981.

9. Prevot, L. R. Bernard, O. Taconet, D. Vidal-Majar and J. L. Thoney, "Evaporation from a bare soil using a water transfer model and remotely sensed surface soil moisture data," Water Resources Research 20(2):311-316, 1984.

SOLAR RADIATION STATISTICS FOR AUSTRALIAN CAPITAL REGIONS

DR M NUNEZ

DEPARTMENT OF GEOGRAPHY & ENVIRONMENTAL STUDIES

UNIVERSITY OF TASMANIA, AUSTRALIA

ABSTRACT

Data from the visible band of the Japanese Geostationary Meteorological Satellite (GMS) was collected for a two degree by two degree region surrounding each Australian State Capital. Data was collected for three years and was used, via a physical model, to derive regional global radiation statistics. Statistics are presented on a monthly basis for each region and consist of four images representing mean global radiation, standard deviation of global radiation, and number of overcast and cloudless days.

INTRODUCTION

Information on solar energy is needed in urban areas for a variety of applications involving solar hot water heating, photovoltaics, urban planning etc. Pyranometer networks are not sufficiently dense to resolve solar energy features at a regional scale (~ 200 km x 200 km). Cloud observations are also scarce so that attempts at modelling solar energy on a regional scale are also limited.

In this project data from the visible channel of GMS (0.55 to 0.75 μm) was used to estimate global solar radiation using a modelling scheme. Data covering a two degree by two degree region surrounding each of the eight Australian capitals was collected three times daily for a three year period (1/3/85-28/2/88). These data were used in a modelling scheme to estimate daily global radiation.

THE MODEL

The daily global radiation (K) can be given in terms of an extra-terrestrial global radiation (K_o), an absorptivity term (ϕ) and an atmospheric reflectivity term (α_A).

$$K = K_o(1 - \phi)(1 - \alpha_A) \qquad (1)$$

The model uses the satellite reflectivity (α'_{EA}) to derive α_A. From pyranometer observations, one located at each capital city airport, α_A can be derived:

$$\alpha_A = 1 - (K/K_o)/(1 - \phi) \qquad (2)$$

From Paltridge and Platt (1976), the earth-atmosphere albedo (α_{EA}) can be given as:

$$\alpha_{EA} = \alpha_A + (1 - \alpha_A)^2(1 - \phi)\,\alpha_G \qquad (3)$$

where α_G is the surface albedo.

The present technique derives a linear regression between the satellite reflectivity α_s and α_{EA} for the pixel containing the pyranometer station. In the operational mode α_{EA} is obtained from α_s, α_A is solved in equation 3(ϕ and α_G are calculated independently), and knowing ϕ and α_A, K is solved in equation 1. ϕ is solved using the technique of Lacis and Hansen (1974) as applied to precipitable water vapour from radiosonde data. α_s is obtained using the technique of Nunez et al (1987).

SOLAR RADIATION STATISTICS

Data is grouped on a monthly basis and consists of mean daily global radiation, standard deviation of of daily global radiation, and number of cloudy and cloudless days. The final step involves dewarping

the image to a Lambert conformal projection and printing each monthly statistic as a grey scale image. A total of 384 images were developed in the study. They have been compiled into an atlas of solar energy statistics (Nunez, 1988).

REFERENCES

Lacis, A.A. and Hansen, J.E., "A parameterisation for the absorption of solar radiation in the earth's atmosphere", J. Atmos. Sci., 21, pp118-133, 1974.

Nunez, M., "Satellite Estimation of Regional Solar Energy Statistics for Australian Capital Cities", NERDDC Report, 1988.

Nunez, M., Skirving, W.J. and Viney, N.R., "A technique for estimating regional surface albedos using geostationary satellite data", J. Climatol., 4, pp573-585, 1987.

OFICINA NACIONAL DE EVALUACION
DE RECURSOS NATURALES
ONERN

Calle Diecisiete N°355
Urb. El Palomar - San Isidro
APARTADO - 4992
CABLES "ONERN" LIMA - PERU

THE TECHNOLOGICAL TRANSFER ON REMOTE SENSING PROGRAMS, A TOOL FOR DEVELOPING COUNTRIES

Presented by :
* Walter Danjoy A.
to : 12th Canadian Symposium
on Remote Sensing
10 - 14 July, Vancouver-Canada

S U M M A R Y

This work show the antecedents and results of the technological transfer process that in matter of remote sensing is carrying out from Canada to Peru, through the execution of an International Technical Cooperation Program (PERCEP PROJECT).

The Technological Transfer on Remote Sensing Program from Canada to Peru, has already get install the technical and physical infrastructure for remote sensing develop and techniques applications. The applications results realized on natural resources and environmental field are varying and has determined the remote sensing use as a very important tool for National Develop.

In accordance to the obtained results during the last five years we have that the remote sensing in Peru has served mainly for cover the national territory on 100 % cartographycally, using satellite data, as well as, supply important information about the agricultural frontier extended, forecast and inventory crops and monitoring of the areas affected by environmental impacts. Likewise, with the remote sensing use has already undertaken ecological monitoring programs in the Amazonian Jungle, which goals is control the deforestation process that been increasing significantly. In the education field, that program has too multiplying effects, because the different universities are now including on its studies programs courses on remote sensing, and also the institutions offer training in specific filds.

This work in general, show the obtained attainments as results of the technological transfer process between Canada - Peru and its applications on the differents activities for the peruvian national develop.

---

*Walter Danjoy A. General Director on Remote Sensing, ONERN, Lima-Peru.

REMOTE SENSING TECHNOLOGY APPLICATIONS FOR THE CARIBBEAN: OVERCOMING ECONOMIC, POLITICAL & INSTITUTIONAL OBSTACLES

C. Specter
Florida International University, Miami, Florida 33199

D. Gayle
Florida International University, Miami, Florida 33199
Caribbean Institute of Social and
Economic Development, Trinidad and Jamaica

ABSTRACT

Remote sensing technology represents a cost efficient tool that can be applied to a host of problems confronting the developing world. However, the technology is not being fully utilized in these endeavors. The countries of the Caribbean represent one such group that could benefit substantially from the application of the technology. However, this group lags behind South America in its involvement in remote sensing activities. This leads us to consider those obstacles that hinder the transfer of available technology from the U.S. and other developed countries. The purpose of this research is to improve the flow of remote sensing technology to and through the Caribbean by developing a series of recommendations that may be useful to policy-makers and managers concerned with remote sensing applications in the region.

Keywords: Technology Transfer, Economics, Coastal Resources

1. INTRODUCTION

Recently, a global survey of remote sensing experts was undertaken to examine major issues in remote sensing technology transfer and applications in developing countries (Specter, 1986; Specter and Amann, 1987). Results revealed seven critical obstacles that must be overcome if the technology is to be successfully utilized within the development context. Certain problems are related to technical aspects of the system, such as the lack of computer equipment and the lack of access to data products, including remote sensing imagery and computer compatible tapes. Other key aspects of the situation extend beyond purely technical considerations. Lack of experienced personnel is a leading example. Organizational issues within the developing countries complicate matters further. Often there is a lack of cooperation among the institutions who could benefit from applications of the technology to their respective programs. These problems may be compounded by a lack of political support for the technology. To the extent that decision makers view current Earth observing satellite systems as untrustworthy in terms of discontinuation of services and/or obsolescence of ground support investments, this recalcitrance may be justified.

In the global study, economic constraints were perceived as the most significant. Related issues include the size of capital investment required, lack of foreign exchange, as well as the lack of funding from developed countries and/or multilateral agencies to support training programs and other remote sensing activities. If remote sensing technology is to be applied within developing country environments, each of these obstacles must be addressed and overcome. This is the real challenge that the community of remote sensing experts must face in the decades ahead. As Sir Hermann Bondi remarked in his keynote address at IGARSS '88 (Bondi, 1988):

> To whom does this knowledge belong? How can it be used in such a way that the benefits of remote sensing are equitably distributed and do not merely accrue to people working on a commodity exchange in a highly industrialized country, thus leaving the government where the plantation is located and the local people - as well as the makers and operators of the satellite - without any of the gains?... Of course, the technical aspects must be studied, improved and made to work together. However, the resolution of the other questions which are legal, institutional and political is equally essential if remote sensing is to thrive. These questions will need at least as much time to resolve as technical questions and do not seem to me at present to receive the urgent and substantial attention that they require.

This paper examines the application of remote sensing technology for coastal resources development and conservation in one developing region that could benefit substantially from such applications, the Caribbean. However, this region lags behind Latin America in its involvement in remote sensing activities (Bartolucci and Adrien, 1978). While the global study referred to above included 45 participants from Central and South America, only four individuals from the Caribbean region responded. This study offers an opportunity to determine to what extent the critical factors identified on a global basis are relevant to the Caribbean transfer experience. In an effort to understand the parameters of remote sensing technology transfer to the region, the research described below focuses our attention on three areas: first, the applicability of remote sensors to the resource development and conservation objectives of the Caribbean countries; second, the major obstacles in transfer and applications processes; and third, a series of recommendations that may prove useful to policy makers and managers concerned with improving the flow of remote sensing technology to and through the region.

2. METHODOLOGY

The methodology used in this research is that of a Delphi study. Delphi study techniques are particularly appropriate in integrating the attitudes and perceptions of relevant experts, especially in relation to the study of issues that necessarily involve a multidisciplinary approach. Experts, knowledgeable about the Caribbean region and remote sensing technology, were asked to participate in this project. These individuals were involved in technology transfer to the region, serving in the roles of policy-makers, managers of technology transfer programs, or advisors. Participants were drawn from recipient organizations in the Caribbean, as well as from donor organizations in the U.S. and Canada. Participants represented government agencies, scientific research organizations, universities, and private sector companies. Twenty-nine experts participated on this Delphi panel. Major topics relevant to remote sensing technology transfer to the Caribbean were covered. A detailed description of first round results was reported at IGARSS '88 (Specter and Gayle, 1988).

One major component of the Delphi instrument dealt with obstacles to technology transfer that impede the application of remote sensing technology to development activities in the Caribbean. This study offers an opportunity to determine to what extent the critical factors identified on a global basis are relevant to the Caribbean transfer experience and to ascertain what special obstacles were encountered in technology transfer projects to the region. To this end, the list of seven critical factors in remote sensing technology transfer was presented to the Delphi participants. They were asked to delete those factors that were not critical to the Caribbean transfer process and to suggest additional factors. Participants were encouraged to discuss their responses. In the second Delphi round, the experts rank-ordered the identified factors to determine their relative significance. The most important work of this round was to develop a series of recommendations that may be used to improve the flow of technology to the region and increase opportunities for successful applications.

3. OVERCOMING THE OBSTACLES TO TECHNOLOGY TRANSFER AND APPLICATIONS IN THE CARIBBEAN

Three of the listed obstacles related to inputs in the technology transfer system: the need for experienced personnel, appropriate hardware, and data inputs. A fourth factor was the role of relevant organizations in the region, including their capability to coordinate and cooperate in the technology transfer process. Another set of factors involved environmental issues: economic constraints, politics and policy/decision-makers, and the uncertain future of Earth-observing satellites. The experts suggested five additional factors: politics and policy/decision makers in developed countries, insufficient communication between scientists and the "political directorate," the small land mass area of any one country in the Caribbean, facilities and physical infrastructure, and the fact that many remote sensing technology projects are research-oriented and technically driven, rather than applications oriented.

A ranking of these factors in round two of this Caribbean study tended to substantiate the results of the earlier, global study concerning critical factors in the technology transfer and applications process. The Delphi participants identified two other problems they viewed as substantial: insufficient communication between scientists and the "political directorate," and the lack of sufficient facilities and physical infrastructure. A summary of the Delphi experts comments concerning the seven most critical problems and their recommendations for necessary actions to overcome them are summarized in Table 1.

4. CONCLUSIONS

This Delphi survey completed by a panel of experts indicates tremendous potential for the technology to be applied to environmental assessment of economic development in the Caribbean. The technology can be used to collect data in areas requiring priority attention: fisheries, coastal waters pollution, coastal erosion, mangrove conservation, and beach development for tourism. The results presented here replicate, to a substantial extent, the results of an earlier, global study. However, the results of this first Delphi round have led to the identification of regional considerations as well. The second round of the Delphi study was used to explore recommendations for improving the flow of remote sensing technology to and through the region. This information may be useful to concerned policy-makers and managers.

5. REFERENCES

Bartolucci, L.A. & Adrien, P.M., "Remote sensing technology and new development trends in Latin America", Proceedings of the Twelfth International Symposium on Remote Sensing of Environment, Ann Arbor, Environmental Research Institute of Michigan, pp. 187-194, 1978.

Bondi, Sir H., "Remote sensing: Moving toward the 21st century", Proceedings of IGARSS '88 Symposium, Edinburgh, pp. 5-7, 13-16 September 1988.

Specter, C., The Transfer of Remote Sensing Technology to Developing Countries: The Landsat Experience, 1970-1985, Ann Arbor, University Microfilms International, 1986.

Specter C. & Amann, R., "Critical factors in the transfer of U.S. space technology to developing countries", Academy of International Business Annual Meeting, Chicago, 12-15 November 1987.

Specter, C. & Gayle, D., "Technology transfer for development of coastal zone resources: Caribbean experts examine critical issues", Proceedings of IGARSS '88 Symposium, Edinburgh, pp. 1449-1552, 13-16 September 1988.

Table 1
A Summary of Major Obstacles and Recommendations for Overcoming Them

CRITICAL FACTOR	RECOMMENDATIONS
Experienced Personnel	-- A 'critical mass' of experienced personnel is needed; people transfer, change jobs, etc. One or two experienced personnel per project is not enough. -- Experienced personnel must be assigned to assist in the development of projects and the training of local personnel. -- Sharing of experienced personnel within the region should be considered. -- More training needs to be provided in all related fields, including remote sensing, GIS, computer science, and other disciplines through which the technology is brought to bear. -- Training programs must be formulated to provide both initial training and follow-up training, so that the "somewhat knowledgeable people" will eventually become very knowledgeable people. -- Local resources are not fully being utilized in education/training efforts; universities and high schools should provide ad hoc courses. -- Some training may be needed from the developed nations: USA, Canada, Europe and Japan. Possibly, post-graduate education may be part of the training to create experienced personnel. -- Experienced personnel should also be able to effect minor maintenance procedures and repairs.
Computer Equipment	-- Micro-based systems should definitely be considered. They can be shared with other departments, as they are not dedicated to remote sensing. A micro-based image or GIS system could provide much more information at less cost and more flexibility. (This system could be used for other administrative and technical jobs thus justifying its purchase). The size of the countries means less data is needed per data set. A big system is not necessary. -- In the Caribbean, large area coverage is not essential for most applications, thus a PC-based image analysis system may be the answer. -- A PC-based system will be useful in the early pilot programs. A more sophisticated VAX-based system can be introduced after on the success of the PC-based image analysis system is established. -- It is better to assemble a reliable system that provides a multiplicity of use modified for the local conditions than to purchase a whole system. -- Harmonization and compatability for inter-linking is the key. -- Computer equipment complexity must be matched by user training/experience. Skilled and experienced personnel are needed not only for data analysis and interpretation, but also for the maintenance of equipment during the operational phase of projects. -- Purchase durable equipment: this is a hot saline environment with a history of power fluctuations. -- The operations and maintenance concept is not ingrained in the Caribbean culture. This concept needs to be instilled for equipment upkeep.
Data Acquisition	-- Since it is often impossible to obtain cloud-free images in the region, more importance should be given to Synthetic Aperture Radar (SAR) and also to equatorial orbits; data acquisition should be more responsive to the needs of those trying to apply the technology in developing regions. -- Remote sensing has shown the greatest benefit when used in the monitoring mode. This implies that projects need to accommodate the acquisition and processing of not 1 or 2, but perhaps 10 or more images. -- It is essential that each country have complete coverage of large scale air photos as a base for interpretation of other types of remote sensing data.
Distribution	-- The regional information to date needs to be documented and analyzed to serve as a guide for future undertakings. -- Remote sensing data should be available to those who really need to know, e.g., farmers, cooperatives, artisanal fishermen, and not just to researchers and policy makers.
Software	-- More software for micro computers needs to be developed. -- Some organizations are now providing value-added products which should be considered before investing in a new data processing system.

Table 1 (Continued)

CRITICAL FACTOR	RECOMMENDATIONS
Economics	-- Good projects will be financed. In developing countries, projects have to compete for funding. Benefit/cost analysis ratios, as well as other economic indicators need to be analyzed to show whether the required instruments can be justified; one needs to make the case that remote sensing can acquire needed data cost effectively (as opposted to traditional surveys). -- The economic problem might be alleviated if regional receiving, processing and training facilities were established, especially in the Central American and Caribbean countries. For the smaller, insular states, groups of nations can initially subsidize one main processing facility. Additional national facilities, thereafter, could be added as the benefits become more apparent and the economic status of a particular state changes. Maintenance costs could be tied into the capital budget of a user nation. Experienced personnel could be shared initially when only a few facilities are available. -- Due to the lack of foreign exchange in some Caribbean countries, external funding should be made available; The World Bank, Canadian International Development Agency (CIDA), USAID, and UNESCO are some agencies that can facilitate the financing for pilot programs to build up local infrastructure. We need to fund an"ad hoc" committee of both Caribbean scientists and resource managers with selective members of the international remote sensing community to determine areas of remote sensing technology transfer. -- Equipment/training could be provided pro bono by donors.
Politics in Caribbean Countries	-- Recognition for RS might be obtained if the earliest projects were based on the most politically sensitive issues in a country. Political backing is more important than economics: it is the bottom line. There is always money to be had if the priorities dictate so. -- An extremely critical factor is government agencies' acquisition/use of such material. This will almost always depend on the politician. Therefore, it is necessary to demonstrate to politicians the multiple-use and potential of the technology in order for them to be motivated to provided funds for remote sensing projects and facilities. -- Remote sensing agencies, user groups, institutions need to learn to market/expose their existence and capabilities to government and schools. -- Programs need to be developed to educate Caribbean decision makers in the potential and relevance of remote sensing to regional development. -- Politics and policy should encourage remote sensing technology in resource management. -- It is important to formulate sound policies in natural resources, development, and information networking.
Insufficient Communication Between Scientists and the Political Directorate	-- A "team effort of these two groups is needed. Scientists should get involved with people who need to know. -- Universities should be involved. Ideas for better communication include forming an "ad hoc" task force to which politicians may be more willing to listen. -- Specific projects employing remote sensing technology could be made flexible enough to provide limited aid to agencies engaged in natural resources, management, planning, etc. in an effort to demonstrate the applicability of the technology and how it can help in the development of the country. -- There should be a greater diffusion of possible applications (less academic) so that political administrators can better evaluate the advantages of this technology.
Facilities and Physical Infrastructure	-- One solution would involve regional facilities for receiving and processing. Use a PC-based "low-cost" system in the first stage of the pilot program. However, the development of an equipment infrastructure should be furthered in order to better utilize the systems in place and depend less on outside sources. -- Secure a steady source of electricity; reduce power fluctuations and failures. -- Justify, then obtain the use of PCs and peripherals for administrative purposes. The add-ons to turn these PCs to image analysis systems are then much less prohibitive to purchase. -- Facilities may be acquired as part of external funding packages for projects.

THE IMPORTANCE OF REMOTE SENSING AND MAPPING FOR RESOURCE MANAGEMENT

A CASE STUDY OF NEPAL

J.E. Ambrose and P.B. Shah

CIDA Advisor and Chief, respectively, Integrated Surveys Section, Topographical Surveys Branch, Government of Nepal

Abstract

The Canadian International Development Agency has sponsored, in conjunction with the Government of Nepal, a major resource mapping program in Nepal. The basis of this program was panchromatic aerial photography supplemented by Landsat false colour infrared imagery. The program resulted in the mapping at a scale of 1:50000, land systems, land use and land capability as well as geology at 1:125000, for the entire country. The applications of the mapped informtion have been many and varied and have assisted government agencies in arriving at sound resource management decisions. Such a program must include a component of guidance to make potential users aware of the value of the information and its correct interpretation and application.

Key words: Nepal, resources, land systems, management

Background

During the mid 1970's Nepal recognized the need for basic resource information for purposes of resource decision making. The Canadian International Development Agency (CIDA) responded to the country's request to provide aerial photo coverage and basic resource mapping of the entire country at a scale of 1:50000. This mapping included geology, land systems, land use and land capability. The Land Resources Mapping Program (LRMP) required five years to complete, resulting in the production of 266 map sheets at 1:50000 scale for the above resource information sets with the exception of geology which was mapped at 1:125000.

Nepal is a relatively poor country with an average annual per capita income of U.S.$160. It is a country whose population is rising at the rate of 2.6% and currently holds a population of 17.1 million. The pressures placed upon the dwindling resource base of the country are mounting and as a result Nepal is facing all of the environmental problems to which mountain systems are prone.

The government recognizes the seriousness of the situation and accordingly has recently completed a comprehensive National Conservation Strategy which is about to enter into the implementation phase; has introduced the Decentralization Act which places planning responsibility where it belongs, in the districts; and has recognized, over ten years ago as mentioned in the above, the need for resource information for making effective planning decisions.

Land Resource Mapping Project (LRMP)

The LRMP was established to provide the Government of Nepal with the necessary information upon which to base its resource management decisions. The exact types of information to be gathered were decided upon through a broad analysis of the government's particular needs, the available systems already in place elsewhere and the time and budget available for the exercise. As well, the nature of the terrain and the working environment of Nepal were taken into consideration.

- remotely sensed information - a foundation

Due to the size of the country, its' terrain and the lack of access, the basis of the program would have to be remotely sensed data. Colour infrared Landsat imagery at 1:250000 was used for obtaining an overall impression of the country's landscape. The major physiographic breaks were differentiated as well gross changes in vegetation and land use pattern. As well, this imagery was used for the interpretation of landscapes above 15000 feet in elevation.

However, the most useful medium for geological mapping as well as for land systems, land unit and land use mapping would be panchromatic aerial photography at a scale of 1:50000. This medium was selected for the majority of the interpretation for a number of reasons including superior resolution, particularly when used in the field, higher accuracy, especially in the

middle hills, the high mountains and the high Himal where distortion presents a serious problem with interpretation and measurement, and the wider practical use that the photographs would have later by a number of government agencies. As the medium is very portable, interpretation is relatively easier for the field investigator, and as well, cost advantages and the unnecessary need for additional sophisticated equipment in the form of image lyzers which require specialty training and care determined the use of this medium over others. In other words, it was found that the standard black and white aerial photography for carrying out the major portion of the work was the most practical tool available.

Aerial photography was available at 1:20000 for only select areas but this was useful for land unit separations and as a backup for the interpretation of smaller scale photography. Later in the project it proved useful for village level management planning when enlarged to 1:5000.

- the geological base

Recognizing that many of the resource decisions to be made in Nepal depend so much on the nature of geological structures and the rock types from which soils are formed, geological investigations and related mapping were carried out to provide the basis for a comprehensive land system analysis of the entire country. The Main Boundary Fault and related faults were mapped along with the Main Central Thrust. As well, other geological features were also indicated. This was all carried out with the aid of the aerial photographs and some ground truthing. Rock types, which form the basis of the land system component of the exercise, were mapped through extensive field sampling and photo - geological interpretation.

- land systems

The main defining criteria for the land system mapping was based upon the country's physiography. The boundaries of the country's five physiographic regions were refined using 1:250000 false colour infrared Landsat imagery in combination with the black and white photography. These regions, defined by their geological charactistics as well as by climate and vegetation, were then used as the framework for the design of the land system.

With basic geological information combined with topographic information available from the national topographical series, and information on soil types and depths, the country's landscapes were then categorized into a land system. The system contains seventeen different homogeneous units based upon recurrent patterns of landforms, geology, slopes and agricultural limits. The system is further broken down into smaller land units which are characterised by topographic position, slope, surface dissection, flooding frequency and soil characteristics such as drainage, depth, texture and profile development.

It was felt that the land system would be useful as a base for carrying out a wide range of resource analysis. Since the major land use activities in the country centre on agriculture and forestry, the land system was used for analyzing and mapping out the country's capability for agricultural and forestry production.

- land capability

Due to the unavailability of other biophysical information and various constraints of the program the land capability classification was based upon limited information which included soil depth, texture, slope, and to a certain degree, present land use. The classification system contains seven classes of which only the first three classes are suitable for agricultural production. The next three classes indicate potential for forestry production/grazing and the last class, covering high elevation areas of rock and snow, has no potential for biological production.

The classification also includes a suitability rating scheme for irrigation. This rating has been applied to all lands which fall within the first two capability classes.

- land use

The fourth component of the program to be analyzed and mapped was current land use. Land use patterns analyzed included agricultural systems, forest cover including forest crown classification, maturity and stand condition, and grazing lands and shrub areas. The 1:50000 aerial photographs were used extensively to delineate both land use patterns and forest stand classification down to a minimal size on the ground of 25 hectares.

Output

The project concluded with the production of 266 topographic maps at a scale of 1:50000 for each of land systems, land capability and land use, and the corresponding number of map sheets at a scale of 1:125000 for geology. As well, a number of reports covering the various aspects of the project were published.

Applications

The LRMP data has a large number of potential uses and users and considering the scale at which the data has been mapped, the data is suitable for resource planning purposes at the national, regional and district levels. Below the district level the data is still useful but would have to be supplemented with more detailed studies in order to obtain information at scales

varying from 1:5000 to 1:25000.

The data has been used by a number of agencies within the Government of Nepal as well as by several donors and international organizations in connection with various development projects within the country. Some examples of the use of the LRMP data are described in the following.

National forestry master plan

The forestry statistics by administrative areas available through the LRMP have been used as the basis for determining long term forestry policy and priority areas for forestry development programs.

settlement

LRMP information has been used by the Ministry of Settlement to determine the most suitable areas in the Teral zone for the resettlment of hill populations.

hydro - electric power

The Karnali power scheme program has used the LRMP data for determining flood areas and the loss of various types and classes of land within the flood areas.

livestock

The Ministry of Agriculture has utilized both the statistical information and mapped references of grazing areas in the High Himal zone for purposes of determining stock carrying capacities and pasture management programs.

horticulture

Various donors, through the Ministry of Agriculture, have utilized the LRMP data in conjunction with more detailed studies to determine the most suitable areas in the country for specialty horticultural crops such as oranges and tea.

road alignment

The different donors involved in the major east - west road across Nepal and the north - south highway which provides access to India have used the LRMP data extensively for assisting in the determination of optimum road alignment.

National parks

The Department of National Parks and Wildlife Conservation is currently using the LRMP data as input into the preparation of a management plan for Langtang National Park and in cooperation with the ISS is experimenting with the use of the ISS's GIS for assisting in the analysis of the park's resources for management prescription purposes.

other

The Department of Irrigation is using the LRMP irrigation suitability ratings as input to the preparation of the national irrigation master plan; the Department of Tourism is following very closely the further analysis of LRMP data that ISS is carrying out for purposes of deriving a tourism attractivity index and in this regard one donor supported rural development program is interested in an application of the index to its geographical area of concern; the National Planning Commission has shown an interest in the LRMP data and now insists that all planning at the local level take into account the information available through the LRMP data bank.

Current Activities

The Canadian International Development Agency (CIDA) recognizes the importance of remote sensing and the use of this valuable tool in providing developing countries with the basic resource knowledge that they require in order to make sound resource decisions. CIDA also recognizes the importance of ensuring that the recipient of the program is aware of how the information can be effectively utilized and in this regard CIDA has extended the LRMP a further two years to include a planning advisory service that will demonstrate to the various resource agencies within government how the information can be interpreted and applied to their various respective programs. Included in this phase is the establishment of a GIS which will facilitate the analysis of LRMP data for various resource planning purposes as well as provide an excellent medium for demonstrating the application of LRMP data to resource management.

Some Lessons Learned

During the life of the LRMP there have been a number of other attempts in Nepal to utilize various forms of remote sensing including sales pressure from donors and commercial firms to utilize landsat imagery and SPOT. With the exceptions described above of applying Landsat, the LRMP has avoided the application of these technologies. Remote sensing availability to Nepal will always come through the donor community simply because the country will never be in a position to afford the required purchases on its own. It is the moral responsibility of the donor agency to ensure that the technology being offered is an appropriate one. The donor must satisfy

itself that the country has the capable people who can interpret the results of remotely sensed information, that the medium will meet the particular needs of the client and that follow - up activities such as training and maintenance are not beyond the ability of the recipient after the donor has left the program.

In the case of the LRMP, standard black and white aerial photography was found to be the most suitable and cost effective means of obtaining the data required. Attempts to demonstrate the utility and economy of other forms of remote sensing have failed.

With regard to the data and information products of the project it is important that the users and potential users of the products have a thorough understanding of the possible applications of such products to their respective programs and that they understand the constraints and limitations when contemplating the use of these products. The Integrated Survey Section has discovered that maps which appear quite straight forward to those who were responsible for their production are quite often misinterpreted and misused by the client. A common misuse is the enlarging of a particular map in order that the information contained within the map can be applied to a large scale activity (e.g. enlarging 1:50000 land use maps and applying the information for village level planning at 1:10000). It is important that such resource mapping programs contain an element of user training and guidance.

Conclusion

A very successful land information system has been implemented in Nepal through the funding of the Canadian International Development Agency and the foundation of the system has been the use of what today would be considered an old tool - standard black and white aerial photography at a scale of 1:50000. This tool has proven to be more useful and economical than more advanced methods of remote sensing, particularly satellite imagery. The system developed has wide spread use throughout Nepal and is being used in a number of areas for arriving at resource management decisions, however, the potential user must be given some guidance in the application of the information.

Donors in the future must be aware of the specific needs of the client and ensure that the technology that they are offering is appropriate. In the developing world it is not always the leading technology of the west that is best.

Referenceś

1. HMG Nepal / Government of Canada, Land Systems Report, the Soil Landscapes of Nepal, Land Resource Mapping Project, 1986.

SUSTAINABLE DEVELOPMENT IN RESOURCE MANAGEMENT THROUGH REMOTE SENSING

JOSEPH EPHSON OTOO

DEPT. OF GEOGRAPHY UNIVERSITY OF BRITISH COLUMBIA

ABSTRACT

The use of remote sensing in many developing countries continues to attract a lot of attention even though the technology is still " remote " in terms of its application. Several studies using remote sensing in Africa have created a gap between the information yielded and what the resource manager can effectively use. In order to bridge the gap between information obtained from remote sensing and what the resource manager needs, answers to the following questions must be sought. what type of information can be obtained at what level of accuracy? Who needs a particular type of information derived from remotely sensed data? A resource simulation model can be used to evaluate some of the parameters involved in the management of forests in Africa, in examining the case for its sustainable development. Remote sensing data will provide the data base and analysed on these and other parameters to measure the level of accuracy and relevance to any resource management technique.

Key words: Africa, resource, sustainable development, remote sensing.

INTRODUCTION

With population sometimes doubling in a matter of 25 years, as in the case of Kenya, Ghana and Nigeria, many governments have realised that resources are not inexhaustible. As rural conditions deteriorate, attention has been given to the use of forests for energy supply and the deforested land for farming. This, in conjunction with severe climatic and physical conditions, has accelerated forest depletion in many parts of the continent. The concept of sustainable development - even though too broad and ill-defined in its generic sense - applies very well to the African situation. It proposes balanced economic development to meet the basic needs of people without causing severe damage to the environment on which the resources are based. As noted in the United Nations Economic Commission for Africa Report, 1988, the results of resource **mismanagement** are desertification, deforestation and pollution.

Some of the issues to be raised in this paper include:

- How severe is resource mismanagement?
- How do we identify and measure resource mismanagement?
- What amount of information on the resource do we possess now, and how accurate is it?
- Can we use remote sensing data to generate information on the resource, given the variability of spatio-temporal characteristics that influence satellite data?
- Can the information derived from remote sensing be used to generate a model for empirical analysis of the resource in its present and future states?
- Can such a model be applied over different spatial and temporal horizons?

Much has been documented in the literature on the severity of resource degradation (ECA, 1988). Usually, in many of these studies, the question asked is - what rate should societies develop to meet basic needs of the citizenry, without necessarily undermining the natural resource base and the environment on which they depend. In its 1987 report, the World Commission on Environment and Development, (WCED, 1987) concluded

"... the time has come to break out of past patterns. Attempts to maintain social and ecological stability through old approaches will increase instability..." This will form the basis of the paper.

2. THE FOREST RESOURCE IN GHANA

In Ghana, forests occupied some 8.2m hectares at the turn of the century but this had systematically been reduced to 1.7m ha. in mid 1987. A valuable resource, forests are cleared for agriculture (commercial and small-scale), logging and fuelwood; they provide abode for wildlife. Indeed, forests, through their genetic resources, have contributed to the growth in modern medicine and industry, yet this is the resource that we understand little about - not even the most basic physio-biotic relationships, community dynamics and energy flows (Myers, 1980). Given the Government's intention to revamp the timber industry, production has been progressing rapidly since 1983. On the other hand the national forest potential is on the decline. Until now it

has been very difficult to monitor private timber companies with respect to the technology used and amount logged, and also their compliance with reforestation regulations.

Various descriptions and classifications of forests have been given in the literature particularly those by UNESCO (1978; Whitmore (1975). I propose to work with a more simplified definition of Ghana forests - " the evergreen (and occasionally semi-deciduous but never completely leafless) forests exhibiting several distinctive strata" (based on Myers, 1980). This definition seeks to divide Ghana into three major forest types:

- the true evergreen forest with annual rainfall of 2500mm
- the intermediate evergreen forest type with rainfall of 1500m
- the semi-deciduous forest that fringes the savanna zone in Northern Ghana with variable rainfall between 750- 1500m.

These classes of forests are not mutually exclusive nor exhaustive since it is likely to find a mixture of these classes in many parts of the country, usually as a result of cultural and physical factors acting over a long period of time. Indeed, the variability in the nature of the resource becomes evident as one travels through the country.

Owing to the extremely complex and diverse nature of Ghana forests, I would like to obtain information and work with the following parameters of forests:

- productivity - measured in terms of regeneration and success of growth rates in the supply of forest products;
- volume - measured in terms of biomass generation and stand quantity;
- types and species identification;
- land use types and other cultural activities and changes;
- infrastructural framework.

The ultimate aim of studying these characteristics will be to estimate within acceptable limits, the aerial extent of forests, the rate of growth and supply in meeting the present and projected demands. On the basis of such information, it should be possible to predict the rate of deforestation and assess its impact on the people, the soil, and wildlife, and the micro-climatological conditions of the area.

The problems of analysing, interpreting and ultimately synthesising the vast quantities of scientific and human knowledge relevant to the sustainable development of the biosphere are daunting. A basic problem that has affected inventory of forest in Africa, and indeed elsewhere in the developing world, is the lack of comprehensive data. The existing information comprises mainly estimates rather than actual surveys, and is fraught with guesses and imprecision. Figures are never agreed upon by independent workers and government sources. Given current research trends, remote sensing should provide a database for the inventory of the resource.

It is against this background of extremely complex issues that the present study seeks to address the question of sustainable development of forest resources.

3. DEVELOPMENT OF A SIMULATION MODEL

A resource simulation model was designed to mimic changes in the forest ecosystem including environmental, social and economic changes. One main advantage of this technique is its flexibility. It must be noted that the forest ecosystem is complex and management practices imposed on it must be within the framework of a resource that is well understood. One requirement of this type of model is that the relevant variables and their relationships must be clearly identified and specified. However, if care is not taken this type of model could grow to a complexity that can exceed the real world it seeks to mimic.

Some writers, for instance Brewer (n.d), are of the opinion that large-scale modelling for decisions and policy making has not fulfilled its promises. He thinks some of the problems involved in the use of models can be traced to basic confusion that appears among those who commission, build and use the tools. Closely related to this is the more fundamental misunderstanding between policy and scientific work.

The key components of a simulation model should identify and classify the elements in the calculated sequence as:

- variables - the things which the model seek to predict;
- parameters - constants such as growth rates which are necessary in the predictions;
- equations - shorthand sentences which state how variables and parameters are related and state the basic assumptions of the calculation;
- driving variables - eg. the harvesting rates that we want to manipulate or vary over time but that are not to be predicted within the basic calculation sequence.

This resource simulation model will be designed as a tool to explore the effects of the current harvesting rates (defined in terms of logging, for agricultural purposes and for energy requirements) on the current forest stock in the country, and determine if this is likely to bring a total depletion of the resource, and in what time frame? The model results provide an objective estimate of the rate of harvesting that can be sustained. Some of the questions to be answered in the use of the model are what are the current effects of reforestation and natural growth on the resource? what will be the impact on the resource if the reforestation and harvesting rates are varied? Then, also, the model will be used to predict the state of the resource over a period of 50 years. A simulation model was used by Daniel and van de Wetering (n.d.) to examine the effects of different harvesting and silvicultural policies on the long term sustainable yield of the British Columbia forests. Using four major policy options:

1. current (forest) policy;

2. increasing the intensity of silviculture;

3. decreasing the harvesting rate;

4. combining increased silviculture with a decrease in harvest.

The model indicated that a dramatic falldown can be expected if the last option is not taken.

4. ROLE OF REMOTE SENSING

Remote sensing will provide the input data for the model. As well, the results obtained from remote sensing data will be used to validate the previously estimated figures of the resource. Remote sensing data will also be fed into the model in order to compare the results of the models based on remotely sensed data and published data. Similarly, given the limitations of remote sensing, it will be necessary to account for all the spurious variability associated with satellite data to evaluate how reliable and accurate the data is in measuring the parameters used in the model.

To begin with, it will be necessary to identify the type of sensor that will be suitable in providing data at the temporal and spatial resolution required in this work. The availability of multisensors (MSS, TM, SPOT, METEOSAT, NOAA'S AVHRR etc.) with their multitemporal characteristics offers the user a wide range to choose in acquiring the information needed.

Tucker et al., (1985) used the AVHRR of the NOAA in land cover classification for Africa; the coarse resolution of the AVHRR may not permit the detection and mapping of small-scale and sporadic activities like fuelwood consumption or bush fires going on in Ghana because of the scale. In combination with ground truth information, an 11-channel multispectral scanner was used to obtain data on the actual state of forest lands in S.W. Germany by Kadro and Kuntz (1986). Also, Daus et al. (1986) derived an inventory of fuelwood volumes in Niger, using Landsat imagery and medium scale aerial photography. Many other studies have estimated the extent and effects of fires on forests, climatic and atmospheric conditions (Maston et al., 1987). The broad range of data available from Landsat Multispectral Scanner (MSS), Thematic Mapper (TM) and SPOT can provide the data for analysis in the model. These sensors should be evaluated to determine which bands best provide the most desirable information on vegetation parameters that would enable the estimation of type, productivity, deforestation and regeneration in the model. Other factors to take into account are the occurrence of frequent cloud cover and the number of overpasses of the sensor in this region of Ghana which is necessary for the monitoring of the resource.

To assess the impact of current and projected deforestation rates on the social milieux, it will be necessary to include population growth, farming practices like shifting cultivation, grazing, bush firing etc. into the model. This work seeks to evaluate how much of the model can be based on remote sensing and at what level of spatial and temporal resolution. Similarly, it seeks to analyse which parameters included in the model cannot be estimated currently by remote sensing means, but the potential is there.

5. LIMITATIONS OF DATA SOURCES

Owing to paucity of data on the resources in Ghana, it was only feasible to work with F.A.O published data for 1980 and projected estimates for 1985. These data were used against the background of the inaccuracies and inconsistencies noted in the FAO publication. It is assumed that a more rigorous model can be developed when more reliable and recent data becomes available.

6. DISCUSSION AND CONCLUSION

This work has attempted to illustrate how remote sensing data can be used to generate a simulation model for resource management over time. As well, data from satellite imagery were used to validate other published statistics and in turn determine what accuracy limits should be expected from satellite data to be useful for resource management. It is not enough to generate maps based on vegetation indices; perhaps, it will be necessary to take a step further in using these indices in generating and simulating some of the resources which are almost in a state of extinction. It is the intention of the author to build on this work by adding other variables into the model such as slope, aspect, soil type and drainage characteristics (whichever can be derived from remote sensing), to generate additional information on the productive behaviour of the resource.

ACKNOWLEDGEMENT:

I would to thank Profs. Graham Thomas, Brian Klinkenberg and Carl Walters for their kind assistance in reviewing the paper and offering useful suggestions. Thanks also go to Geoff Cushon for his computing assistance.

REFERENCES

1. Myers, N., 1980, The present status and future of tropical moist forests, Environmental Conservation, vol. 7, no. 2.
2. Whitmore, T.C., 1975, Tropical Rain Forest of the Far East, Clarendon Press, Oxford, UK.
3. Tucker et al., (1985), African Land Cover Classification Using Satellite Data, SCIENCE, vol. 227, no. 4685.
4. Daus et al., 1986, A remote sensing aided inventory of fuelwood volumes in the Sahel region of W. Africa - a case study of five urban zones in the Rep. of Niger, In Remote Sensing for Resources Dev't and Environmental Management, vol 1, Rotterdam, Netherlands.
5. Kadro, A. and Kuntz S., 1986, Experiences in application of multispectral scanner-data for forest damage inventory, In Remote Sensing for Resource Dev't and Environmental Management, vol 1, Rotterdam, Netherlands.
6. Brewer, G.D., (n.d.), Methods for synthesis: policy exercises.
7. Daniel, C. and van de Wetering, D., (n.d.), The effects of forest policies in British Columbia on the long term sustained yield.
8. UNESCO, (1978), Tropical Forest Ecosystems: A statement of knowledge Report, Natural Resources Research XIV, Paris, France

9. Matson, M. et al., 1987, Fire detection using data from the NOAA-satellites, Int. J. Remote Sensing, vol. 8, no. 7.
10. FAO, (1986) FAO Production yearbook, Rome, Italy.
11. W.C.E.D., (1987) World Commission on Environment and Development, Final Report. 12.
E.C.A., Economic Report on Africa, 1988, Addis Ababa, Ethiopia.

USE OF REMOTE SENSING SATELLITE DATA FOR INVESTIGATING WATER RESOURCES IN AFRICA

A. Pietroniro, W. Wishart, and S.I. Solomon*

Department of Civil Engineering, University of Waterloo,
Waterloo, Ontario N2L 3G1, Canada
Tel:(519)885-1211 Telex:069-55259 Fax:(519)888-6197

* MacLaren Engineers Inc., 320 Adelaide St. S., London,
Ontario N5Z 3L2, Canada

Abstract- Investigation to alleviate the consequences of prolonged and frequent drought periods in the Sahel areas of West Africa are hampered by the lack of data on the water resources Conventional techniques of investigating these resources in the Sahel are expensive and difficult to implement given the particular conditions of the area. Remote sensing by satellite coupled with ground truth data from the past and obtained currently at a limited number of gauging stations may provide a viable alternative to the conventional techniques For this purpose, a simple rainfall-runoff model, whose parameters have clearly defined physical meaning was developed calibrated and validated A technique for estimating monthly rainfall values by using Meteosat data was also developed and tested By using the satellite derived rainfall data as input into the rainfall runoff model, all the components of the hydrologic cycle were estimated and the estimates verified using measured runoff data The results indi cate that the application of these techniques yield satisfactory estimates at least with respect to surface runoff

Keywords – Rainfall, Runoff, Model, Meteosat, Africa.

INTRODUCTION

As is well known, the Sahel area of West Africa has experienced prolonged and frequent drought during the last two decades As the periods followed a period of relatively abundant rainfall and significant improvement in sanitary conditions, the consequences of the drought have been particularly disastrous. A large number of national, regional and international organizations have been investigating the problem associated with Sahel drought with the purpose of developing action plans aimed at alleviating the consequences. A significant difficulty encountered by all these organizations is the scarcity of data on the natural resources of the area, in particular the water resources. Whereas the larger rivers have been generally gauged, occasionally for an extended period of time, the gauging of the smaller rivers has been to a large extent neglected. Future plans involving the use of surface resources within the area, in particular for irrigation purposes, should consider the use of the smaller rivers, as these are likely to be close to the demand areas and suffer much less losses due to evapotranspiration and infiltration than the larger rivers Inferring the hydrological regime of smaller rivers from that of larger rivers is not possible because of the large unaccountable and often variable water uses and because of the existence of numerous negative tributaries. The problem is further complicated by the fact that changes in land use in the area result in non stationary hydrologic regimes. These circumstances make the use of conventional hydrologic analysis and synthesis for the Sahel impractical.

An investigation into the hydrologic regime of the Sudano-Sahelian zone of West Africa carried out in the mid seventies (Davy, Mattei and Solomon, 1976) indicated that precipitation is related to runoff by a relationship which varies mainly as a function of river basin vegetation. This relationship suggests that it might be possible to develop a conceptual (deterministic) rainfall - runoff model with parameters varying precisely with the river basin vegetation. Such a model could be used to synthesize runoff time series from a corresponding time series of precipitation, to explain past non-stationarities related to changes in river basin vegetation and predict the future hydrologic regime assuming future changes in the vegetation cover. Since vegetation variation in time and space can be ideally monitored by means of satellite remote sensing, the parameters of the model could be readily estimated using remote sensing data. Furthermore, since precipitation variation in time and space can also be estimated from satellite remotely sensed data, the model could be used to estimate the variation in time and space of runoff and also other components of the hydrologic cycle. It was therefore considered of interest to proceed with research to assess the feasibility of the above described methodology for estimating the time - space variation of the hydrologic cycle in small river basins in the Sahel.

THE CONCEPTUAL RAINFALL-RUNOFF MODEL

Many hydrologic models in use today require extensive historical runoff data as well as detailed physiographic data before adequate calibration can be carried out. This data, however, is not always available in developing areas. In addition, these models also have a large number of parameters. These parameters are sometimes not based on physical phenomena but are introduced to provide calibration opportunities. Several researches (Ragan and Jackson, 1975, Link, 1980) have shown the potential for using remote sensing to estimate parameters for hydrologic modelling.

This phase of the research has as its primary goal, the development of a simple and rational hydrologic model that has the potential of using remotely sensed data for the estimation of the model parameters. The hydrologic model developed for this investigation is of a parametric type and based on a mass balance approach and a monthly time step. The model postulates that moisture enters the basin as precipitation and either percolates to groundwater, becomes runoff or is used for evapotranspiration. It assumed that there are no other sources of moisture available

within a river basin.

The model has four parameters, soil moisture capacity, percolation capacity, subsurface flow coefficient and a soil evapotranspiration coefficient. The soil moisture capacity (SMC) is the maximum possible amount of moisture contained within a soil profile (mm). Percolation capacity (PERK) is the maximum amount of water that can percolate to deep groundwater in any given month and the subsurface flow coefficient (SSFC) represents a proportion of the soil moisture that can flow to rivers through seepage or subsurface flow. In many arid regions, this is the only flow that occurs. The soil evapotranspiration coefficient (β) is a proportion of the soil moisture that is available for evapotranspiration. Evapotranspiration is estimated using a modified version of a formula described by Turc(1956).

For the purpose of calibration, a subroutine called CLIMB (Rosenbrock, 1960), based on the method of steepest gradients for optimization, was used to speed the optimization of model parameters. The four model parameters were calibrated using the meteorolgical and hydrological data for 57 small (less than 100 sq.km.) river basins in the west and equatorial regions of Africa. Optimization of parameters was done by comparing historical runoff data with those estimated by the model and then minimizing the standard error of measurement. Timing of peak runoff and dry periods was also examined. The optimized parameter values were then grouped according to hydrologic characteristics such as soil type, texture, slope, vegetation and lithology.

Data for the 57 small river basins were available for between 24 month (41 river basins) to 48 months (2 river basins). The parameters were calibrated using the first 24 months of available data. Validation in time was then carried out using the last 12 months of data that was available for each of fourteen remaining basins. Validation in time indicated that the model was reasonably estimating the flows of the basin independently of the time period used for calibration. In addition, monthly estimates follow the pattern of rainy and dry seasons well and provide reasonable estimates of monthly runoff.

The model parameters were also grouped together according to basin characteristics. This was done in order to determine if there were any definitive patterns between model parameters and their soil, vegetative or any other grouping. Model parameters were then averaged according to a particular group within each

General Vegetation Group	β	SMC (mm)	PERK (mm)	SSFC (mm)
Grassland	10.3	52	58.2	0.55
Woodland	8.0	185	43.6	0.34
Rainforest	6.3	374	32.4	0.22

Table 1: Model Parameters and Vegetation Categories

characteristic. For example, the 57 river basins were grouped into vegetation categories based on FAO vegetation classifications available on large scale maps. The calibrated model parameters were then grouped together according to the particular vegetation classification of the basins. The averaged parameters were then run through the model (without recalibration) using all of the basins and the average standard error calculated. It was found that grouping by soil texture and slope provided very inaccurate results. The groupings of soil type, lithology and vegetation were found to have potential use for the estimation of model parameters and subsequently estimation of runoff. For obvious reasons, the vegetation category lends itself the best to remote sensing.

Model parameters were then examined to see if they correspond to expected physical phenomena. For example, Table 1 shows the parameter values for three general vegetation categories. As expected, soil moisture capacity and the proportion of soil moisture available for evapotranspiration decreases with the increase in the depth of the soil. Similarly, soil percolation capacity decreases with the increase in soil depth. Also, the subsurface flow coefficient increases as the vegetation type changes from grassland to rainforest. This can be explained by the increased porosity of the rangeland soils as well as the more arid climate of the region which reduces the possibility of direct runoff.

The parameter values obtained using the vegetation and FAO soil classification system was validated in space using data for Cote d'Ivoire that became available towards the end of the investigation. It was found that the model provided reasonable estimates of total annual runoff and monthly runoff of the test basins despite the basins used for validation being generally much larger than the 100 square kilometer limit imposed on the initial model calibration.

RAINFALL FROM METEOSAT DATA

Since the advent of meteorological satellites, many methods for estimating rainfall from space have been developed. A description of some of the most successful methods can be found in Barrett and Martin (1981) and Atlas and Thiele (1981). Each of these methods developed over the years may vary substantially from one another, depending on the initial user requirements, time frames of measurements and the format and type of satellite data used. For example, some methods rely on visible channel data, cloud type classifications and raingauge data to develop a rainfall estimation model . Others may use the bi-spectral characteristics of the clouds in conjunction with weather radar to develop the model algorithm . In almost all cases, the information derived from the meteorological satellite supplements the existing raingauge network and by no means replaces it.

With all of this in mind, a rainfall estimation algorithm was developed based on the following criteria;

- The rainfall-runoff model dictated a monthly time step for the rainfall estimates.
- The rainfall model should be flexible enough to allow the user to estimate rainfall for any given basin within the study region
- The model would require archived satellite data since the basin runoff data was only available for 1986.

In examining the rainfall patterns and cloud behavior within the Sudan-Sahel zone during the rainy season (May - October), one quickly realizes the speed with which the systems travel across this part of the continent. During the early months of the rainy season, the main rain bearing system within the zone occur along squall lines which travel from the east along the trade winds Using Meteosat, these squall lines can be monitored and reports typically show them traveling across continental West Africa in 24 hours (AGRHYMET,1986). Also, storms within the region tend to be very spotty and of low duration and high intensity For these reasons, it is important to monitor the situation over the entire region relatively frequently Given the situation, Meteosat data is seen as the logical choice for satellite data

Attempts were made to obtain visible or infrared imagery (and/or CCT's) spanning a long enough period of time to de-

velop a rainfall model. Unfortunately obtaining such information at the required intervals (maximum of 6 hour intervals was chosen as an extreme outside limit) was not possible. Such a high frequency of images are simply not readily nor easily available. Although daily images are available, they are obviously totally unsatisfactory for proper monitoring of meteorologic information for an area such as West Africa. Fortunately, data for the entire rainy season of 1986 (May to October) was stored in the form of cold cloud top occurrences (O_c) and Meteosat maximum surface temperatures (T_{max}), both based on data obtained from the Meteosat infrared channel. The ORSTOM research station in Lannion, France in conjunction with the Centre Météorologique Spatiale (CMS) has developed an algorithm which analyses the corrected infrared data from Darmstadt and keeps an accounting of the cold cloud ($\leq -40°$ C) occurrences and the maximum surface temperatures for 5 day periods within the rainy season. This data is obtained for all pixels within the Meteosat image frame at intervals of three hours. Although Meteosat images are available every 1/2 hour, computing power limited the algorithm to accounting at intervals of every three hours.

Much work has been done over the years using this form of satellite data for developing rainfall algorithms, particularly for this region of Africa. It has been shown (Carn & Lahuec, 1987) that monthly rainfall amounts, using regression equations based on cold cloud occurrences and known rainfall levels, could be estimated successfully In this case, it was felt that a strong relationship exists between the frequency of cold clouds passing through a particular pixel and the known rainfall amount (based on a rainfall gauge somewhere within the pixel). Similar work has been done at the University of Reading (Milford & Dugdale, 1985) however, the threshold temperature was set $\leq -60°$ C.

In parallel with the attempts to determine the relationship between cold clouds and rainfall, there have also been attempts to determine whether or not a relationship exists between the the changes in maximum surface temperature and rainfall amounts. What is expected is that during and after rainy periods, the soil temperature becomes cooler due to the lower reflective characteristics of the soil and the actual evapotranspiration process. Studies show a correlation coefficient of - 0.88 between the Meteosat surface temperature summed up for the entire rainy season and the corresponding rainfall amounts. (See Assad et. al., 1986)

The rainfall data obtained for this study is made up of 142 rainfall stations located within the entire Sudano-Sahelian zone. Ninety three of the stations are synoptic weather stations scattered throughout the zone, with the bulk of the stations located in the more southern reaches. The remaining stations are all secondary meteorolgical stations located in Niger.

Using a navigation algorithm developed by ESA and implemented at the CMS, it was possible to determine the location of each rainfall station within the meteosat image frame. The accuracy of these results are considered good to +/- 1 pixel. With this information it is possible to extract both the cold cloud occurrences and maximum surface temperatures for each of the 143 stations Since the cold cloud top occurrences and the 5 day maximum Meteosat surface temperature data sets were both available, it was felt that a statistical regression equation using the above mentioned values as the independent values and the rainfall as the dependent could be developed.

Histograms of the spatial distribution of the rainfall data for the six months within the rainy season were plotted to determine how the monthly data were distributed. Based on these results and the values obtained for the coefficients of skew and kurtosis, the distributions fit a log-normal distribution for the 'dry' months and a normal distribution for the 'wet' month. This lead to the development of the regression equations in the following form.

Statistic	MAY	JUN	JUL	AUG	SEP	OCT
α_1	1.431	1.693	5.648	4.46	3.26	0.83
α_2	-0.080	-1.372	-6.118	-7.951	-9.193	-4.12
α_3	0.086	2.131	257.4	326.5	372.6	6.77
r	0.74	0.69	0.79	0.79	0.76	0.83

Table 2: Rainfall Regression Results

$$PPT = \alpha_1 O_c + \alpha_2 T_{max} + \alpha_3 \quad (1)$$

for the wet months of July, August and September. The regression equation for May June and July are given as,

$$\log(PPT) = \alpha_1 \log(O_c + 1) + \alpha_2 \log(T_{max}) + \alpha_3 \quad (2)$$

Where $\alpha_1, \alpha_2, \alpha_3$ are the regression coefficients.

The results of the regression show a correlation coefficient value ranging from the lowest in June of 0.70 to the highest in October at 0.83. A summary of the regression results are listed in Table 2.

ESTIMATING BASIN RAINFALL AND RUNOFF

The rainfall - runoff model described previously estimates monthly runoff values using the lumped parameter values given in Table 1 or by calculating these parameters using the previously described calibration technique (provided a large enough sample of historic data is available). In either case, once the 4 model parameters are established, it is possible to estimate monthly runoff using only monthly rainfall and temperature data as model inputs. In our case, runoff data from 5 small basins within the zone were available for validation purposes. This enabled the verification of the rainfall runoff model with respect to runoff estimates and established a basis for comparing conventional and remote sensing estimates of basin rainfall.

River	Station	Size (sq.km.)	β	SMC (mm)	PERK (mm)	SSFC (mm)
Douni	Point398	120	12.0	311.6	54.4	0.14
Kobani	Nimbrini	342	8.0	185.0	58.2	0.55
Degou	Diolala	450	8.0	185.0	58.2	0.55
Doundian	Wahire	640	4.7	106.0	94.1	0.28
Ouaireba	Wahire	810	1.9	93.6	94.2	0.32

Table 3: Rainfall-Runoff Model Parameter

The size of the basins used in this verification varied from 120 sq.km. to 810 sq.km. with all of the basins located in Northern Cote d'Ivoire. The general vegetation category for all the basins fell into the Sudanian Woodland category with the associated model parameter values as given in Table 1. Because data for 1982 and 1983 were available for three of the five basins a calibration of the four model parameters was performed. Since the calibration was possible, these calibration basins were then validated in time alone while the remaining two basins were validated in both space and time. Table 1 gives the rainfall-runoff model parameters for the 5 study basins.

In order to compare the rainfall-runoff model results using both conventional and remotely sensed rainfall as an input, an areal basin estimate was required. The conventional method was based on a weighted average where the weight is a function of

the distance between the gauge and the center of the basin. This method was chosen because in all cases, the raingauges were located outside the basin. The remote sensing rainfall estimates are based on the previously mentioned regression technique with α_1, α_2 and α_3 as given above. The five basin boundaries were delineated within the Meteosat image frame and the T_{max} values along with the O_c pixel values were extracted for each pixel within

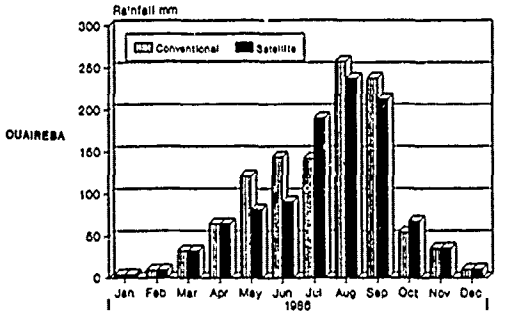


Figure 1: Basin Rainfall Estimates (Ouaireba)

each of the 5 basins. The regression equation was then applied to each of the pixels, giving a rainfall estimate for that pixel. The basin estimate was then calculated by averaging the rainfall estimates of the pixels within their respective basins. Areal basin rainfall estimations results for Ouaireba@Wahire are given in Figure 1. With both the remotely sensed and conventional rainfall estimates available, it is possible to compare the rainfall-runoff model response to these two inputs The rainfall-runoff model monthly hydrographs for Douni@Point398 are given in Figure 2.

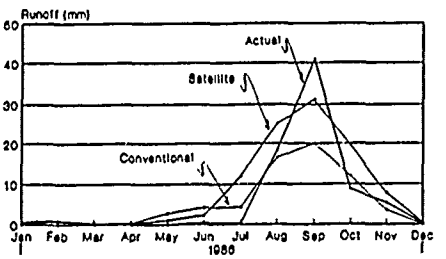


Figure 2. Basin Runoff Estimate (Doundi@Pt398)

In all 5 cases the remotely sensed basin rainfall estimates yielded better model results, with respect to the standard error calculated from known runoff values. The standard error of validation of runoff for all 5 study basins are given in Table 4. Based on the standard errors calculated for the monthly hydrographs a trend towards an improved method of rainfall interpolation is demonstrated. This is of particular interest to areas such as the Sahel, where there is only a sparse network of existing rainfall gauges. Also demonstrated is the possibility of assessing monthly flows for small basins in the region where no hydrometric data already exists.

Basin Name	Station Name	Conventional Runoff Error (mm)	Remote Sensing Runoff Error (mm)
Douni	Point398	6.39	5.76
Kobani	Nimbrini	31.07	28.53
Degou	Diolala	8.34	7.13
Doundian	Wahire	16.92	9.89
Ouaireba	Wahire	11.96	11.75

Table 4. Rainfall-Runoff Model Results - Standard Error

CONCLUSIONS

There is a significant relationship between vegetation, precipitation and runoff in the Sahel area. This relationship can be used with a conceptual (deterministic) rainfall-runoff model to estimate monthly flows in small (less than 500 sq. Km.) ungauged river basins. The results, based on monthly hydrograph estimates, demonstrate a trend towards improved rainfall monitoring by using satellite remote sensing as a means of interpolation between known raingauge values. Also, by using the rainfall-runoff model in conjunction with the remotely sensed data, the possibility of reliable runoff data for basins with little or no existing hydrometeric data is introduced. The error of estimate of precipitation values is of the order of (+/-) 40%. The error of estimation of monthly runoff values is the order of (+/-) 20%. The accuracy obtained by the technique is equivalent to that which could be obtained if a network of about 300 stations could be installed on small river basins in the Sahel area.

REFERENCES

[1] AGRHYMET, Synthèse regionale des situation agrométéorologiques de la saison des pluies, No 197,1986.

[2] Assad, E: Seguin, B. Kerr, Y: Freteaud, J.P.. and Lagouarde, J.P. , Veille Climatique Satellitaire, No 13. pp.21-29, 1986.

[3] Atlas, D. and Thiele, O.W., Precipitation Measurements from space - Workshop Report. Greenland, Maryland: NASA, 1981.

[4] Barrett, E.C. and Martin, D., The Use of Satellite Data in Rainfall Monitoring. Academic Press, 1981.

[5] Carn, M. and Lahuec, J.P., Estimation des précipitations au Niger au cours de la saison des pluies 1986, Veille Climatique Satellitaire, 17, pp 40-48, 1987.

[6] Davy, E.G., Mattei, F. and Solomon, S.I. (1976), An Evaluation of Climate and Water Resources for Development of Agriculture in the Sudano-Sahelian Zone of West Africa. Special Environment Report No.9, World Meteorological Organization, Geneva, 1976.

[7] Link, L.E , "Compatibility of Present Hydrologic Models with Remotely Sensed Data", Proceedings of the 17 International Symposium on Remote Sensing of Environment, Ann Arbour, Michigan, 1983.

[8] Milford, J.R. and Dugdale, G., "Rainfall estimation over the Sahel using Meteosat Thermal Infrared Data," in ISLSCP - Proceedings of an International Conference held in Rome, ESA, 1985.

[9] Ragan, Robert, M. and Jackson, Thomas, J. "Runoff Synthesis Using Landsat and SCS Model," Journal of the Hydraulics Division. (ASCE), pp. 667-678, 1975.

[10] Rosenbrock, H.H., "An Automatic Method for Finding the Greatest or Least Value of an Function," Computing Journal. Vol. 3, pp. 175-184, 1960.

[11] Turc, L., "Le bilan d'eau des sols. Relations entre les précipitations, l'évaporation, et l'écoulement " Annuaire Agronomique. Vol. 5, pp. 491-596, 1956

Remote Sensing for Development in Ethiopia

Asfaw Fanta

Ethiopian Mapping Authority

ABSTRACT

In view of the size, the natural resource base and the changing weather conditions of Ethiopia, the application of remote sensing techniques for rapid mapping and environmental monitoring is very suitable. The National Centre for Remote Sensing, although at its inceptive stage, is making good progress in strengthening its capability. Some assistance, by way of technology transfer, implementation of resources studies using remote sensing data and acquistition of equipment, has been received by the centre through Technical Cooperation Programmes with international & governmental institutions outside Ethiopia. However the centre is not as yet in a position to segnificantly broaden the terrain data base of the country and to cope with the ever increasing demand by users in a variety of the economic sectors for timely and reliable terrain information. These and some of the constraints hampering the much desired promotion of remote sensing in Ethiopia are discussed and possible solutions suggested.

I. INTRODUCTION

Ethiopia is situated in the Eastern Horn of African with an area of 1,223,000 sq km stretching from 3° N to 18° N latitude and from 33° E to 48° E longitude. It has a population of about 48,000,000 which makes it the 3rd largest in Africa after Nigeria and Egypt.

The topographical features are diverse with high mountains, plateaux,plains, gorges and valleys. The altitude ranges from 4,620 metres above sea level at Mount Ras Dashan to about 110 metres below sea level at the depression of the Kobar sink. The Western and South Eastern Highlands which are distinctly separated by the central lowland areas of the Rift Valley are conspicuous features. The areas surrounding these two major highland systems are lowland areas. The absolute altitude of these lowlands are higher than several mountains and it is relative to the overall topography that they are referred to as 'lowlands' for example at some places the Rift Valley floor rises to a height of 2,000 meters.

The general geology of the country suggests the existence of several types of minerals, oil and gas deposits. Ethiopia has many lakes and big rivers. Some of the rivers flow beyond the boundaries of the country washing away enormous amount of valuable soil. Many of these rivers have their sources on the high mountains thus cutting the land into deep gorges and valleys.

The landscape presents widely varied aspects of vegetation ranging from scanty desert vegetation to Afro-Alpine vegetation type. Heavy deforestation for cultivation, settlement and fuel wood has diminished the high forest region from 40% to 4.4% over a period of about 50 years. The continuing deforestation, overgrazing, traditional cultivation methods, etc, accelerate erosion. This can clearly be seen in the northern parts of the country where farming has been practised for a long time with deforestation completed and agricultural potential almost exhausted. In general, the high population growth rate (2.9%) and the misuse of the resources together with the adverse natural processes have literally changed the landscape of the country.

As a result, the government's immediate

conern is to rehabilitate the land, prevent further damage and develop its natural resources. In this respect, the timely and regular supply of reliable terrain information is an absolute necessity to restore the ecological balance and to be able to exploit the natural resources for the benefit of the people. Remote sensing is an appropriate technology for environmental monitoring and rapid mapping and evaluation of the resources. Its wider application in Ethiopia therefore appears to be very promising.

II. THE NATIONAL CENTRE FOR REMOTE SENSING (NCRS)

The NCRS is the national focal point for remote sensing. It is responsible for the planning and execution of remote sensing activities and ensures that the requirements of user agencies are well coordinated. The NCRS functions under the auspices of the Ethiopian Mapping Authority (EMA) which in turn comes directly under the Office of the Council of Ministers. It has full mandate for the collection, processing and supply of terrain information for the planning, development and management of the natural resources of the country.

In the organizational structure of EMA there is a Map & Remote Sensing Advisory Committee composed of representatives from major user institutions and chaired by the General Manager of EMA. The Committee is responsible for setting priorities for mapping and remote sensing projects in pursuance of the guide lines in the National Economic Development Plan, and in ensuring that all users have balanced access to the centre's facility. The immediate objectives of the NCRS may be summarized as follows:-

- Promotion of the application of remote sensing technology through training and technology transfer.
- Processing of satellite data and establishing Geographic Information System.
- Establishment of a national archive for all satellite data covering the country and
- Provision of user's assistance services in applied remote sensing.

The centre is managed by a Senior Expert and has 10 technical staff including a United Nations consultant who is mainly engaged in training the local staff and supervising projects. The Head directs and coordinates all technical activities and EMA's central administrative office renders all logistical and other general services. Very Soon the NCRS will have its own administrative service. Although the centre is at its inceptive stage, it is making a good progress in building up capability in trained manpower and the acquisition of necessary equipment for analogue and digital image processing and interpretation.

III. THE NEED FOR, AND ABAILABILITY OF, TERRAIN DATA BASE

It is only through proper planning and development of the natural resources of the country that the economic well-being of the people can be improved. Sound economic planning, be it at macro or micro level, must be based on complete and reliable information about the location, extent, type and subsequent exploitation of the natural resources. With the advent of remote sensing, man's knowledge about the earth's resources and the changing environment has expanded greatly. Satellite remote sensing in conjunction with aerial photography and other traditional mapping techniques providesnew possibilities for quick, inexpensive and effective means of terrain data collection and evaluation.

Even though the application of remote sensing in Ethiopia started late in 1984, it is now becoming more and more popular among planners and earth scientists. Landsat MSS data have been used for general landuse and land reclamation studies, land cover classification and woody biomass assessment. The preparation of the National Landuse Plan was based on the result of the studies. But as the scale of the imageries was 1:1,000,000 and the method of interpretation was analogue, the information is found to be inadequate. Plan is, therefore, underway to up date the study by using Landsat TM and Spot XS and digital image processing method. Some pilot projects have been carried out to evaluate the cost and time effectiveness of map revision and production using satellite imageries. One map sheet at the scale of 1:50,000 has already been revised using Spot pancromatic imageries and a new mapping of an area of 3,780 sq km using Spot pancromatic imageries is in progress.The ground control required for the geometric & precision correction of the imageries will be acquired by GPS and the actual image restitution is being undertaken on the digital system of the Swedesh Space Corporation. The results of the evaluation is pending the completion of certain activities.Use of the Spot pancromatic and high resolution Soviet space photographs will be made at EMA for mapping at 1:50,000 scale applying advanced photogrammetric techniques and for revising the 1:250,000 scale maps. Other satellite data such as the Soviet KATE are also available for some portions of the country in pancromatic and multispectral films.

The weather satellite data is received regularly by the National Meteorological Agency and this together with the NOAA-AVHRR data is becoming very useful for the National Early Warning System.

The satellite data for digital and analogue processing are purchased abroad and are available in CCTs and other photographic products. The nonavailability of a ground receiving station for high resolution satellites such as Landsat and Spot, neither in Ethiopia nor in the neighbouring countries, has caused inconvenience in acquiring the desired Sat-data in real or in near real time. In general there is a growing interest in the application of remote sensing by users in a number of sectors and this is a big challenge for the NCRS to cope with.

The whole of Ethiopia has been covered by topographic maps at scale 1:250,000 with contour interval at 100 metres. This maps are being used extensively for pre-feasibility studies for resources development and as a base for other surveys such as geological, soil and water resources.

Since maps at this scale cannot portray detailed and accurate information for final planning and project preparation and execution, EMA has embarked upon 1:50,000 scale topographical mapping project with the aim of covering selected areas of the country. So far only about 29% of the country has been mapped at this scale completion of this mapping programme will take 20 to 25 years even assuming considerable increases in productivity. Other small scale maps of the order of 1:500,000, 1:1,000,000 and 1:2,000,000 and a variety of thematic maps are also available. In spite of the need for revision of most of the existing maps, so far no attempt has been made to embark upon map revision programmes since the available resources are limited and the priority is the extension of the map coverage of the country. At present possibilities of launching map revision programmes for map scales of 1:50,000 and 1:250,000 with the Spot and the Soviet KATE & KFA photographs are being considered in light of the time and cost effectiveness of this approach.

A multicolour edition of the First National Atlas of Ethiopia has been published recently. The contents of the Atlas cover 76 topics dealing with environmental, resources, socio-economic and historical aspects of the country. The availability of such a comprehensive data source on these subjects has strengthened the existing data base and has proved very useful for macro-economic planners, researchers,decision-makers and the public at large.

Aerial photography is normally flown as a component of mapping projects. A complete photographic coverage at the scale of 1:50,000 has been acquired in 1972 for the 1:250,000 scale mapping programme. New photography at the same scale is being flown for the 1:50,000 scale mapping. Some areas where intensive development activities are taking place are covered by large scale photographs, ie. scales 1:20,000, 1:10,000 etc. These photographs are also used for interpretation purposes.

IV. TECHNICAL COOPERATION PROGRAMMES AND TECHNOLOGY TRANSFER

EMA has been cooperating with international and governmental organizations out side Ethiopia for carrying out mapping and remote sensing projects. Essentially the technical cooperation programmes aim at producing maps and/or carrying out resources studies using remote sensing techniques in a given time frame. The equipment and expertise which are necessary for the job are usually provided by the cooperating institutions and in most cases these programmes have been acomplished sucessfully. It was through such arrangements that the topographic mapping at scales 1:250,000 and 1:50,000 and that the landuse studies and the woody biomass assesement using landsat data were implemented. Needless to say that technical cooperation programmes are necessary to acquire in time the terrain data needed for the various economic sectors as the local capability in the fields of cartography and remote sensing is not yet adequate.

In an environment where the technical base for developing ones own indigenous technology is non-existent, the obvious option is to judiciously apply proven exogenous technology through technical cooperation programmes, Such an approach, however, should not preclude the possibility of creating the condition for developing indigenous technology. Neverthless experience reveals that the positive impact of technology transfer through technical cooperation programmes has not been so encouraging. After the projects foreseen in the cooperation agreement have been completed and the programmes terminated, until very recently, it has not been possible for EMA to undertake similar projects independently simply because the participation of the indigenous staff in the various areas of the project activity has not been sufficent and emphasis has not also been given for local manpower training and technology transfer during the implementation of the projects. The foreign experts who are assigned to the projects are primarly concerned with the fulfillment of their responsiblities according to work plans and naturally cannot attach great importance to the training of local staff.

It is often the case that the equipment purhased for executing such projects rarely extend their services beyond the duration of the projects for lack of maintenance and repair facilities locally. The only way to keep the equipment working is by bringing service technicians and spare parts from

outside the country which, of course, is a very costly and time consuming option. This option itself becomes impossible if the failure rate of the equipment happens to be high. Delay in project implementation is caused due largely to lack of maintenance and services in the local enviromental especially for electronic outfits. Technical cooperation programmes usually donot give serious consideration to this aspect while planning purchase of equipment.

In order for technical cooperation programme to foster effectively technology transfer it should aim at:-

- Building local capability at all levels, ie, from routine maintenance of equipment to project management.
- Giving attention to a continuous exchange of experiences through study tours and consultancy services.
- Designating specific duties to local personnel so that they feel a sense of responsibility and this will guarantee the success of the learning process.
- Ensuring that cooperating institution provide support services in a sustained manner so that the recipient institution has a fall back position whenever its own system fail to operate.

In this context, the technical assistance programme between IDRC and EMA,although limited in scope, may be cited as a good example in technology transfer. IDRC provided financial assistance for the preparation and publication of the First National Atlas of Ethiopia. The data collection analysis and compilation of the maps including the writing of the explanatory texts have been carried out by the staff of EMA. IDRC's involvement was confined to giving technical advice on the design and cartographic presentation of the Atlas and some occasional supervision as and when required. The work of the Atlas has been accomplished successfully leaving behind a big confidence and experience among the Ethiopia experts and this will certainly enable EMA to undertake in future similar projects on its own.

V. CONCLUSIONS

The existing terrain data base couldnot satisfy the requirements of the various sectors of the economy. There is therefore an urgent need for speeding-up the process of terrain data collection and evaluation. The conventional method of mapping is very costly and time consuming; furthermore, it is not suited for monitoring the environment and for interpretation of the different land features, Remote sensing has certain capabilities such as synoptic view, repetitivity and multispectral interpretability which the conventional method is lacking. The conventional method will,of course, continue to play an important role for large scale mapping. Remote sensing and the conventional method should be regarded as complementary techniques for the acquisition, processing and evaluation of resources and environmental information.

The NCRS which has the responsibility for promoting the application of remote sensing is making an encouraging effort to strengthen its capacity. Some technical assistance has been received from the United Nations Development Programme and the Food and Agriculture Organization for training the staff of the centre, purchase of equipment and materials as well as for carrying out some studies using remote sensing techniques. The results of these studies have proved to be very useful for the development planning and management of the natural resources. The centre has succeeded in serving as the national focal point and as a coordinating body for remote sensing activities in the country.

The centre has difficulty in getting satellite data easily from the different sources abroad. The fact that there is no ground receiving station for capture of high resolution Sat-data in Ethiopia and the neighbouring countries is a draw-back for the centre. The centre is also facing other constraints such as shortage of trained personnel, equipment and local manitenance and services. The technical assistances received so far have been very useful in providing satellite data and undertaking some remote sensing activities. However, they donot seem to have put enough emphasis on technology transfer so the centre is not as yet in a position to implement remote sensing projects without some help from experts on consultancy services. It is therefore suggested that Technical Cooperation Programmes should play a more important role to accelerate the process of technology transfer and thereby enable the recipient institution become self-reliant in a reasonable period of time.

REFERENCES

1. A Feasibility Study for the Establishment of the Ethiopian Centre for Remote Sensing and Geographic Information - Addis Abeba, Dec. 1987.
2. National Atlas of Ethiopia - Addis Abeba, Jan. 1988.
3. Proceeding of Development & Applications of Remote Sensing for Planning, Management and Decision-making - Beijing, Jan. 1985.

EVALUATION OF SATELLITE DATA FOR TOPOGRAPHIC MAP UPDATING IN KENYA

Hussein O. Farah, Philip J. Howarth and Ronald A. Bullock

Department of Geography, University of Waterloo
Waterloo, Ontario N2L 3G1, Canada

Abstract

Reliable up-to-date topographic maps are essential for the orderly development of a nation. Many developing countries, including Kenya, are now completing their national map coverage at 1:50,000, 1:100,000 or 1:250,000 scale. However, due to an increased pace of economic development and high population pressure on the land topographic maps are rapidly becoming out of date. Traditional updating methods using ground surveys and aerial photographs are slow and exceedingly expensive for a country with limited resources. Therefore faster and more economical methods need to be developed. Landsat Thematic Mapper (TM) and SPOT panchromatic (P) imagery are evaluated for information content and geometric fidelity to aid in the 1:50,000 map revision programme. Preliminary results demonstrate the usefulness of these data.

Keywords: Kenya, Landsat, Map revision, SPOT

Introduction

Map revision is a major task of mapping organizations (Datta,1980; Thompson, 1980; Usery et al., 1984; van Zuylen, 1980). In developing countries, basic mapping has been undertaken over a long period of time with the assistance of foreign mapping agencies. In Kenya the whole country is covered at 1:250,000 scale and two thirds at 1:50,000. The remainder of the area being arid and sparsely populated is mapped at 1:100,000. The Directorate of Overseas Surveys (DOS) of Great Britain was the main agency which helped Kenya to accomplish this mapping.

The maintenance of topographic maps, which is a continuous task, is not considered aid worthy. It is, therefore, the responsibility of the local mapping agency to keep the map series in a healthy state using funds from its own meagre recurrent budget. Such agencies need to develop revision capabilities which can be supported with limited resources and at the same time permit the regular updating of maps. This calls for methods which are simple and cost effective.

An efficient revision programme requires:

- A logical and systematic approach to determine which maps are revised and when.
- Cost effective data acquisition.
- Simple instrumentation (Thompson and Crane 1986).

The establishment of an "intelligence" gathering centre is also important. A lot of information on changes in the landscape is held by various government departments. This information, if used efficiently, can guide the map-updating programme, avoiding the unnecessary flying of certain areas where information is already available or in which no significant changes have occurred. Data acquisition for mapping new detail is the most expensive component in the revision exercise. Conventional ground surveys and the acquisition of aerial photographs are exceedingly expensive. Hence the need for newer less costly sources of information. High resolution satellite data is now available with the operation of the SPOT 1 and Landsat-5 satellites. The suitability of such data for topographic mapping has already attracted researchers. In this study SPOT P and Landsat TM are evaluated for map updating at 1:50,000 scale in Kenya.

Methodology

The study area is a 30 km x 30 km region centred near Nakuru in the central Rift Valley of Kenya. This area has a wide range of cultural and natural features. Topography is varied, most areas are of low relief, but some rugged areas are found both within and along the edge of the Rift Valley.

Landsat-5 Thematic Mapper (TM) digital data and SPOT P digital data (Level 1b), imaged in July 1984 and March 1986 respectively, were acquired. A 1:50,000 scale map of the area, published in 1975 but based on 1969 aerial photographs, was available. A Dipix ARIES III image analysis system was used to enhance, geometrically correct and register the data to the UTM grid. Image enhancement was used to

improve the interpretability of detail. The system-corrected histograms were linear stretched to cover the full brightness range. The maximum and minimum values were assigned to the mean brightness value +/- 3 standard deviations. Visually, this appeared to give a better result than a simple linear stretch.

The simplest method to rectify satellite data is to apply polynomials to calculate the parameters of a two-dimensional geometric transformation using ground control points (GCPs) and then to resample the image to match the desired map projection (Dowman, 1987; Welch et al., 1985). The rectification involved matching GCPs in the image to the 1:50,000 map. A least-squares solution was used to evaluate the polynomial coefficients. Identification of the UTM coordinates of GCPs in the map was done with a digitizer. For the images, line and pixel values were identified on the ARIES display system. About 60 potential GCPs were located on the map. However, only 22 and 20 points were found to be reliable for TM and SPOT, respectively. This was because the points were either poorly defined in the images or changes in the positions of natural features (e.g., rivers) caused unacceptable errors at these points. Of the reliable points, 10 were used for the transformation (control points), the balance being used to check the accuracy of the rectified image (check points).

Data were then processed on the FIRE-240 colour film recorder to obtain positive transparencies. For the Landsat TM image, Bands 2, 3, and 4 were used to create a false-colour transparency. Hard copies at 1:50,000 were produced and features were directly mapped on overlays. The geometric accuracy and the information content for the 1:50,000 scale mapping were then analyzed. This was done by comparing the information on the overlays with the existing 1:50,000 scale map.

Results and Discussion

The rectification was performed with first-, second- and third-order polynomials. Ten control points were used for both data sets. Table 1 illustrates the root mean square errors (rms) in planimetry at 10 and 12 check points for SPOT P and Landsat TM, respectively. As can be seen, subpixel accuracies were obtained for Landsat TM and second-order polynomials gave the best results of 11.6 m and 13.6 m for SPOT P and Landsat TM, respectively.

For SPOT P, first-order polynomials gave rms errors of 7.3 m in the x-direction and 18.9 m in the y-direction. This suggests that the error in the y-direction is non-linear and hence cannot be rectified by a first-order polynomial. On the other hand, the third-order polynomial was not suitable because of the limited number of control points available. Ten points is the minimum number required for third-order caiculations. Normally higher accuracies can be obtained if more control points than the minimum are used. For these reasons, the second-order polynomial was used in the final rectification of both data sets.

The above results are well within the acceptable accuracies for 1:50,000 scale mapping. This is significant because some parts of the study area had moderate relief. This means that such relief does not introduce intolerable errors. Therefore, complex procedures involving digital elevation models are not necessary in the rectification. Further, the transfer of information from the images can be done accurately with simple instruments.

The completeness of the main planimetric details such as communication routes, drainage networks and settlements was evaluated to determine the interpretability of the images. Table 2 shows some of the preliminary results. To evaluate the maps, it was necessary to select areas of minimum change. Two areas, a 10 km x 10 km section of rural land and a 6 km x 4 km area located within Nakuru town were chosen. To check other features, the whole study area was used. Routes were classified into three levels: primary roads, other roads and tracks, and railways. The primary roads can be easily identified on both the TM and SPOT images. However, smaller roads, such as tracks, pose some difficulties especially in the TM. While 78% of these roads were identified on the SPOT image only 55% appeared on the TM image. For the urban region, all primary roads and railways could be plotted from the SPOT image.

Table 1. Rectification of Landsat TM and SPOT P Images

Polynomial Order	RMS (m) Landsat TM	SPOT P
1	19.1	20.2
2	13.6	11.6
3	24.6	13.9

Some rivers can be seen directly. However, the locations of smaller tributaries are often recognized by the presence of riverine woodland. Where this vegetation is lacking, it is difficult to identify the location of the river due to a lack of contrast between the river bed and the surrounding land. The edge of the built-up area of Nakuru town is well defined. It

has expanded by 19 % and 18 % for TM and SPOT, respectively, over the period of about 10 years between the dates of the map and the images. Airport runways and sewage ponds also display sharp boundaries and are easily identified. Other features which are clearly observed include golf courses, crop-land, field boundaries, parks, quarries, forests and scrub-land. Large buildings such as factories and other areas with dense buildings are visible. However, most of the isolated buildings in the rural areas are difficult to detect.

Table 2. Preliminary Results of Feature Identification from Landsat TM and SPOT P

Feature Type	Topographic Map	% Identification TM	% Identification SPOT P
Primary road (rural)	40.8 km	94	100
Other roads and tracks	189.0 km	55	78
Primary roads (urban)	123.9 km	-	100
Railway	6.4 km	-	100
Rivers	174.4 km	76	87
Built-up area	35.6 (km^2)	119	118
Airfields	2	150	150
Sewage ponds	1	200	200

The above initial results and others from different parts of the world indicate that, together with suitable field work, most of the important features required in map revision can be obtained cost-effectively from satellite data (Dowman, 1987; Moore, 1988). Satellite data could also be used to significantly reduce costs in conventional revision methods by directing acquisition of photography to areas where enough change has occurred to warrant revision (Gregory and Moore, 1986). However, some questions need to be raised. Should we not reassess the specifications of our topographic maps in order to accommodate newer sources of data? Current specifications have been established in an *ad hoc* manner over the years and are mainly oriented towards what can be extracted from aerial photographs, rather than what the map user needs. By redefining the existing standards, better use could be made of newer data. There are also new formats of cartographic products such as image and space maps. These products, although cartographically less perfect than traditional line maps, contain most of the recent information required by users and can be obtained cheaply and at more frequent intervals. Mapping organizations, especially those in the developing countries, should therefore think of ways to introduce less costly data into their mapping programmes.

Conclusion

The objective of this study, investigations of simple and cost-effective methods to allow developing countries such as Kenya to maintain their topographic series, is still on-going. Preliminary results from the analysis of Landsat TM and SPOT P data indicate that it is possible to plot most of the important features. Even with moderate relief, simple rectification methods give acceptable accuracies. However, not all specified 1:50,000 map detail is easy to detect. It is suggested that we need to reassess the specifications for topographic maps to accommodate newer sources of information such as satellite imagery.

Acknowledgements

Mr. Farah is jointly supported by the Kenya Government and.CIDA. Special thanks to the Regional Centre for Surveying, Mapping and Remote Sensing, Nairobi and SPOT Image Corporation for making the data available.

References

1. Datta, M.M., "Map revision in developing countries", XIV ISP Comm. IV Supplement, Hamburg, 130-139, 1980.

2. Dowman, I.J., "The prospects for topographic mapping using SPOT data", Image Utilization, Assessment, Results, Cepadues-Editions, Toulouse, pp. 1163-1172, 1988.

3. Gregory, A.F. and H.D. Moore, "Economic maintenance of a national topographic data base using Landsat images", Photogramm. Eng. and Remote Sensing, Vol. 52, No. 4, pp. 519-524, 1986.

4. Moore, H.D., "SPOT and Landsat TM for the maintenance of topographic databases", Proc. Internat. Symp. on Topographic Application of SPOT Data, Sherbrooke, pp. 297-308, 1988.

5. Thompson, W.A., "Maintenance of Australian 1:100,000 National Map Series", XIV ISP Comm. IV Supplement, Hamburg, pp. 715-721, 1980.

6. Thompson, W.A. and K.C. Crane, "Report of WG IV-1 on map revision", ISPRS Comm. IV Symp. Proc., pp. 511-518, 1986.

7. Usery, E.L. and R. Welch, "Map revision using digital image processing techniques", Tech. Papers 50th Meeting Amer. Asso. Photogramm., Vol. 2, pp. 753-756, 1984.

8. Welch, R., L. Jordan and M. Ehlers, "Comparative evaluation of the geodetic accuracies and cartographic potential of Landsat-4 and Landsat-5 image data", Photogramm. Eng. and Remote Sensing, Vol. 51, No. 9, pp. 1249-1262, 1985.

9. van Zuylen, L., "Map revision", ITC J., Vol. 1, pp. 130-139, 1980.

EVALUATING LAND RESOURCE INFORMATION WITH GEOGRAPHIC INFORMATION SYSTEMS IN NEPAL.

P.B. Shah,

Head, Integrated Survey Section, Topographical Survey Branch, Box 1611, Kathmandu, Nepal

ABSTRACT: Micro-computer based GIS techniques are proving to be excellent tools to evaluate resource information in Nepal. Examples are provided from a watershed study, where soil, land use, topographic and hydrological data were evaluated with GIS to quantify the test area for agricultural and forestry suitability assessments and to determine changes in land use, erosion, and sedimentation within the basin. Sequential aerial photos are proving to be an essential data source for land use change assessment, and the subdivision of the watershed into elevation/slope/aspect categories facilitates the establishment of micro-climatic zones.

Keywords: GIS, Watershed, Land Resources

BACKGROUND:

Between 1978 and 1985 an Integrated Resource Survey was carried out for all of Nepal under CIDA sponsorship. The landforms, soils, agricultural and forestry land resources and land capability were mapped at the 1:50000 scale and the data provided baseline information on the resources conditions in Nepal. Forest degradation, erosion and soil fertility losses are critical issues in determining the sustainability of agriculture and forestry in Nepal, were population growth is currently at 2.8 % and where resource conditions in the mountains are marginal at the best of times. The land resource mapping project clearly pointed out that the fodder situation is most critical, while food and fuelwood problems are of concern in selective areas of the country.

To address the rate of land use change, soil degradation and sedimentation, a watershed project was initiated in the Jhikhu Khola river basin to quantify processes using a micro-computer based Geographic Information System (GIS). Current and historic aerial photography and ground observations were used to quantify the study area since the use of satellite imagery is inappropriate in this high relief environment due to distortion problems, adverse weather conditions, micro size field parcels and the high cost of such imagery. The Jhihu Khola watershed, located in the Middle Mountain physiographic region of Nepal was chosen for this study because land use pressure is very high. Triple annual crop rotations are widespread in the lower portions of the basin and the demand on the forest for fuelwood, organic fertilizer and animal feed is enormous. Given the natural instability, the friable bedrock, steep slopes (internal basin relief of 1200 meters), and the concentration of rainfall over the monsoon season, soil erosion and slope stability is a very critical issue. To sustain future production in agriculture and forestry and to prevent flooding problems downstream it is imperative to carry out land use practices that conserve the soils and the GIS technique can greatly facilitate the evaluation of the current conditions and recent changes.

METHODOLOGY:

The available resource was digitized using a micro-compu GIS system (Terrasoft) and a resou io was produced which included a e, forestry, landforms, soils and c: .ty. Individual resource maps were com. ed to arrive at a spatial representation to show where in the watershed the high capability forests and agricultural lands occur and what crops, trees and soil conditions are common at such sites. The micro-climatic conditions were determined by dividing the watershed into four slope-elevation-aspect categories which will later be calibrated with available climatic data. The land use change was examined by comparing the 1950 resource data with the 1982 land use inventory and with the GIS. Once all resource information is entered into the GIS system it is now possible to produce any combination of resources in a quantitative manner. The watershed was divided into five sub-basins and the area for each sub-basin was calculated with the GIS system. Each sub-basin has a unique land use and the stream flow and sediment transport is measured on a daily basis at the mouth of each sub-catchment. Two additional stations are used to obtain the integrated input of the basin. The data is then incorporated into

the GIS to determine a waterflow and sediment mass balance for the basin. This will then allow the quantification of erosion in relation to land use.

RESULTS:

The GIS assessment showed that 55.5% of the watershed is under agriculture, 24.5% under forestry and 20% under shrub. Based on the GIS land use change map a decrease in forest land amounting to 50 % was observed between 1950 and 1982. Agricultural land increased by 22% and shrub by 138 % over the same period of time. This suggests that considerable forest degradation has ocurred over the recent past. These findings are significantly different from the changes observed by Mahat et al.(1986, 1987) and Griffin et al. (1988) but their observations are over a smaller time increment. What is of additional interest is that the changes appear to be more dramatic before 1972 and some improvement in area under forest cover is apparent subsequent to the 1972 aerial photographs. This indicates that it is essential to examine sequential aerial coverage over shorter time intervals in order to determine land use dynamics. The land use interpretation of the 1972 images are currently incorporated into the GIS system and this should give us a more refined picture of the land use changes.

A more important problem relates to the deterioration of the forest quality and this is more difficult to assess from aerial imagery. The decline in quality is a result of heavy grazing pressure, collection of animal fodder in the forest, the removal of forest litter for input into agricultural, and the need for fuelwood. The fodder deficiencies have been amply described by Panday (1982), Carson et al. (1986), and Hrabovszky and Miyan (1987), and the introduction of fodder trees is a high priority. The decline in forest quality is known as peripheral degradation and is well documented by Gilmore (1988) and Monch and Bandyopadhyay (1986). The forest floors in many of the forests in the watershed are very deficient in organic matter due to intensive collection by the farmers for animal bedding and organic fertilizers. This is likely to result in a long term decline in the forest production capability and the maintenance of a good soil cover with organic matter is imperative to prevent erosion and to maintain a balance in nutrient cycling.

The introduction of native fodder trees into the forest and into agriculture would reduce the fodder demand to some extent and the GIS based assessment of micro-climatic conditions serves as an excellent example on how GIS can be used to facilitate the quantification of appropriate site conditions for afforestation.

The watershed was divided into two elevation classes (< 1200 m and > 1200 m) to coincide with natural boundaries in vegetation zones. With the GIS all slopes less than 35% were eliminated from the classification because lower slopes are automatically put under agricultural production as a result of food demand by the growing population. Finally, the watershed was divided into north and south aspect exposure which reflects differences in moisture conditions and intensity of rainfall. With the GIS technique the combination of elevation, aspect, and slope could readily be displayed and the summary information is provided in table 1. The four classes reflect unique site conditions which are available for afforestation programs. The low elevation south facing sites are hot and dry while the high and north facing sites are cool and moist. The foresters can now select different types of fodder tree species which best suit the conditions and this should improve survival rates after planting and future growth. The GIS combination map indicates where in the watershed this combination of conditions occur and native, nitrogen fixing fodder tree seedlings with the appropriate tolerance range can then be planted in the appropriate areas.

Similarly the soil conditions can also be combined with the micro-climatic site conditions to further refine the classification. This is of particular importance since the watershed contains a large component of red soils which are notorious for exhibiting poor phosphorous and nitrogen conditions.

Table 1. Micro-Climatic classification of the watershed based on elevation, slope, and aspect.

Elevation	Area in ha	% of Watershed
< 1200 m	9155	65 %
> 1200 m	4905	35 %
Aspect		
North (NW & NE)	4575	32 %
South (SW & SE)	2906	21 %
Slope		
> 35 %	9560	68 %
< 35 %	4500	32 %
Class 1	487	
Class 2	380	
Class 3	534	
Class 4	392	

1 = < 1200 m / North / > 35%
2 = > 1200 m / North / > 35%
3 = < 1200 m / South / > 35%
4 = > 1200 m / South / > 35%

With the GIS display maps it can be shown where the red soils occur in relation to the topo-micro-climatic categories. It is obvious that the micro-climatic subdivision based on topography, slope, and aspect alone is insufficient and refinements will be made once the climatic data for the watershed has been tabulated.

Another GIS application relates to the hydrometric stations at which the waterflow and sediment output is measured on a daily basis in five separate sub-basins. The sub-basins were digitized from the topographic map and table 2 provides a brief description of the areas involved and the major land uses within the catchment areas.

Table 2: Distribution and land use of five sub-basins which are monitored on a daily basis.

Sub-basin	area		Dominant Land Use
1	114	ha	Natural Landslide, Grazing
2	116	ha	Agriculture (1-2 rotations)
3	78	ha	Forest
4	613	ha	Mixed Agric. & Forest, North Facing
5	505	ha	Mixed Agric. & Forest, South Facing

With the GIS analysis it is now possible to characterize the sub-basins by determining length of drainage channels, channel density, relief, slope, and dominant aspect, percent forest cover, shrub, and percent of crops and cropping systems. This can then be related to water flow patterns, and sediments loads. This will provide a crude evaluation on how much water and sediments are produced in each sub-basin and which land use and site conditions are most responsible for producing high sediment loads.

Using the land use, soil, and capability maps a number of combination maps were produced using the overlay or extraction techniques and these prove to be an excellent source of information to determine where in the watershed the most capable agricultural and forestry sites are located. The search program within GIS will allow the characterization of the bio-physical conditions such as slope, aspect, elevation and soils and this can serve as a model to find all other sites within the watershed which have similar site conditions. This can serve as a suitability assessment for agriculture and forestry and the experience from a known site can then be transferred to sites with analogue conditions. The same process can be used in forestry where afforestation success can be extended to areas where conditions are unique and biophysical conditions most appropriate for a given set of tree species. As shown in these examples the GIS technique shows great potential and this micro-computer based technology will greatly facilitate land evaluation and resource management in the Middle Hills of Nepal.

CONCLUSIONS:

1. Information extracted from aerial photographs is becoming a major data source for GIS analysis and the use of sequential photos is of particular interest since it gives quantitative information on rates of land use change.
2. Combining resource information with GIS greatly facilitates land use assessments and provides us with a tool to find potential areas which are not used optimally. This is of particularly interest in afforestation programs. Once the optimum site conditions for tree species have been identified it is relatively easy to search with GIS and find all site conditions in the watershed which fit the required conditions.
3. Dividing the watershed into elevation-slope-aspect combination classes provides a tool to arrive at a crude micro-climatic subdivision of the watershed, and these divisions can then be calibrated using available climatic data.
4. Combining hydrology, sediment data and land use on the GIS system provides us with a means of determining hydrological response to site conditions and land use and this will form a basis to evaluate rates and causes of erosion.
5. The micro-computer based GIS technology is most appropriate for assessing resource problems in watersheds in Nepal since the technology is relatively cheap, user friendly, and produces output in a format that is readily understandable to most resource people and government officials.

REFERENCES:

Carson, B., R. Nield, D.B. Amataya, and G. Hildreth. " Land Resource Mapping Project, Agriculture Forestry Report, Present land Use and the Potential for Improvement. Governments of Canada and Nepal, 82 pp. 1986.

Gilmore, D.A. Not seeing the trees from the forest: A reappraisal of the deforestation crisis in two districts in Nepal. Mountain Research and Development, Vol. 8, pp343-350, 1988.

Griffin, D.M., K.R. Shepherd, and T.B.S. Mahat. Human impact on some forests of the Middle Hills of Nepal. Comparisons, concepts, and some policy implications. Mountain Research and Development, Vol8. pp 43-52.

Hrabovszky, J.P. and K. Miyan. Population growth and land use in Nepal, "The great turnabout". Mountain Research and Development, Vol. 7, pp 264-270, 1987.

Mahat, T.B.S., D.M. Griffin, and R.K. Shepherd. " Human impact on some forests of the Middle Hills of Nepal. Mountain Research and Development, Vol. 6 pp 223-232. 1986.

Mahat, T.B.S., D.M. Griffin, and K.R. Shepherd. "Human Impacts on some Forests of the Middle Hills of Nepal. Part 3. Forest in Subsistence Economy of the Sindhu Palchok and Kabre Palanchok. Mountain Research and Development, Vol. 7, pp 53-70, 1987.

Moench, M. and J. Bandyopadhyay. People-forest interaction: A neglected parameter in Himanayan forest management. Mountain Res. Dev., Vol. 5 pp 3-16, 1986.

Panday, K.K. Fodder trees and tree fodder in Nepal. Swiss Development Corporation & Swiss Federal Institute for Forest Research, Birmenstorf, Switzerland, 107 pp, 1982

MICROPROCESSOR-BASED WATER RESOURCES REQUIREMENTS SIMULATION ANALYSIS

Miguel Ortiz ·
Dr. Ramon Vasquez-Espinosa

University of Puerto Rico at Mayaguez
Department of Electrical and Computer Engineering
Mayaguez, P.R. 00708
(809) 832-4040 X-3094

Rapid Urban Development and population growth in Puerto Rico during the past decades have limited the availability of land and water resources for the needs of the State. Puerto Rico is characterized by a northern coast with plentiful rainfall and short, deep basins, and a southern coast with limited rainfall and long, shallow catchments. A coherent GIS for a selected basin in Puerto Rico will allow the rapid estimation of runoff parameters, determine the connectivity, and estimate crop water requirements from land use and land cover databases. A methodology will be developed to transfer available data from a selected basin into a PC readable form for a water-resources simulation system. The data required for water resources planning is contained within various maps or electronic databases. These information can then be combined or manipulated to simulate different situations, yielding resource allocation plans, forecasts, and future requirements. Finally, an expert system will be developed to guide the user and to make up a smart database that will improve the success of the system by supplying or pointing out missing data. These will automate the identification and quantification of various basin resources like precipitation, land use, soil types and others that may be deemed required.

DEVELOPMENT OF AN ENVIRONMENTAL INFORMATION SYSTEM FOR VENEZUELA

C. Gosselin[1], S.R. Haja[1], R. Simard[1], S.I. Solomon[2], R. Salcedo[3]

1. Digim Inc., Montréal, Canada
2. Solomon and Associates, Toronto, Canada
3. Ministry of Environment
and Renewable Natural Resources
Caracas, Venezuela

ABSTRACT

The implementation of a PC-based environmental information system within the Ministry of the Environment and Renewable Natural Resources of Venezuela was carried out together with the demonstration of application projects in hydrometeorology, agrometeorology and topographical mapping.

The main system functions include: data capture, image analysis, spatial analysis and cartographic data output as well as a specialized hydrometerological modelling software. With this system, currently used techniques can be replicated at higher levels of precision and accuracy and more advanced techniques and models can be more readily tested using both conventional and remotely sensed data.

The results obtained with the hydrometeorological model for a large scarcely gauged tropical river basin are presented in this paper.

INTRODUCTION

A project was undertaken in cooperation with the Canadian International Development Agency (CIDA) to assist the Venezuelan government in the development of an environmental information system (SICA). The results of this project will aid in the establishment of the terms of reference for a full-scale system to be developed for the Ministry of the Environment and Renewable Natural Resources (MARNR).

After a thorough analysis of the needs of the concerned Directorate from this Ministry, the architecture for a prototype PC-based environmental information system was proposed and prioritary applications were selected to demonstrate the capabilities of the system.

One of these four applications projects consisted in the application of a distributed rainfall-runoff model to a large river basin in the tropical area of Venezuela, to assess the possibility of estimating flows from meteorological data.

BACKGROUND

During the last fifteen years or so, hydrometric activities in large areas of Venezuela have been somewhat neglected. Meteorological data, on the other hand, have been collected to a reasonable extent thereby making the indirect estimation of flows the only avenue for generating historical time series of hydrologic data for the above-mentioned areas. The basin selected for the model application is that of the Caura River which, because of its hydroelectric potential, has been less neglected from the viewpoint of hydrometric activities. The location of the basin is shown in Figure 1.

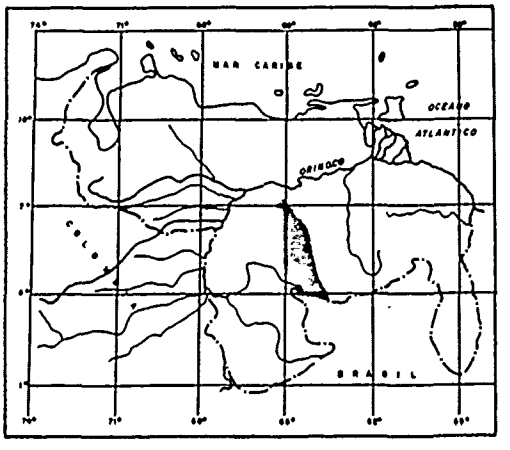


FIGURE 1. Location of study area: Caura River basin

MODEL DESCRIPTION

The distributed rainfall-runoff model used for this application was initially developed for the Ontario Ministry of Housing (Solomon and Associates, 1974) and later adapted to the modelling of water quality (Solomon and Associates, 1976) and of suspended sediments (Gupta and Solomon, 1977, Solomon and Gupta, 1977). However, the model which is the object of this paper refers only to the estimation of flow.

The model requires a square grid data bank of the river basin which can be developed by extracting the data from the corresponding topographic, land-use/land-cover, soils and geological maps. In addition to the latter, inputs to the model are precipitation and temperature data at hourly intervals at all meteorological stations inside or near the given basin and flow data at hourly intervals at all stations inside the given river basin.

Precipitation and temperature data are required for each square of the river basin and the values for a given square at a given time are interpolated using a weighted average technique described in more detail in Solomon and Gupta (1977). A water balance is calculated for the time interval considered in each square. Evaporation is estimated from a formula developed by Turc (1959) and adapted for the purpose of the model. Interception and infiltration are estimated taking into account the soil, land-use/land-cover and geological characteristics. After soil saturation, percolation and runoff are assumed to occur in accordance with the geological and topographical conditions. The runoff is then concentrated into the channel, augmented with contribution from groundwater where appropriate, and routed using a combination of the Muskingum and kinematic wave routing technique (Solomon and Gupta, 1977).

The model is normally calibrated starting from model parameters obtained in other river basins. Each parameter is modified separately in a model run and the modification continued in the same direction if improvement is achieved, or reversed in the contrary case. Only 3 to 4 parameters are usually subjected to the calibration process, although the number of model parameters is larger (9). The parameters usually modified for calibration purposes are :

FC: a parameter representing the percolation rate, its increase reduces the runoff volume and vice versa;

CLAG: the channel lag whose increase reduces the routing velocity and increases the lag between the occurrence of precipitation and of the flood peak resulting from it;

GWRFAC: whose increase produces an increase in the amount of groundwater contributed to the channel flow; and

GWLAG: whose increase slows down the subsurface and groundwater contribution to surface flow and may also be used in conjunction with GWFRAC to fine tune the regression curve.

The calibration process requires an initial priming of the model to establish initial surface, soil and groundwater storages. It is achieved by assuming some arbitrary storage values and running the model for a period of a few days prior to the calibration period. Initial storage values that are too high would be normally corrected by high releases from the storage and vice versa.

The main calibration criterion is the sum of the squares of the errors, i.e. of the differences between the observed and the estimated values. Additional qualitative criteria such as the shape of the hydrograph, the time and value of peak flow, the total runoff volume can also be used.

Validation in space is obtained by comparing the flows estimated by the model in the squares of the validation stations to the observed values. Validation in a neighbouring basin, however, should be preferred when affordable since it represents a completely independent check of the model in the given area.

RESULTS

The data base required to run the model was obtained by digitizing the corresponding 1:500 000 topographic, land-use/land-cover and geological maps and storing the required information in a square grid matrix of 5 X 5 km^2. The flow direction results obtained on the basis of the interpretation of the topographical information within each grid square are illustrated in Figure 2. The absence of data such as soil field capacity and soil depth resulted in the use of uniform values for the whole basin thereby introducing estimation errors.

The model was calibrated using data fromat the station Pie de Salto (drainage area: 29 980 km^2) and validated in space at the staiton Entrerios (drainage area: 14 830 km^2). The calibration and validation in space were carried out for the month of June 1986, and the validation in time was carried out for the months of July and August 1986.

Since only three meteorological stations (Amandina, Entrerios and Pie de Salto) were available, all of them located along the main river, much of the precipitation data is very roughly estimated. Given the strong convectional component of the precipitation in this region, the variation of the precipitation is quite large from one point to another. This is illustrated by the precipitation records shown in Figure 3 for two of the three stations. Similarly, only extreme daily values of temperature are given for those same three stations. The actual time of occurrence of the minimum and maximum temperature was assumed to be 02:00 and 14:00 hrs respectively, and the variation in between was assumed to follow roughly a sine curve. This of course introduces further errors in the model inputs which are, however, much less significant than in the case of precipitation.

The lack of data on river flow from small to large rivers also imposed the use of default values for the hydromorphological relationships necessary for the application of the model. These default values were obtained from previously tested basins in Canada.

Considering all the assumptions and approximations mentioned above the results of the model application in the Caura River basin appear indeed remarkable. The model calibration (Figure 4) resulted in a standard error of 27,3 % (after elimination of the first four days of model priming). The model validation in time, illustrated in Figure 5, carried out for 62 days resulted in a standard error of estimation 28,0%, well consistent with the calibration results. Validation in space, shown in Figure 6) resulted in a standard error of estimation of 26,4 %. This type of performance has been observed in other areas of the world where the model was applied.

CONCLUSIONS AND RECOMMENDATIONS

It may be concluded from the above-mentioned results that the distributed rainfall-runoff model based on the square grid information system can be used as a practical tool for flow synthesis in areas with scarce hydrometric networks but with reasonable

meteorological networks (one meteorological station for every few thousand square kilometres).

Further work is required to improve the application of the model will be facilitated by SICA after its final implementaiton, namely:

a) use of remotely sensed satellite (particularly the meteorological ones) images to obtain a more realistic estimation of the space and time variation of precipitation and temperature;

b) use of remotely sensed data to update the vegetation and land use maps;

c) use of remotely sensed data and conventional data to prepare soil maps providing data on fileld capacity, porosity and soil depth;

d) development of regional hydromorphological relationships;

e) automatic determination of flow direction using a digital terrain model;

f) use of stereoscopic SPOT images for the generation of medium-scale (1:50 000) topographical maps for smaller basins lacking such data.

Work carried out in other similar tropical regions (Shawinigan and Solomon, 1983) indicates that more conventional models to estimate daily flows with errors of the order of 20 to 30% require a much denser hydrometric network (at least one station for every 2 000 km^2). This translates itself for the basin under consideration in a requirement of at least 20 new stations). The cost of operation, in this area, of one hydrometric station is at least 20 000 $US/year and a network has to be operated for at least 5 years before meaningful results can be obtained. It can therefore be safely estimated that the application of the distributed rainfall-runoff model could result in savings of 2 M $US in the case of the Caura River basin. The savings that may result from the extension of the application of the model to the whole tropical area of Venezuela can be estimated at about 30 M $US.

This should not be construed as an exhortation to reduce the operation of hydrometric networks. Higher accuracies than those obtainable by means of the most sophisticated rainfall-runoff models are needed in many areas, and direct measurements are irreplaceable in all such cases. This is merely an appeal to the rational use of limited available funds for operating information systems on environmental variables in a coordinated manner, with the rational and timely transfer of information between related fields. This is of particular importance when time is an essential factor and data of lesser accuracy are acceptable for planning and pre-feasibility investigations.

REFERENCES

Gupta, S.K., 1974. A Distributed Digital Model for Estimation of flows and Sediment Load from Ungauged Watersheds, Ph.D. Thesis, Dept. of Civil Engineering, Univ. of Waterloo, Ont.

Gupta, S.K. and S.I. Solomon, 1977. Distributed Numerical Model for Estimating Runoff and Sediment Discharge of Ungauged Rivers, 1: The Information System, Water Res. Research, Vol.13, No.3, pp.613-618.

Shawinigan and S.I. Solomon and Associates, 1983. A Square Grid Hydrological Study of the Amazon River Basin, Report for UNDP/WMO.

Solomon, S.I. and Associates, 1974. A Rainfall-Runoff Model for Environmental Impact Assessment in the Basins of the Rouge, Little Rouge, Duferin, and Peticoat Rivers, Report for the Ministry of Housing, Toronto, Ontario.

Solomon, S.I., and Associates, 1976. Application of WATMAP-WATFILE Data System for the Development of a Distributed Water Quantity-Water Quality Model for South Nation River Basin, Rep. for Environment Canada, Water Res. Data Syst. Section, Planning and Management Branch, Ottawa, Ontario.

Solomon, S.I. and S.K. Gupta, 1977. Distributed Numerical Model for Estimating Runoff and Sediment Discharge of Ungauged Rivers, 2: Model Development, Water Res. Research, Vol. 13, No. 3, pp.619-629.

Turc, L., 1959. Le Bilan d'Eau des Sols, Relations entre les Précipitations, l'évaporation et l'écoulement, Ann. Agron.,5.

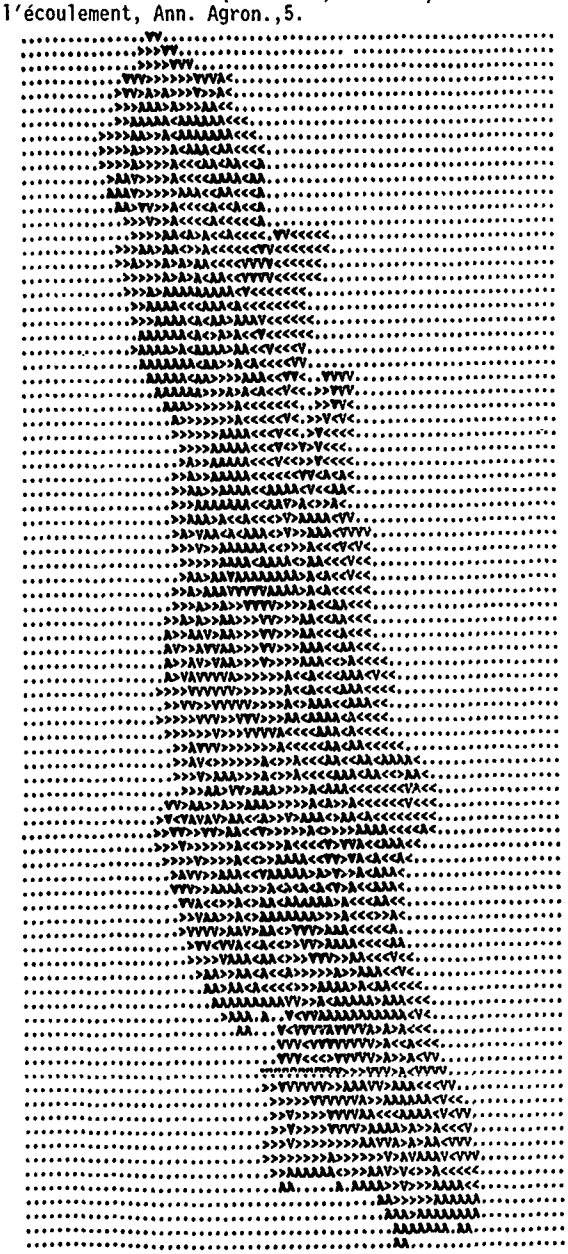
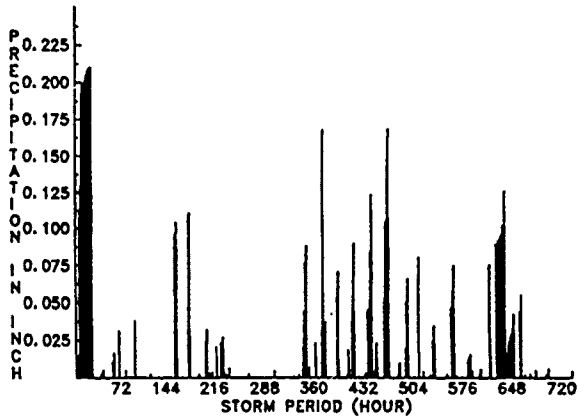
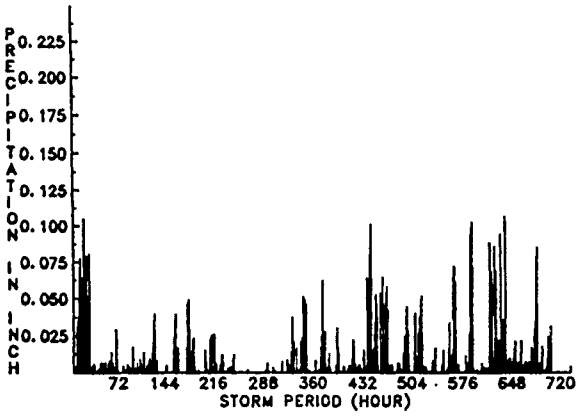


FIGURE 2. Flow direction calculated for each grid square.

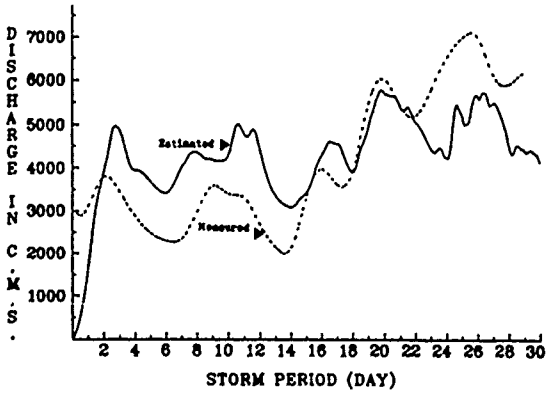


Station: Pie de Salto
Duration: 86-06-01 to 86-06-30



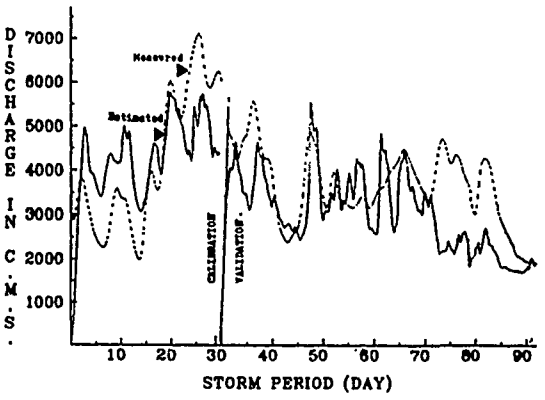
Station: Entrerios
Duration: 86-06-01 to 86-06-30

FIGURE 3. Hourly precipitation records



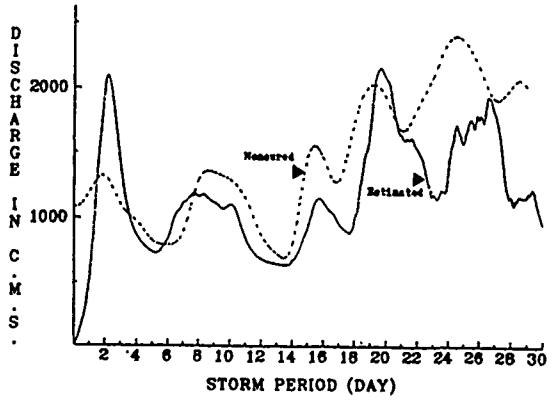
Station: Pie de Salto
Duration: 86-06-01 to 86-06-30

FIGURE 4. Model calibration



Station: Pie de Salto
Duration: 86-06-01 to 86-08-31

FIGURE 5. Model validation in time



Station: Entrerios
Duration: 86-06-01 to 86-06-30

FIGURE 6. Model validation in space

THE DEVELOPMENT OF A RESOURCE INFORMATION MANAGEMENT SYSTEM A PILOT PROJECT IN NORTHERN THAILAND

P. Vincent, M. Beaudoin, C. Prévost, G. Lemieux and R. Simard

Digim inc.
1100 René Lévesque Blvd. West
Montreal, Quebec
H3B 4P3

S. Vilbulsresth
National Research Council of Thailand
Bangkok, Thailand

ABSTRACT

The development of remote sensing applications in Thai agencies involved in natural resources management is currently being supported by the Canadian International Development Agency through a technology transfer program.

The first phase of the project consisted of basic training and demonstration projects in which remote sensing technology was applied to offer suitable (alternative) solutions to the production of maps using conventional methods. The second phase involved the use of satellite-derived information in combination with conventional data for the management of natural resources. The aim of this second phase was to generate integrated land and resource management alternatives.

A micro-based computer system (SPANS) was used to build the data base and perform data modelling. Information on phenomena of a dynamic nature such as land use change, derived from satellite imagery, as well as information on phenomena of a more static nature such as geology and topography was integrated in the data base using computer interfaces and digitizing procedures. The crop suitability evaluation is the field of application to be presented. The study area selected to carry out this project corresponds to a rectangle astride the Chiang Mai basin in northwestern Thailand.

The application chosen served to demonstrate, in a Thai context, the potential of remote sensing and GIS technology in improving the management of natural resources. This model could be further apllied to the whole country.

INTRODUCTION

The realization of a pilot project relating to the development of a resource information management system (RIMS) in northern Thailand is part of a large training and transfer of technology program in the field of remote sensing, supported for the last three years by the Canadian International Development Agency (Rochon et al., 1988).

During phase I, 23 representatives from various Thai ministries, including the National Research Council of Thailand, have been involved in data gathering and transfer, cartography from satellite image, and the preparation of geo-referenced data for the management of the country's natural resources and land development.

The second phase was concerned with the realization of various demonstration projects in agriculture, forestry, topography, water and land resources, all sectors in which remote sensing technology offers a promising alternative to conventional mapping methods. This phase was concluded by the installation of a natural resources management information system.

The objective of this last demonstration project was to provide concise information that would help decision-makers assess the present state of natural resources and thus improve the management process. The integrated management of resources has been the guiding principle to the setting up of a biophysical data base made up of existing maps, record data, and most of all data obtained from satellite imagery taken during phase I. A resource management information system in a data processing environment proved to be an excellent tool for coordination between specialists in charge of satellite data and decision-makers responsible for broad policies relating to the management of national resources.

METHODOLOGY

Chiang Mai in northern Thailand was the site selected for the RIMS project (figure 1). It is 85 km wide across the Chiang Mai valley and extends 50 km south from the town of Chiang Mai. It is a flat valley surrounded by high mountains degraded by deforestation. Among those the 2 590 m high Mount Doi Ithanon is the highest mountain in Thailand.

Some maps drawn on a scale ranging from 1:50 000 to 1:250 000, were integrated into the data base. They included the geology and geomorphology maps; the topography map; the drainage and water system map; district and road maps and temperature and rain data taken by local meteorological stations.

Maps obtained during Phase I of the project completed the preliminary data set. They comprised a forestry map generated through the automatic classification of LANDSAT MSS imagery, land cover and land use maps drawn up from a SPOT multiband imagery.

The above data were integrated into the digital data

base by means of various interfaces as shown on figure 2. The SPANS spatial analysis system was used for data modelling and analysis. Most of the existing maps were digitized by means of a digitizing table. Digital maps derived from satellite imagery and the digital terrain model obtained from the digitization of contour lines, were transferred through electronic links and commonly geo-referenced into the SPANS data base.

Many fields of applications were examined for the project. The example presented herein refers to a land evaluation application by the production of maps establishing the potential for agriculture involving more than 20 differents crops.

LAND EVALUATION APPLICATION

The main goal of this application was to establish the land suitability ratings for rainfed agriculture. Many crops similar to those growing in the area were studied. Other crops, such as coffee and tea, that could cope with the environmental conditions found in the area studied, were evaluated for further alternatives to actual land use or agricultural practices, especially where shifting cultivation is common practice.

The land evaluation methodology is based on the one adopted by the Food and Agriculture Organization of the United Nations (FAO, 1988). This methodology, recognized and used by almost all nations, provides guidelines for this type of task. In Thailand, this approach has been successfully tested over many years and adapted to fit specific environments throughout the country.

A major output of this land evaluation is a map representing the suitability classes for specific crops according to the land qualities or potential of the study area. Information for such a task can be easily obtained from the Resource Information Management System data base since the primary data are already entered. The computer makes over-lay combinations less time-consuming and more flexible as parameters and weighting factors can be changed and tested before the final map production.

STEP BY STEP PROCEDURE

Figure 3 shows the most important stages followed to establish the suitability ratings and the inputs required in the preparation of resource management plans.

The first step was to determine the land qualities considered for the purpose of the evaluation. They represent criteria for the analysis of each land unit. Figure 4 shows a list of the land qualities that were taken into account, the diagnostic factors or land characteristics that have been used for their measurement, the measuring units and the value differences observed in the study area as well as the cartographic data source.

Once the land qualities had been determined, crop requirements for each type of crop were compiled. The crops listed in the following table were analyzed:

cocoa	cotton	ground nut
vegetable	palm tree	sugar cane
flooded rice	pineapple	coconut
cashew nut	tea	banana
pepper	soya bean	coffee
maize	pasture	sweet potato

TABLE 1 : CROPS RETAINED FOR THE LAND EVALUATION

Data were compiled for each land characteristic by establishing a value difference corresponding to four suitability classes described in the next table. Figure 5 shows the crop requirement values for tea.

S1 : high
S2 : medium
S3 : low
N : unsuitable

TABLE 2 : CROP SUITABILITY CLASSES

While crop requirement was established, a data base was developed and a land unit map with a corresponding value for each land characteristic was produced. The next step consisted in matching each crop requirement and land unit characteristic of the study area so as to establish the crop suitability map (high, medium, low and unsuitable), for a specific crop. Figure 6 shows the procedure followed to generate such a map.

A comparison between the crop suitability map and the present land use, helped to identify areas that can be developed for other purposes or cultivated in a different way. Figures 7 and 8 respectively show a typical land unit map and the map of land suitability for the tea crop.

RESULTS AND DISCUSSIONS

This pilot project has served to establish a comprehensive remote sensing data base for an area of more than 4 250 km^2. The application chosen, the crop suitability evaluation of land, served to demonstrate the advantage of applying the integrated approach of RIMS to the development and management of land resources.

RIMS is a very convenient technology for environmental specialists who have a huge volume of data to process. They can thus provide a more pertinent and clearer information to decision-makers. The manager, responsible for the region in the above example, has all the necessary information to evaluate the differences between the present and the optimal land use. This type of information is often hard to obtain when the data base processing technology is inadequate.

CONCLUSION

Specialists trained through the technology transfer program have been able, during this demonstration project, to integrate into the same data base many of their maps produced using satellite imagery and to take part in a multidisciplinary analysis for the development of natural resources in the study area.

Moreover it has been demonstrated that RIMS can combine conventional data with remote sensing data to produce relevant information that is very important to decision-makers.

We will never emphasize enough the strong synergy between remote sensing and RIMS technologies. It particularly worths when applicable in part of the world where urgent environmental actions are to be considered with minimum financial resources.

ACKNOWLEDGEMENTS

The authors wish to express their gratitude to the participants in this project, in Thailand and in Canada for their recognized contribution to its success. They are also grateful to the Canadian International Development Agency for its support and useful collaboration in the realization of the project.

REFERENCES

FOOD AND AGRICULTURAL ORGANIZATION. 1988. Guidelines: Land Evaluation for Rainfed Agriculture. No. 52, Rome, Italy.

ROCHON, G., R. Simard, C. Prévost, P. Vincent, S. Vibulsresth and D. Srisaengthong. 1988. "Thai-Canadian Cooperation in Remote Sensing Application Projects - Improved Access to Natural Resources Information." Proceedings of the ninth Asian Conference on Remote Sensing, Bangkok, Thailand.

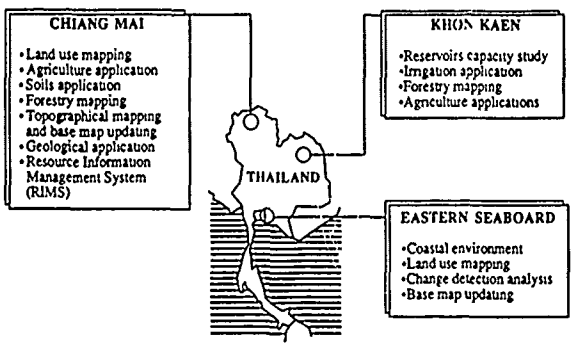


FIGURE 1 : LOCATION OF THE DEMONSTRATION PROJECT STUDY AREAS

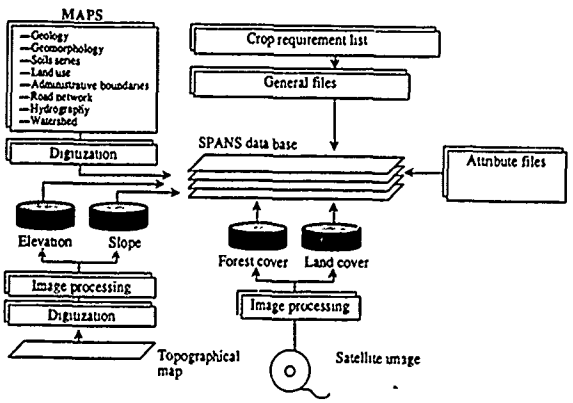


FIGURE 2 : FLOW CHART OF THE DATA BASE CONTENTS AND THE DATA INPUT PROCEDURE

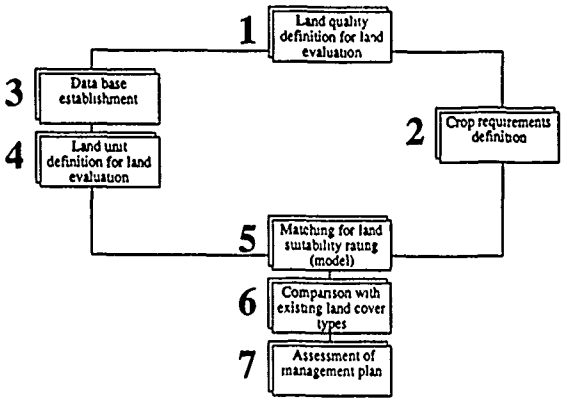


FIGURE 3 : METHODOLOGY OF THE LAND SUITABILITY EVALUATION FOR CROP

MAP ID	LAND QUALITY	DIAGNOSTIC FACTOR	UNIT	RANGE
Mean annual rainfall 1951-1970	Moisture availability	Total rainfall	mm	0 — 1750
(Agricultural stats) + topographical map	Temperature regime	Annual average temperature	°C	7 — 26
	Terrain characteristic	Slope	%	0 — 50
Soils series (geology and geomorphology)	Oxygen availability	Soil drainage	Level	6 classes (Qualitative)
	Rooting condition	Effective soil depth	cm	0 — 150
	Nutrient availability	Soil reaction	pH	4.5 — 8.0
		Available phosphorus	ppm	0 — 45
		Available potassium	ppm	0 — 120
	Nutrient retention	Texture	Type	15 classes (Qualitative)
		Cation exchange capacity (CEC)	meg 100g	0 — 30

FIGURE 4 : LAND QUALITY DATA CHARACTERISTICS

CROP IDENTIFICATION: LAND UTILIZATION TYPE:				
Land characteristics	Factor rating Suitability classes			
	S1	S2	S3	N
Temperature regime Annual mean T°	19-22	22-24 17-19	24-27 14-17	27 14
Moisture availability Annual average rainfall	25-40 (dm)	18-25 40-50	13-18 50-60	> 60 <13
Terrain slope (%)	0-8	8-16	16-35	> 35
Oxygen availability Drainage	well	m.well s exce	poor s.poor	v.poor exces.
Rooting condition Depth (cm)	>150	100-150	50-100	< 50
Nutrient availability pH	4.5-5	6-5.5 4.5-4	5-5.5 4-3.5	6.5 3.5
Nutrient availability Phosphorus, P (ppm)	> 15	6-15	3-6	< 3
Nutrient availability Potassium, K (ppm)	>30	< 30		
Nutrient retention Texture	l,scl, sil,cl	sl,sc	ls,slc	s,c
Nutrient retention Cation exchange capacity	> 5	< 5		

FIGURE 5 : LIST OF THE CROP REQUIREMENTS FOR TEA

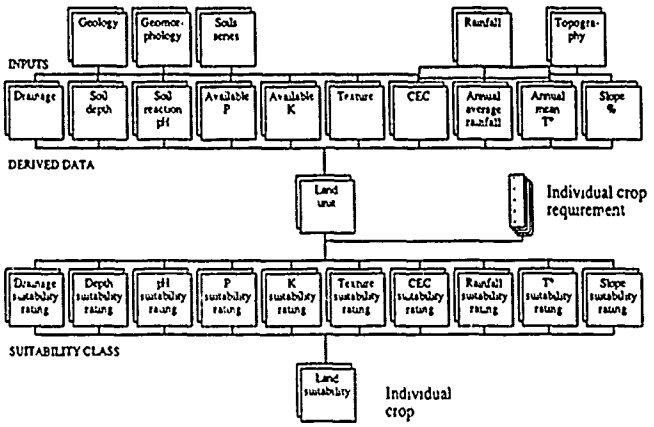


FIGURE 6 : FLOW CHART OF THE LAND SUITABILITY MAP GENERATION

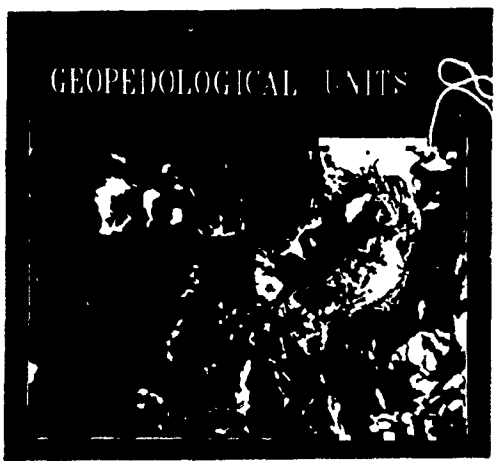


FIGURE 7 : LAND UNIT MAP COMBINING SOILS, GEOMORPHOLOGY AND GEOLOGY

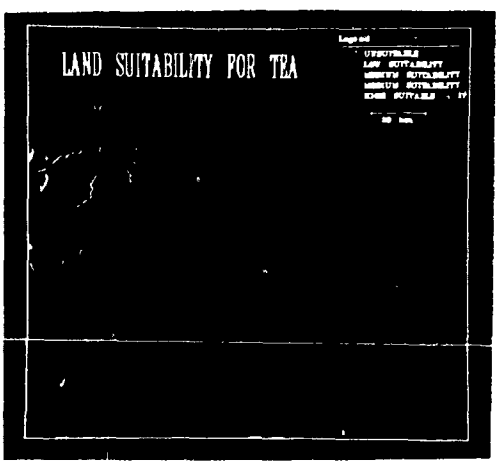


FIGURE 8 : LAND SUITABILITY MAP FOR TEA

A NEW AIRBORNE SAR FOR ICE RECONNAISSANCE OPERATIONS

by

J. BRYAN MERCER
INTERA TECHNOLOGIES LTD.
#2500, 101-6th Avenue S.W.
Calgary, Alberta, Canada T2P 3P4
Telephone: (403) 266-0900
Fax No: (403) 265-0499

ABSTRACT

In order to improve its capability to provide ice information to the Canadian Coast Guard (CCG) and to fulfill other ice-related requirements within its mandate, the Atmospheric Environment Service (AES) of the Canadian Government has awarded a contract for the development and operation of a new SAR-based, airborne reconnaissance system to INTERA Technologies Ltd. This paper will describe the performance specifications and operational profile of the system, which has been under development since mid-1987 and is scheduled to be operational by July-1989.

Data acquisition is based upon dual side-looking X-band, digital, real-time, multi-look SAR's carried aboard an executive jet aircraft. In its wide-swath mode (100 km per side), it is capable of imaging more than 600,000 km^2 per mission from altitudes typically above 11 km.

Data delivery requirements necessitate three modes of on-board data processing, which are accomplished in a custom-design Data Management Unit.

In the first mode, data are transmitted to CCG vessels at full resolution in near-real time over an S-band data-link where they are received, displayed, manipulated and stored on an image analysis system referred to as STARVUE. Among other features, the imagery can be approximately geo-referenced using transmitted navigation data and through split screen display, can be presented as an overview plus local enlargement. In this mode, the imagery will be used in direct tactical support of ten CCG vessels and three CCG offices.

The second delivery mode involves compression of the data prior to downlink to Alert ground stations which re-transmit the data via ANIK satellite to the AES Ice Centre in Ottawa. The data are received, decompressed and quality controlled at this point prior to being handed over to the Ice Centre computer for subsequent analysis.

The third mode requires that the data be archived, on disk initially, and subsequently on high volume tape media. This raw data will ultimately serve the archival requirements of Ice Centre.

The system has a minimum data acquisition requirement of 70×10^6 km^2/year over the six year life of the contract. In fact, this represents less than 50% of its design capacity. In response to AES needs, the aircraft will be based in varying locations in the Arctic and East Coast, but due to its speed and range, it will be capable of missions far removed from the operational bases.

INTERNATIONAL ICE PATROL SIDE-LOOKING AIRBORNE RADAR (SLAR) EXPERIMENT 1988

Stephen R. Osmer and Michael A. Alfultis

International Ice Patrol
U. S. Coast Guard
Avery Point
Groton, CT 06340

Abstract

During the period 7 through 16 June 1988, the International Ice Patrol conducted an evaluation of the AN/APS-131 Side-Looking Airborne Radar (SLAR). This SLAR, installed as part of the multi-sensor surveillance AIREYE system onboard the U. S. Coast Guard HU-25B medium endurance aircraft, was evaluated for its ability to detect icebergs. The data collection occurred in an iceberg infested area off the coast of Newfoundland, Canada.

The fundamental goal of this research was to provide guidance on the ability of the AIREYE-equipped HU-25B to perform the iceberg detection mission of the International Ice Patrol. Specifically, there were two objectives:

1. Determine the optimum altitude for iceberg reconnaissance, and predict the probability of detection as a function of sea state, lateral range, and iceberg size.

2. Compare the iceberg detection capability of the AN/APS-131 SLAR with the AN/APS-135 SLAR currently used on the International Ice Patrol's HC-130 long range reconnaissance aircraft.

Ground truth (i.e. iceberg dimensions and positions, and environmental conditions) were collected by the U. S. Coast Guard icebreaker NORTHWIND (WAGB 282). The HU-25 and HC-130 aircraft flew a box pattern around the iceberg search area. Several different altitudes were used.

The Ice Branch of the Atmospheric Environment Service of Canada also had two of its SLAR-equipped ice reconnaissance aircraft (an Electra and a Dash-7) participate in the experiment.

Results indicate the AN/APS-131, while not having the azimuth resolution of the AN/APS-135, is capable of performing the iceberg reconnaissance mission. These preliminary results indicate an altitude of 4000 to 6000 feet is best for the AN/APS-131 for this mission.

Current plans for the 1989 iceberg season are for the HU-25B to compliment the HC-130H reconnaissance aircraft. Due to its limited endurance, the HU-25B aircraft will not be able to replace the longer-range HC-130. However, during certain times of the year and in certain light ice years, the HU-25B should be able to conduct the International Ice Patrol mission.

Key Words

Aerial Ice Reconnaissance
Iceberg
International Ice Patrol
Side-Looking Airborne Radar
SLAR

Introduction

After the sinking of the RMS TITANIC on April 15, 1912, an International Ice Patrol Service was created to monitor the presence of icebergs near the Grand Banks of Newfoundland, and to warn mariners of these hazards. The International Ice Patrol (IIP), a unit of the U. S. Coast Guard, has provided this service since its initiation in 1914. From 1914 to 1945, IIP used visual reconnaissance from ships to monitor the icebergs. After World War II, and up to 1983, IIP used aircraft visual reconnaissance as its primary method of iceberg detection. Since 1983, IIP has utilized a Motorola AN/APS-135 Side-Looking Airborne Radar (SLAR) onboard HC-130H Hercules long-range aircraft as its primary method of iceberg reconnaissance.

In 1983, the U. S. Coast Guard installed the Motorola AN/APS-131 SLAR as part of the airborne multi-sensor surveillance AIREYE system on its HU-25B Falcon medium-range aircraft. The AN/APS-131 SLAR is very similar to the AN/APS-135 SLAR on the HC-130, except that the antenna length of the APS-131 is half that of the APS-135. This results in the APS-131 having a lower azimuth resolution than the APS-135. Although the iceberg detection ability of the APS-135 SLAR has been previously evaluated, no evaluation of the iceberg detection ability of the APS-131 SLAR has been made.

This report presents the results of an evaluation of the AN/APS-131 SLAR to detect icebergs. This evaluation was

conducted by IIP from 7 to 16 June 1988 in the North Atlantic Ocean off Newfoundland, Canada. The fundamental goal of this research was to provide guidance on the ability of the AIREYE-equipped HU-25B to perform the iceberg detection mission of the International Ice Patrol. Specifically, there were two objectives:

1. Determine the best altitude for iceberg searches, and predict the probability of iceberg detection as a function of sea state, lateral range, and iceberg size.

2. Compare the iceberg detection capability of the APS-131 SLAR with the APS-135 SLAR.

This report will also compare the results of this evaluation with the results of two previous SLAR iceberg detection evaluations.

Background

Previous SLAR Studies

Two previous SLAR studies have been conducted to evaluate the ability of the AN/APS-135 SLAR to detect icebergs. During April 1984, BERGSEARCH '84 was conducted to evaluate the ability of three SLARs and two Synthetic Aperture Radars (SAR) to detect small icebergs and growlers. The M/V POLARIS provided surface truth data. Results of the data analysis reported in Rossiter et al (1985) show greater detectability is obtained with lower sea states, at lower altitudes within the operating envelope of each system, and when viewing targets across rather than up or down wind and sea. BERGSEARCH '84 data also demonstrated that ships and iceberg targets generally do not have different SLAR signatures.

The 1985 SLAR Detection Experiment was designed to determine SLAR's ability to detect various search and rescue and iceberg targets at all ranges out to 50 km. The iceberg detection results reported in Robe et al (1985) indicate medium icebergs are detectable nearly 100% of the time in up to 2 m seas, small icebergs are easier to detect at lower altitudes and with a smaller swath width, and growlers are detectable more than 90% of the time in seas less than 1 m. Also, both growlers and small icebergs in seas less than 1 m appear to be just as detectable at lateral ranges between 25km and 50 km as they are at ranges less than 25 km. Finally, they noted similar iceberg detection performance of the AN/APS-135 SLAR in this experiment and in BERGSEARCH '84.

Description of Aircraft

The HC-130H is a long-range four engine turboprop reconnaissance aircraft, whereas the HU-25B is a medium-range twin engine fan jet aircraft. CG-1503 from Coast Guard Air Station Elizabeth City, North Carolina, was the HC-130H aircraft in the experiment, and CG-2103 from Coast Guard Air Station Cape Cod, Massachusetts, was the HU-25B.

Description of AN/APS-135 and AN/APS-131 SLAR

Significant system parameters of each SLAR are presented in Table 1. The major difference between the two systems is the antenna length of each. The APS-135 has a 4.8 m long antenna, while the APS-131's antenna is 2.4 m long. This results in the APS-131 having a lower azimuth resolution than the APS-135.

Table 1. SLAR Operating Parameters.

Aircraft	HC-130H	HU-25B
SLAR (Real Aperture)	Motorola AN/APS-135	Motorola AN/APS-131
Frequency	X-Band (9250 MHz)	X-Band (9250 MHz)
Peak Power	200 Kw	200 Kw
Pulse Width	0.2×10^{-6} sec	0.2×10^{-6} sec
Antenna Characteristics		
Length	4.8 m	2.4 m
Polarization	VV	VV
Elevation Coverage	-1.5 to -45 deg	-1.5 to -45 deg
Depression Angle	1.5 deg	1.5 deg
Azimuth Resolution	0.47 deg	0.8deg
Range Resolution	30 m	30 m
Receiver Bandwidth	6 MHz	6 MHz
Swath Widths	25,50,100,150 km	25,50,100,150 km
Look Direction	L & R	L & R
DataFormat	Negative Film	Negative Film VHS video tape

Description of Targets

The USCGC NORTHWIND was the only surface vessel used as a SLAR target during the evaluation. The NORHTWIND is a U.S. Coast Guard wind-class icebreaker, and is 82 m long and 19 m wide.

IIP classifies icebergs into five size categories: growler, small, medium, large, and very large. Table 2 lists IIP's length and height parameters for each size category. A total of 44 icebergs were used as targets during the evaluation. No growlers were used. 13 of the iceberg targets were small icebergs, 27 were mediums, and 4 were large. No very large icebergs were used.

Table 2. IIP Iceberg Size Categories

Descriptive Name	Height (m)	Length (m)
Growler	<5	<15
Small Iceberg	5-15	16-60
Medium Iceberg	16-45	61-122
Large Iceberg	46-75	123-213
Very Large Iceberg	>75	>213

Description of Environmental Conditions

Seas were generally 1-2 meters during most of the experiment.

Data Collection Procedures

General

Data for this evaluation were collected on 7-16 June 1988. The exact location of data collection varied with ice movement within a box bounded by 51 N to 52 N and 52-30 W to 53-30 W (Figure 1). Each evening, IIP personnel on USCGC NORTHWIND were responsible for selecting the next day's area of study around a group of iceberg targets, and passing the study area coordinates to the aircraft using VHF radio prior to the aircraft's departure. The next morning, the aircraft would confirm the location of NORTHWIND, and the study area, after takeoff.

During the data collection runs, the IIP crew on NORTHWIND monitored the positions of each iceberg in the study group, and recorded surface environmental data.

Search Patterns

Two search patterns were used by the aircraft during the evaluation. Most of the searches were an area (type 1) search consisting of a square with 54 nm (100 km) sides. Both the HU-25B and HC-130H flew this pattern at four altitudes: 4000, 6000, 8000, and 10,000 feet. All these searches were conducted using the 27 nm (50 km) range scale on the SLAR.

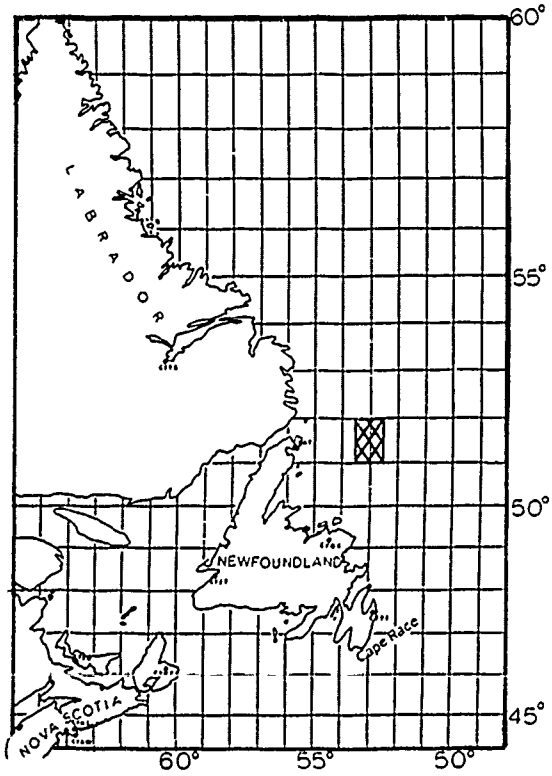


Figure 1. Cross-hatched area depicts study area.

Only the HU-25B flew the second search pattern. It was a parallel line (type 2) search. This pattern was flown at 4000 and 8000 feet. The track spacing was 6 nm (11 km) for the first three legs, and 12 nm (22 km) for the last two legs. Both the 27 nm (50 km) and 54 nm (100 km) range scale on the SLAR were used on this pattern. Because of the limited amount of data collected with the 54 NM scale, no discussion on the use of this scale for ice reconnaissance will be made in this report.

SLAR Data Format

The output analog imagery from both the APS-131 and APS-135 SLAR was recorded on 23 cm wide dry-process film. The HU-25B also used a video recorder to record the analog imagery in VHS tape format.

Results and Discussion

Table 3 summarizes the total number of detection opportunities for each SLAR. Table 4 shows the distribution of the detection opportunities with altitude and lateral range for each SLAR. For the APS-135, the detection opportunities were distributed evenly between the four altitudes 4,000, 6,000, 8,000, and 10,000 feet and in the mid to far-range (5 to 27 NM). For the APS-131, most of the detection opportunites were at 4,000 and 8,000 feet and in the far range (15-27 NM).

Table 3. Total Detection Opportunities
(Type 1 and Type 2 Searches, 27 and 54 NM Scales)

	HU-25B (APS-131)	HC-130H (APS-135)
S	133	48
M	230	132
L	45	17
Northwind	76	31
Total	484	228

Table 5 lists the number of SLAR detections over the number of detection opportunities, and the probability of detection (POD), for each SLAR as a function of iceberg size (and for the NORTHWIND). The two SLARs have a very similar iceberg detection capability. The iceberg POD (for small, medium, and large icebergs) for each SLAR is 99 percent.

Table 6 lists the iceberg POD as a function of lateral range and altitude. Again, both SLARs have a similar iceberg detection cabability at all altitudes and lateral ranges. There is no significant variation in iceberg POD with lateral range or altitude. There is a small decrease in iceberg POD at 8,000 and 10,000 feet for the APS-131 SLAR, however.

These results are similar to the results obtained during BERGSEARCH '84 and the 1985 SLAR Detection Experiment, for the given type of targets and sea conditions.

Table 4.

HC-130H (APS-135) Total Detection Opportunities
(Type 1 Searches)

	Range (NM)					
Altitude (FT)	0-4	5-9	10-14	15-19	20-27	Total
4,000	1	15	14	7	18	55
5,000	0	0	0	0	0	0
6,000	0	14	12	14	20	60
8,000	0	13	8	15	17	53
10,000	0	13	13	15	19	60
12,000	0	0	0	0	0	0
Total	1	55	47	51	74	228

HU-25B (APS-131) Total Detection Opportunities
(Type 1 and Type 2 Searches)

	Range (NM)						
Altitude (FT)	0-4	5-9	10-14	15-19	20-27	27-50*	Total
4,000	16	27	19	31	29	0	122
5,000	0	9	2	11	11	0	33
6,000	2	16	10	17	17	0	62
8,000	8	37	39	38	52	28	202
10,000	3	1	5	10	22	0	41
12,000	0	0	0	6	0	18	24
Total	29	90	75	113	131	46	484

*27-50 NM ranges for Type 2 Searches only.

Table 5. Detection Data (Type 1 and Type 2 Searches, 27 NM Scale)

	Detections/Opportunities (POD)	
	HU-25B APS-131	HC-130H APS-135
Small	114/116 (.98)	47/48 (.98)
Medium	200/203 (.98)	132/132 (1.00)
Large	36/36 (1.00)	17/17 (1.00)
Total	350/355 (.99)	196/197 (.99)
Northwind	53/53 (1.00)	31/31 (1.00)

Table 6. Iceberg POD (Type 1 and Type 2 Searches, 27 NM Scale)

Range (NM)	HC-130H APS-135	HU-25B APS-131
0-4	1/1 (1.00)	24/24 (1.00)
5-9	53/53 (1.00)	82/83 (.99)
10-14	42/42 (1.00)	62/63 (.98)
15-19	43/43 (1.00)	85/88 (.97)
20-27	57/58 (.98)	97/97 (1.00)

Altitude (FT)	HC-130H APS-135	HU-25B APS-131
4,000	47/47 (1.00)	105/105 (1.00)
5,000	-	29/29 (1.00)
6,000	52/52 (1.00)	52/52 (1.00)
8,000	46/46 (1.00)	130/134 (.97)
10,000	51/52 (.98)	31/32 (.97)
12,000	-	3/3 (1.00)

Conclusions

For the given sea conditions and size of targets, all altitudes and lateral ranges out to 27 NM appear suitable for iceberg searches. For the APS-131 SLAR, an altitude of 6,000 feet and lower appears to be a slightly better altitude for iceberg reconnaissance than 8,000 feet or higher. More study is needed at higher sea states and with smaller ice targets before any final conclusions can be drawn regarding the optimum iceberg search altitude and the POD as a function of sea state, lateral range, and iceberg size.

The APS-131 is very similar in its iceberg detection capability to the APS-135. These results indicate it is capable of performing the iceberg reconnaissance mission of the International Ice Patrol.

Current plans for the 1989 iceberg season are for the HU-25B to complement the HC-130H reconnaissance aircraft. Due to its limited endurance, the HU-25B aircraft will not be able to replace the longer-range HC-130H. However, during certain times of the year and in certain light ice years, the HU-25B should be able to conduct the International Ice Patrol mission.

Bibliography

Rossiter, J.R., L.D. Arsenault, E.V. Guy, D.J. Lapp, and E. Wedler, "BergSearch '84: Assessment of Airborne Imaging Radars for the Detection of Icebergs. Summary Report - Results, Conclusions, and Recommendations," Environmental Studies Revolving Funds Project, CANPOLAR Consultants, Ltd, Toronto, Ontario, 1985.

Robe, R.Q., N.C. Edwards, Jr, D.L. Murphy, N.B. Thayer, G.L. Hover, and M.E. Kop, "Evaluation of Surface Craft and Ice Target Detection Performance by the AN/APS-135 Side-Looking Airborne Radar (SLAR)", U.S. Coast Guard, Washington D.C., 1985.

LIGHTWEIGHT AIRBORNE RADAR FOR MARITIME SURVEILLANCE

David C. Cantrill
Paul A. Gorton

Litton Systems Canada Limited
25 City View Drive
Etobicoke, Ontario
Canada, M9W 5A7

ABSTRACT

An overview of the features of the Litton APS-504(V)5 Maritime Surveillance Radar System is presented and a fully coherent variant of the APS-504(V)5 Transmitter/Receiver subsystem which has been developed for Synthetic Aperture Radar (SAR) applications described. The application of coherent techniques to the detection of moving targets is then discussed and Pulse Doppler processing is shown to provide an effective means of meeting requirements for moving target detection. Lightweight technologies are identified which enable a new maritime surveillance radar, with capabilities previously available only in larger and more expensive systems, to be provided in a system which is suitable for installation in a small, twin-engined turboprop aircraft or helicopter. The Litton Lightweight Radar (LWR), currently under development, is based on the demonstrated technology of the coherent variant of the APS-504(V)5 Transmitter/Receiver subsystem and the proven APS-504(V)5 signal processing algorithms. In addition, a Pulse Doppler processor is being included.

Keywords: Airborne Maritime Surveillance Radar

INTRODUCTION

Airborne radar provides a powerful means of satisfying a wide variety of maritime surveillance requirements which cannot be met by surface-based alternatives. In particular, the 360-degree, scanning radar provides rapid, wide-area coverage in real-time and is suitable for a variety of maritime surveillance missions, including:

- Search and Rescue
- Policing of maritime traffic
- Fisheries protection
- Surveillance and control of the exploitation of the continental shelf
- Customs and immigration surveillance
- Detection of ships discharging at sea
- Surveillance of all sources of pollution
- Checking maritime navigation aids
- Iceberg patrol
- Oceanography
- Ensuring compliance with the rules of navigation

Many of these applications reflect the importance of monitoring Territorial Waters (12 nautical miles wide) over which the littoral state has complete sovereignty and the Exclusive Economic Zone (EEZ) (188 nautical miles wide) over which the littoral state retains a certain number of exclusive rights as regards exploitation and exploration (Salvy, 1986). The above missions therefore require the surveillance of long stretches of coastal waters for which economical, medium-range maritime patrol aircraft with low-operating costs are ideal. A wide variety of such aircraft are fitted with Litton radars and such systems are in use in every continent from the Arctic to the Tropics.

The physical dimensions and weight of a maritime surveillance radar are required to be compatible with the payload capacity of medium-range maritime patrol aircraft types, small, twin-engined turboprop aircraft, commuter jets, aerostats and helicopters. The constraints on the size and weight of the radar in maritime surveillance applications to some extent dictate the characteristics and specifications of the radar. The most recent Litton maritime surveillance radar (designated APS-504 (V)5 in Canada and AN/APS-140(V) in the U.S.) is a 360°, scanning, X-band radar with a peak power of approximately 8 kilowatts. (The maximum effective peak power with pulse compression is approximately 4 megawatts). The APS-504(V)5 provides a long-range maritime surveillance capability to 200 nautical miles and includes features which are designed to provide detection of small targets in sea clutter at shorter ranges. In addition, the APS-504(V)5 provides weather avoidance and beacon interrogation modes and a (real-beam) land mapping mode.

The capability of detecting aircraft and small, high-speed sea-surface moving targets in clutter is of increasing importance for maritime surveillance applications. It is now possible to provide an economical radar which has the means of reliably

detecting such targets using newly available technologies. Furthermore, the additional capabilities can be included in a radar comparable to the APS-504(V)5, with a significant reduction in weight.

LITTON APS-504(V)5

The APS-504(V)5 is an advanced, frequency-agile, X-band pulse compression radar which provides the following features:

- Travelling Wave Tube (TWT) based transmitter
- Two pulse compression modes with compression ratios of 500:1 and 210:1
- Wide bandwidth frequency agility (500 MHz)
- Advanced Surface Acoustic Wave (SAW) technology
- Two or three-axis antenna pedestals with multiple antenna scan rates up to 120 rpm
- Low sidelobe, broadband, parabolic or flatplate antennas
- Sector scan with selectable range delay
- Low noise receiver
- Pulse-to-pulse and scan-to-scan integration with Constant False Alarm Rate (CFAR) processing
- High-resolution, TV type multi-function raster display with alphanumeric overlay and cursors
- Track-While-Scan (TWS) for up to twenty targets
- Extensive Built-in-Test (BIT)

In addition, the following optional features are available:

- Electronic Support Measures (ESM) interface
- Data link interface
- MIL-STD-1553B interface
- Pilot's repeater
- Identification Friend or Foe (IFF) interface

A block diagram showing the baseline configuration of the APS-504(V)5 and optional features is given in Figure 1. The APS-504(V)5 is packaged into eight Line Replaceable Units (LRUs) with approximate weights as shown in Table 1.

A variety of processing options are available to the operator to maximize the probability of detection of maritime targets. Two pulse compression modes are available. A 200 nanoseconds resolution compressed pulse width is provided for the detection of larger targets at ranges of up to 200 nautical miles and a fine resolution pulse with 30 nanoseconds compressed pulse width is provided for the detection of small targets such as cigarette boats and motor launches in clutter at shorter ranges. The Digital Filter Preprocessor (DFP) is used to preprocess the radar video when the fine resolution pulse width is in use. Pulse-to-pulse integration and Constant False Alarm Rate (CFAR) processing are selectable and scan-to-scan integration may be performed with a fully ground-stabilized display.

SAR TRANSCEIVER SUBSYSTEM (STS)

A fully coherent variant of the APS-504(V)5 RF subsystem has been developed for the MacDonald Dettwiler Integrated Radar Imaging System (IRIS) Synthetic Aperture Radar (SAR) (Akam et al, 1988). The STS functions entirely under the control of a data link and provides two fully coherent pulse compression modes using similar SAW technology to the APS-504(V)5 system. A pulsed TWT optimized for the SAR requirement is used in the coherent transmitter. The receiver output to the SAR processor is in the form of In-Phase and Quadrature components of the coherent radar video at baseband. The appropriate coherent timing signals are also provided by the STS. High-quality SAR imagery has been produced from raw video provided by the STS.

LIGHTWEIGHT RADAR (LWR)

The target environment in which maritime surveillance radars must operate is an increasingly demanding one. It is required that such systems be capable of reliably detecting high-speed surface targets of low Radar Cross Section (RCS). Future radars should also be capable of detecting aircraft which may cross land/sea boundaries. Such target detection requirements demand a high degree of clutter rejection which may be obtained by discriminating between targets and clutter on the basis of their Doppler frequencies. Pulse Doppler techniques (Morris, 1988) provide an effective means of meeting the target detection performance requirements but impose stringent requirements on almost all aspects of the design of the radar system, including:

- High degree of coherence
- High degree of linearity
- Wide dynamic range
- Low antenna sidelobes
- High-speed signal processing
- Post-detection target data processing requirements

The techniques employed in the fully coherent STS variant of the APS-504(V)5 RF subsystem which have already been demonstrated will form a cornerstone of the new Litton lightweight airborne maritime surveillance radar system.

The LWR Receiver Unit (RU) will provide both noncoherent and coherent radar video outputs. The noncoherent video (in the form of the logarithm of the magnitude of the envelope of the IF signal) will be processed using the proven noncoherent signal processing techniques employed in the APS-504(V)5. Newly available VLSI devices allow the word length and processing accuracy to be increased in some areas of the signal processing. The DFP Unit of the APS-504(V)5 will be integrated into the Signal Processing Unit (SPU) of the LWR thus eliminating one of the APS-504(V)5 LRUs and providing a saving in the weight of the system.

The coherent radar video (In-Phase and Quadrature components) will be processed by a new Pulse Doppler processor to provide improved detection of moving targets. As a result of newly available VLSI devices for digital signal processing, the Pulse Doppler processor may be implemented in a sufficiently compact form to be located within the LWR SPU.

Pulse Doppler processing can provide improved performance against moving targets for the following reasons:

- Coherent integration is more efficient than noncoherent integration
- The area of land or sea contributing to the clutter level in a resolution cell is reduced by formation of the Doppler spectrum
- Target returns and clutter returns are separated in the Doppler spectrum.

Pulse Doppler systems are usually classified according to Pulse Repetition Frequency (PRF) as follows:

- Low PRF systems are unambiguous in range and highly ambiguous in Doppler
- Medium PRF systems are ambiguous in range and Doppler
- High PRF systems are unambiguous in Doppler and highly ambiguous in range.

The selection of PRF is a complex issue which depends on a wide variety of factors (Williams and Radant, 1983). The achievable target detection performance is ultimately determined by the degree of spread of clutter in the Doppler spectrum. However, the spread of clutter in the Doppler spectrum depends on the flight envelope of the radar platform, antenna radiation pattern, operating frequency of the radar and other factors. The extent of the Doppler spectrum (and hence the extent of regions of the spectrum which are clear of clutter) is determined by the PRF. In airborne maritime surveillance radar applications, the appropriate selection of PRF tends to differ from application to application and even from mission to mission. Both low and medium PRF Pulse Doppler capability will be required to meet the requirements imposed by the full range of applications anticipated for the LWR. The Pulse Doppler Processor is therefore being designed with a modular and programmable architecture in order to provide the required flexibility.

The weight and volume of the transmitter and receiver are being reduced in comparison to the APS-504(V)5 units by adopting newly available technologies, while still maintaining sufficient linearity and dynamic range to provide a high-performance Pulse Doppler capability. The reductions are being effected by careful layout and the increased use of hybrid and stripline circuitry. Plug-in modules are being used extensively in order to achieve high packing density and improved maintainability.

A detailed trade-off study has shown that a further reduction in the weight of the system may be made by reducing the peak power of the radar from 8 kilowatts to 4 kilowatts. This results in a significant reduction in the required weight of the Travelling Wave Tube heat exchanger. The excellent long-range performance of the APS-504(V)5 system may be maintained by making use of improved Low Noise Amplifier (LNA) technology to further reduce the receiver noise figure. Furthermore, a smaller and lighter TWT power supply with increased capability may now be designed as a result of improvements in power semiconductor technology.

A block diagram showing the LWR system configuration is shown in Figure 2. The LWR will be packaged into 4 LRUs with weights as shown in Table 2. The weight of the LWR will be increased in applications which require standalone control and display capability.

CONCLUSIONS

The airborne, scanning maritime surveillance radar is a versatile, cost-effective and efficient means of providing real-time, wide area surveillance of coastal waters in a variety of applications. Newly available VLSI signal processing devices and lightweight radar technologies enable future maritime surveillance radars to provide increased performance and to enter new application areas involving the detection of aircraft and sea surface moving targets. The gains in performance and versatility can be achieved with a radar similar to the APS-504(V)5 with reduced weight. Capabilities which were previously available only in much larger and more expensive systems will be available in a lightweight maritime surveillance radar that may be installed in light aircraft or helicopters.

REFERENCES

1. Akam, W.R., Deane, R.A., Sartori, M., Lowry, R.T. and Mercer, J.B., 'An Integrated Radar Imaging System for the STAR-2 Aircraft', IEEE 1988 National Radar Conference Proceedings, University of Michigan, Ann Arbor, Michigan, April 1988.

2. Morris, G.V., 'Airborne Pulsed Doppler Radar', Artech House, Inc., Norwood, MA02062, ISBN 0-89006-272-2, 1988.

3. Salvy, R., 'Airborne Surveillance of the EEZ', International Defense Review, Vol. 12, pp. 1765-1772, 1986.

4. Williams, F.C. and Radant, M.E. 'Airborne Radar and the Three PRFs', Microwave Journal, July 1983. '(Reprinted in: Skolnik, M.I., (Ed.), 'Radar Applications', IEEE Press, ISBN 0-87942-223-8, pp. 272-276, 1988.

APS-504(V)5	Weight (lbs)
Power Amplifier Unit (PAU)	99
Exciter/Receiver Unit (ERU)	54
Antenna/Pedestal Unit	39
Antenna Control Unit (ACU)	18
Digital Filter Preprocessor (DFP)	16
Operator's Display Unit (ODU)	31
Radar Control Unit (RCU)	22
Synchronizer/Processor and Converter (SPC)	42
Equipment trays and miscellaneous	20
Total Weight (approx)	341

TABLE 1 APS-504(V)5 Line Replaceable Units (LRUs) and Approximate Weights.

Lightweight Radar	Weight (lbs)
Scanner Unit (including Antenna/Pedestal)	44
Transmitter Unit (TU)	48
Receiver Unit (RU)	38
Signal Processing Unit (SPU)	48
Equipment trays and miscellaneous	7
Total Weight (estimated)	185

TABLE 2 Lightweight Radar (LWR) Line Replaceable Units (LRUs) and Estimated Weights.

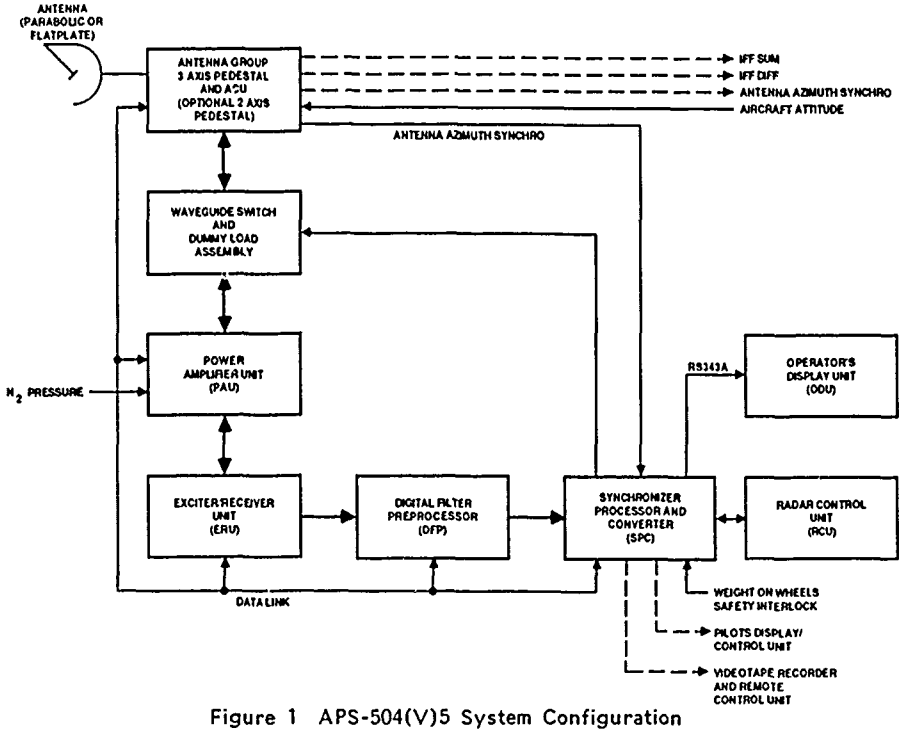


Figure 1 APS-504(V)5 System Configuration

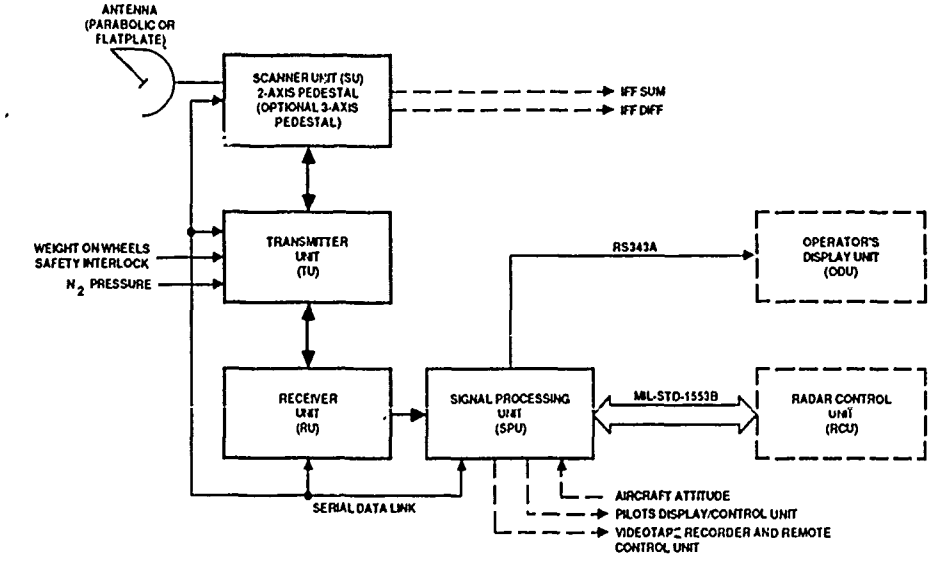


Figure 2 Lightweight Radar System Configuration

ICE MANAGEMENT FOR THE OFFSHORE FISHERY

PHILIP RUDKIN HERBERT RIPELY KEN LUDLOW

ATLANTIC AIRWAYS LIMITED CANADA

Abstract

Modern technology, a shrinking resource and economic pressure have forced modern fishing vessels to routinely enter and work in ice covered waters. Working in these conditions is not only time consuming and expensive, but can occasionally be hazardous.

Site specific ice management was initially developed for the offshore oil and gas industry, and over ten years has evolved into an effective and reliable method of handling operations in ice infested waters, resulting in the minimum amount of down time.

The need for effective offshore ice management has been highlighted by the recent quota system implemented by the Canadian Fisheries Department.

Through the joint efforts of Fishery Products International and Atlantic Airways, a limited ice management system was tested during the 1988 winter fishery.

This paper will detail the need and economic benefits involved in an integrated ice management system specifically designed of Canada's North East Coast offshore fishery.

Clear demand, Gain information, Data management, Operational costs.

Ice Conditions with Respect to the East Coast Fishery

The fishery areas of the North East coast of Canada extend from latitude 44 degrees North to 70 degrees North. Within these boundaries two distinct ice areas exist: to the North, Artic pack, and to the South the highly mobile Marginal ice zone. The Northern sector within the bounds of 60 - 70 degrees North is totally impenetrable from December to May, being totally covered with heavy Artic packice.

The main fishing effort in this area is held between July and September depending upon ice conditions.

The Southerly section is divided into two sub areas, the Labrador margin and the Grand Banks.

Due to recent DFO quota regulations (Optimum sustainable yield) the fishing effort on the Labrador margin is now concentrated in February and March, with the shrimp fishery extending into June. While the fishing effort on the Grand Banks is year round.

Data Requirements

The Captain of a trawler operating in the seas off Newfoundland and Labrador can gain information on ice conditions from (3) three sources in normal operating conditions.

1. The vessels radar will provide limited information on ice coverage out to a maximum distance of 10 miles.
2. Area ice charts, compiled by AES are transmitted daily via weather fax. They contain information of a general nature and often show only the previous day's ice situation.
3. The Canadian Coast Guard's ice control centre makes daily broadcasts of ice conditions based on the best information they have available.

These three items on their own do not provide enough information to conduct effective large scale vessel operations in ice covered waters. It has been stated that to efficiently manage vessel operations in a winter fishery. There is a frequent and clear demand for detailed site specific, real time data (Hogan 87) on ice conditions in three areas.

1. Immediately upon departure or return to/from a Newfoundland port.
2. During the passage to and from the Northern fishing locations.
3. During the fishing operation.

Equipment

In the 1988 fishery a specially equipped

remote aircraft owned and operated by Atlantic Airways Limited was used to assist Fishery Products International in its demand for real time ice data. This aircraft is specially designed for offshore ice surveillance and features a completely self contained ice management system designed to relay real time ice data to surface vessels.

Onboard data collection is accomplished using a Litton APS 504(V) 5,360 degree surveillance radar. Radar range is user selectable between a very high resolution 3 mile scale, to a 200 mile look, covering an area of 94,000 square miles with every sweep. Radar resolution below the 50 mile range scale is capable of reliably detecting ice in as little as one tenth concentration and as thin as new ice.

The (V) 5 is a digital radar making it ideal for ice surveillance. The operators display shows target returns in various gray levels depending upon the strenght of the target (thickness of ice). This makes it possible to classify and type ice based purely on radar imagery.

Airborne data management is accomplished by the use of a computer permanently mounted in the aircraft. Custom computer software, designed and written by Atlantic Airways Ltd. allows the computer to receive information directly from the radar, and the aircraft's navigation system. Additional programs then turn this data into easily understood ice data listings. Ice charts are also produced (to AES standards) using this digital data in conjunction with ice coverage and thickness information entered into the computer by an experienced ice observer. Supplementary low level visual information is included for areas of specific interest.

The aircraft is equipped with a complete array of communications equipment. One of the most important parts of which is the HF facsimile transmitter. This unit is frequency selectable to cover the entire HF range, making it possible to simultaniously pass ice charts to all vessels equipped with a weather fax receiver.

Atlantic Airways has been operating this ice surveillance system for the offshore oil and gas industry for the past four years. At the beginning of 1988 , ice surveillance operations were begun for Fishery Products International.

Coverage

Coverage concentrated on all three areas previously identified. Vessel routing flights where flown resulting in one case of transit time savings of over forty eight hours.

Several flights were conducted to assist trawlers extend the fishable time in a specific area, one case in area 2J off the Labrador coast resulted in the extention by (11) eleven vessel days and 749,00 lbs of Northern Cod.

The results of the above examples and the work for the Oil industry over the past years have proven that this system is capable of providing detailed and highly accurate ice information and is an extremely cost effective way to gather ice information over a large area.

Conclusions

To receive the maximum benefits from an ice management system the utilization period would be in the order of three to four months and should included full ground based support.

While our test period covered two and one half months and less than 50 hours flight time, the results were very successful. Acceptance by the vessel masters was very good and all vessel crews where keen to receive the ice charts from the aircraft. All feed back received from the fishing vessels was positive.

We would like to take this opportunity to thank all people involved in this test and especially the vessel crews for their patience and help when setting up the first few data down links.

Cost factors associated with operating this ice management system are fairly high, even so it has been stated in a benefit assessment conducted by the client (Hogan 1988) that each deep sea trawler's contribution to profit per hour of harvesting is $1,000.00 Canadian and through out the test period substantial savings on opertional costs were realized.

The offshore oil and gas industry operate a joint ice management system. Data collection and management is conducted from one central location and the data products are then distrubuted to all participants (in their own perferred format). This would seem an ideal solution for the offshore fishing industry, and by offsetting the cost against vessel operating budget, the figure comes in at about $50.00 per vessel day for a full ice management system.

As demonstrated the potential cost savings far exceed this figure, not to mention the increase in safety of operations in ice conditions.

Bibliography

Rudkin,P. et al "Surveillance radar a new tool for ice surveillance". Proceedings or The 21st International symposium of remote sensing, 1987.

Rudkin, P. et al, "Integrated aerial surveillance system designed for ice management in the marginal ice zone". Proceedings of POLAR TEC 88. Trondheim, Norway 1988

Hogan M. et al "The use of fixed wing aircraft in ice reconnaissance for a deep sea fishing fleet" Proceedings of The World symposium on fishing gear and vessel design, 1988.

DIGITAL PROCESSING OF X-BAND VARAN-S AIRBORNE SAR IMAGES

R.Albrizio(*), P.Blonda(*), A.Mazzone(*), F.Pasquali(^)
G.Pasquariello(*), F.Posa(@), N.Veneziani(*)

(*) Istituto Elaborazione Segnali ed Immagini (Bari,Italy)
(@) Dipartimento di Fisica, Universita` di Bari (Italy)
(^) Centro di Geodesia Spaziale (Matera,Italy)

ABSTRACT

This paper describes the techniques developed to process SAR data obtained by the CNES X-Band VARAN-S Airborne SAR during the Agrisar-86 campaign on the Italian test-site. The processing has been performed starting from the sensor "raw-data", whose main characteristics are the analogical range compression, the azimuthal prefiltering and decimation, all executed in real-time onboard.
The SAR data have been firstly compressed in azimuth by means of a matched filter in the frequency domain. The filter parameters were derived from standard clutterlock and autofocus algorithms.
With regard to the autofocus, two partial Doppler bands of some azimuth lines at different ranges were used for a correlation process to find the true sequence of the Doppler rate values.
Point and extended targets have been selected on the resulting full resolution images in order to assess both the data and processing quality.
In order to fully exploit the use of these SAR images, their radiometric calibration has been performed utilizing the echo signal from trihedral corner reflectors
The availability of ground truth enabled to assess the validity of the textural characterization of crop types, where the textural analysis is based on the co-occurrence matrix method.

1. INTRODUCTION

The growing interest for microwave remote sensing, the future programs of Earth observation from space using SAR sensors for their all-weather capabilities, the opportunity to participate to Agrisar-86 campaign, offered by the JRC-ISPRA Establishement of EEC (Economic European Community), led this group to focus on SAR data processing and gain experience from the analysis of data produced by the VARAN-S Airborne SAR.
The effort has been devoted to implement and test a processing chain able to obtain calibrated SAR images, useful for classification purpose and for integration with images in other spectral bands.
The approach used is flexible enough to handle data from various SAR sensors, either airborne and spaceborne.

2. SENSOR CHARACTERISTICS AND DATA COMPRESSION

The french X-band airborne Varan-S SAR was used during the AGRISAR-86 campaign, the sensor was mounted on WWII B17. The antenna pointing was stabilized by means of servo system connected to the LITTON LTN 96-100 gyrolaser INS, the latter provided also the flight parameters to be recorded on tape togheter with the antenna corrections.
However this part of ancillary data were not on the CCT raw-data tapes available to us.
In tab.1 the main sensor characteristics and flight parameters, [1], are indicated.

INS : LITTON LTN 96-100
Radar : THOMSON-CSF VARAN-S

Radar carrier	=	9.37 GHz
PRF	=	400 Hz
Chirp bandwidth	=	67 MHz
Trasmitted pulse length	=	10 μs
Polarization	=	HH/VV
Antenna aperture (azimuth)	=	2.3 deg
Antenna aperture (range)	=	41.7 deg
Platform speed	=	103 m/s
Squint angle	=	90 deg
Look angle	=	55 deg
Flying heigth	=	5790 m

Tab.1 Flight parameters and system technical characteristics

For each PRF pulse, the echo signal from the illuminated ground area was range-compressed in real-time at the receiver by means of the analogic SAW circuit. After this compression, the resulting signal was sampled at 50 MHz to obtain the complex sample sequence with 5 bit/sample.

2.1 Raw-data characteristics

The CNES-Toulouse raw-data CCT's contained records of 14648 bytes; each record is subdived in 7 sub-records of 2092 bytes and other 4 header bytes. Each sub-record has a

92 bytes header and 2000 bytes of data.
Fig.1 gives an example of the raw-data before di azimuth compression relative to the flight of June 26th (OP4211HH).
In the azimuth direction the useful Doppler band and PRF were reduced by a factor of 8 using a filtering (presummer) that degraded the theoretical azimuth resolution of 0.5 mt to about 4 mt.

Fig.1 AGRISAR Raw-data, range compression on board (corner reflectors window)

2.2 Azimuth compression

The azimuthal presumm cancelled the information relative to the frequency Doppler centroid (fdc) as verified by means of the clutterlock algorithm. As consequence only the Doppler rate (fdr) could be extracted from the raw-data in order to obtain the Doppler information [2,3].
The knowledge of the fdr has to be as precise as possible to focus accurately an airborne SAR image. Unfortunately the fdr itself is strongly dependent on the flight geometry and varies from near to far range [4], requiring the fdr determination and updating in function of the corresponding range gate. In our processing the precise fdr behaviour has been obtained by means of a very accurate geometrical model and an autofocus procedure, [1,4], as shown in fig.2.
The results show an fdr linearly varyng between -108 Hz/s and -90 Hz/s on the selected azimuthal lines.
Fig.3 reproduces the same area in fig.1 after the focusing. Part of our focused image, containing a cross-like array of corner reflectors (C.R.s) is shown in fig.4 in comparison with the same area processed by the CNES.

3. QUALITY CHECK OF PROCESSED IMAGE

A first level quality check of a SAR processed image can be performed, in general, by means of radiometric resolution measures on omogenous extended targets and geometric resolution measures on artificial point-like targets such as C.R.s [4,5].
In our case, the radiometric characterization of the single look processed image, fig.3, has been carried out by measuring the

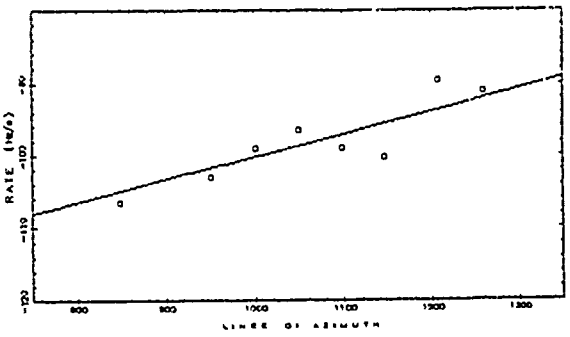


Fig.2 Fdr behaviour as obtained by means of the autofocus procedure and linear fitting result of the experimental values (indicated with "□")

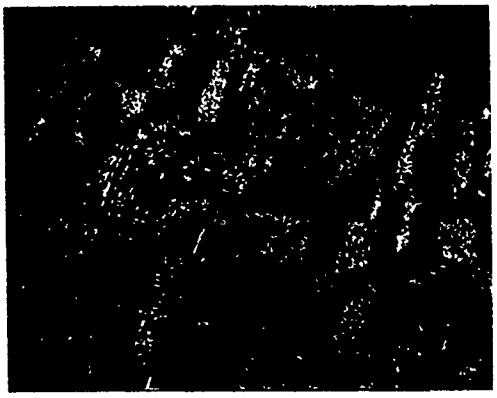


Fig.3 Result of azimuth compression

Fig.4 Top: cross-array image processed by our processor; bottom: cross-array image processed by CNES

radiometric resolution γ, defined as:

$$\gamma = 10 \text{ Log } (1 + \sigma/\mu) \qquad (1)$$

where μ is the average value of the levels distribution (intensities) on uniform areas and σ is the related standard deviation.
The obtained value is $\gamma = 2.976 \pm 0.092$ dB. This result is in excellent agreement with the expected value of 3.010 dB, [6].

The geometric resolution measures have been obtained form the response of C.R.s array aligned along the azimuth and range direction.
Furthermore, the Peak Side Lobe Ratio (PSLR) has been measured for a better check of focusing process. The results are reported in tab.2, in tab.3 the corresponding results on the CNES processed image are also reported. In such tables Rsr and Ra are the slant-range and azimuth resolution respectively. By compairing the two set of measures, it can be observed that the value of geometric resolution are in good agreement with our image, showing a smaller variance with respect to the CNES image.
However, it must be pointed out that the obtained Rsr and Ra values are not as good as one could expect from the theory that gives Rsr = 3 m and Ra = 4 m as limit values. A possible explanation is the influence of: the analogic compression and sampling procedure utilized for the range processing, the true aircraft speed varations and the azimuthal prefiltering (presumm). In fact, it should be considered, for the range direction, that the trasmitted signal bandwidth was 67 MHz and the sampling occurred at 50 MHz [1,7]; therefore as a consequence the signal bandwidth after the SAW was reduced and the measured Rsr = 4.5 m indicates that effective bandwidth was 33 MHz.
Concerning the processing in azimuth, it has been verified experimentally that the presumming filter bandwidth is not larger then 25 Hz; with such value the maximum achievable azimuth resolution is 4 m in agreement with the measured values.

	Rsr [m]	PSLR [dB]	Ra [m]	PSLR [dB]
CR1	4.528	9.117	4.305	10.272
CR2	4.345	13.268	4.406	7.295
CR3	4.691	11.041	4.339	10.950
CR4	4.162	14.483	4.345	12.225
CR5	4.077	8.908	4.641	13.287
CR6	4.518	10.443	4.401	8.330
CR7	5.108	0.941	4.670	7.972
CR8	3.657	8.937	4.387	13.054
CR9	4.984	6.279	4.138	3.886
CR10	4.919	2.531	4.288	10.032
CR11	4.535	7.371	4.425	8.346
< >	4.50	8.5	4.395	9.6
err.	0.13	1.2	0.046	0.8

Tab.2 Measurements result on our image

	Rsr [m]	PSLR [dB]	Ra [m]	PSLR [dB]
CR1	4.868	11.912	6.371	12.342
CR2	5.916	10.376	4.947	13.887
CR3	5.756	13.312	4.796	8.695
CR4	5.407	16.151	2.914	6.999
CR5	4.138	5.654	7.517	4.982
CR6	4.769	9.960	2.756	5.096
CR7	5.085	7.058	2.639	5.879
CR8	5.700	6.706	2.956	4.913
CR9	5.346	3.694	5.893	12.265
CR10	-	-	-	-
CR11	7.472	7.114	3.001	4.340
< >	5.45	9.2	4.4	7.9
err.	0.28	1.2	0.6	1.1

Tab.3 Measurements result on CNES-image

4. SPECKLE REDUCTION

The adopted approach to speckle reduction is based on the "multilook" or "subaperturing" technique. Same indipendent images (looks) of the same scene have been produced by using a limited, without overlap, part of the synthetic aperture; this implies that looks have a limited doppler bandwidth and hence a degraded spatial resolution [6]. The look extraction, [5], operates in the frequency domain and uses data already focussed in azimuth. In particular, N looks are obtained by subdividing the azimuthal spectrum in N segments with a band-pass filtering operation, the applied filter is N times ad hoc shifted.
Processing wuth full geometric resolution and N = 2,3,4 has been done, for each set of data the radiometric resolution has mesured, according to (1) and compared with the expected value, (tab.4). The presence in the scenes of C.R.s enabled us to verfy also the degradation of the azimuth resolution with the increasing N of looks (tab.4).
The experimental results are in good agreement with the predicted values, so it can be concluded that our processor does not introduce any significant bias on the raw-data.
A further test of the processor is illustrated in fig.5, where for a four looks the standard deviation values are plotted in function of the average reflectivity value for extended targets, resulting fit is linear and its slope value is in agreement with the expected value.

	Theoric [dB]	Experimental [dB]	R.az. [m]
N = 1	3.010	2.997 $\pm$ 0.083	4.40
N = 2	2.323	2.283 $\pm$ 0.034	8.83
N = 3	1.979	1.978 $\pm$ 0.030	13.29
N = 4	1.761	1.774 $\pm$ 0.024	17.56

Tab.4 Behaviour of the radiometric and geometrical resolution versus N (number of look) as evaluated on our image

5. RADIOMETRIC CALIBRATION

A radiometric calibration has been attempted using: the echo signal from C.R.s deployed on the test-site on the test-site; the backscattering coefficient of known extended target; the antenna elevation gain model, [8].
The aim has been the extraction of σ° for extended homogenous targets from the

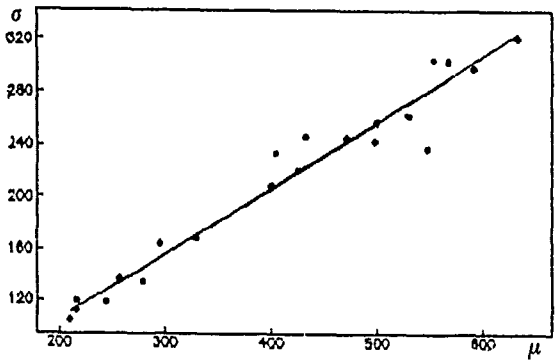


Fig.5 Result of linear fitting (4-looks AGRISAR image)

corresponding SAR power image. Because the SAR imagery is a scaled and sampled representation of the received power, the first step has been the determination of the relationship between the grey level ΔN on the image and the received power. Considering $y = \sigma^0/\cos\theta$ as the radar cross section (R.C.S.) per projected resolved area, [9], and $\langle\Delta N\rangle$ the average grey level the equation used for calibration has been:

$$y = \langle \Delta N \rangle / [\ \bar{K}\ G_0^2 H\ F'(R)\ ] \qquad (2)$$

where: $\bar{K}$ is considered constant over the extended target and contains scaling factor, SAR parameters and geometric resolution; G_0 is the antenna gain at center beam; F'(R) the normalizing function determined by modeling the antenna pattern and varies with the slant-range value; H is the aircraft altitude. The constant $\bar{K}$ has been computed from the point target (C.R.) ΔN. In this respect it has been preferred to integrate the response of C.R.s, assuming their R dependence, ruther than using the peak value [10], because the total C.R. reflected energy results to be indipendent of focus variations in the SAR processor, random phase varations in the reflector and multilooking. By applying this calibration procedure to the CNES slant-range images (fig.6) it appear an overcompensation in the middle-near range areas probably due to processing and pour modelling of the antenna pattern.

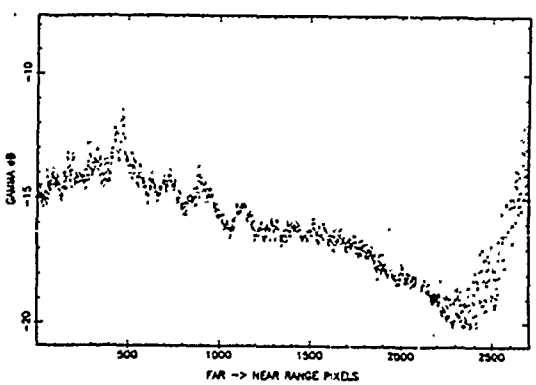


Fig.6 Gamma profile in range (PAV211)

6. TEXTURE ANALYSIS

The investigation of new methods of data processing, such as texture analysis, finalized to land cover, land use discrimination for thematic maps production, is one of the primary objectives of the European Campaign "Agrisar 86", [7]. In order to fulfill this goal, the SAR images corresponding to the HH and VV polarization of the June 26th and July 16th acquisition flights over the Italian testsite, have been analysed. For each image some fields corresponding to three selected crop types (ready for harvest wheat, sugarbeet and grass) have been identified and, for each field, statistical characteristics, both in terms of first-order statistics (mean and standard deviation of intensity values), and in terms of second order textural features, have been collected.

6.1 Testsite description

The Italian testsite covers an area (known as "Oltrepo` Pavese") of about 100 sq. Km, located on a region of primary agricultural use, within the Upper Po Valley. From a morphological point of view, mostly of the overall area lies in a flat region, chracterized by intensive land use practices; in this subpart are located the fields under investigation.
From a thematic point of view, a large variety crops can be found (near 30 different crops, including wheat, corn, sugarbeet, vegetables, grass, vineyard and so on).
The average field size is small, tipically less than 1 ha. The last two features are well representative of the overall Italian land use situation.
The attention of our investigation has been focused on a rectangular area of 1*2 Km: this sub area was selectd for the large mean crops extension, in order to facilitate the comparison between the ground truth, represented by a crop map at 1:10,000 scale and the SAR images.

6.2 Texture characterization

Following the ground truth map, 18 training fields have been selected on the SAR images. More specifically has been possible to identify 5,6 and 7 field for wheat, grass and sugarbeet crops, respectively.
For each field, in tab.5, are reported the number of pixels, the mean intensity , expressed in digital counts, and the associated standard deviation, relatively to the HH polarization for the June 26th image. The intensity values show a large degree of overlapping, i.e. the variance of a crop signature is comparable with the distance between two different crops. Moreover, in the last column of the tab.5, the normalized variance values (ratio between variance and the square mean) are reported. These values essentially match the value of 0.273 expected [2] to explain the variance duo to the speckle for an 1-look intensity radar image. For this reason, a 4-look image has been generated: the corresponding statistical features are shown in tab.6, for each field.

FIELD	PIX #	MEAN	ST.DEV.	$(\sigma/\mu)^2$
W1	703	55.56	31.40	.320
W2	607	61.82	36.13	.341
W3	645	70.54	41.35	.344
W4	816	76.12	41.37	.295
W5	904	74.31	39.57	.283
G6	448	78.48	43.14	.302
G7	586	89.10	47.42	.283
G8	324	76.69	44.12	.330
G9	394	74.82	42.25	.319
G10	437	118.78	62.01	.272
G11	349	97.08	51.99	.287
S12	1108	104.47	58.01	.308
S13	966	112.26	58.00	.267
S14	743	121.83	62.98	.267
S15	1787	118.60	63.29	.285
S16	1142	124.05	64.02	.266
S17	1571	127.62	66.12	.268
S18	647	124.00	61.13	.243

Tab.5 Statistical features for June 26th HH 1-Look image

FIELD	PIX #	MEAN	ST.DEV.	$(\sigma/\mu)^2$
W1	703	57.32	28.81	.253
W2	607	61.09	33.50	.301
W3	645	76.11	37.84	.247
W4	816	79.25	39.29	.246
W5	904	80.57	38.94	.234
G6	448	81.65	40.10	.241
G7	586	99.44	43.77	.194
G8	324	82.59	38.52	.217
G9	394	83.76	40.54	.234
G10	437	149.64	52.01	.121
G11	349	106.20	43.21	.166
S12	1108	117.38	56.33	.230
S13	966	128.50	55.37	.186
S14	743	145.82	54.72	.141
S15	1787	139.00	57.35	.170
S16	1142	143.97	57.67	.160
S17	1571	147.57	56.68	.147
S18	647	143.44	56.12	.153

Tab.6 Statistical features for June 26th HH 4-Look image

In the case of a 4-look image, the expectd speckle dependent normalized variance is 0.273/4=0.068: the values shown in the last column of tab.6 are clearly larger, revealing that the intrinsic spatial variability, i.e. the "texture" dependent information, predominates when the speckle is reduced by a multilooking processing. In order to give a quantitative measurement of the texture associated to each crop type, the second order Gray Level Co-occurrence Matrix (GLCM) method has been adopted.
For each field, the co-occurrence matrices have been computed for two main directions (range and azimuth) and for ten inter-pixels distances d (ranging from 1 to ten). From each co-occurrence matrix two parameters have been extracted: the Contrast (CON) and the Angular Second Moment (ASM). In fig.7 and 8, the values of Contrast in range and azimuth direction are reported versus d, for the three analyzed crops.
In fig.8, the effect of redundancy due to the

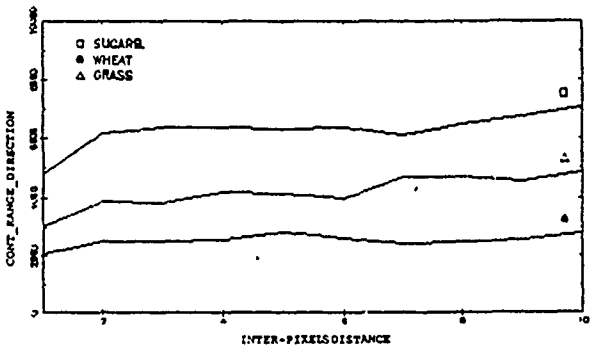


Fig.7 Contrast values in the range direction versus inter-pixel distance for June 26th HH 4-Look image

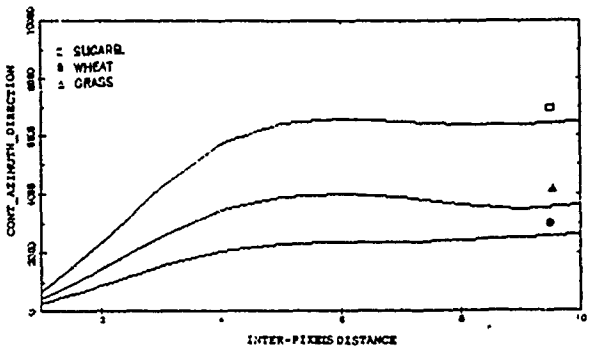


Fig.8 Contrast values in the azimuth direction versus inter-pixel distance for June 26th HH 4-Look image

4-looks processing, is well emphasized, corresponding to low values for contrast in azimuth direction and for an inter-pixels distance less than 4. In fig.9 the distribution of the 18 fields is shown, with respect to the values of CON and ASM copmuted in azimuth direction, for a inter-pixels distance equal to 5. In order to give a numerical estimate of the separability between different crop types, achievable using textural segnatures, the distance D has been introduced:

$$D(i,j)=|M(i,k)-M(j,k)|/(S(i,k)+S(j,k)) \quad (3)$$

where M(i,k) and S(i,k) refer to the mean and standard deviation, computed on a per-field base, of the feature kth for the ith class.
The values of D for the wheat, sugarbeet and grass cultures are reported in tab.7, relatively to the first order mean intensity values and the second order CON and ASM

features.
The co-occurrence matrix parameters are computed for two directions and at a distance of 5 pixels. Two main conclusions can be extracted, looking at this table: preprocessing, such as multilooking tecniques, can greatly influence the features extraction step; the statistical separability achievable between different crops is slightly better when computed starting from second order segnatures; in other words, the texture dependent characteristics, such as those extracted from co-occurrence matrices, can usefully improve the process of land use classification using SAR image.

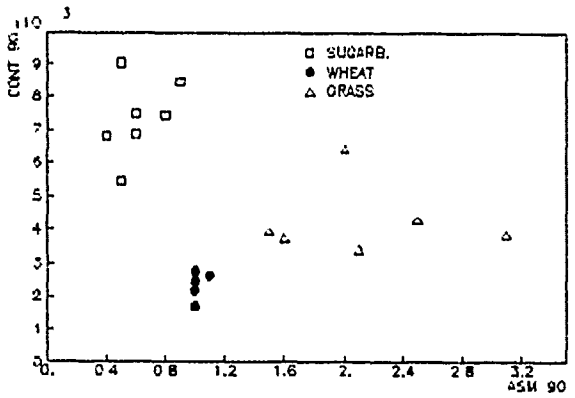


Fig.9 "Textural" Characterization of crop fields for June 26th HH 4-Look image

CL1-CL2	MEAN	CON0_D	CON90_D	ASM0_D	ASM90_D
WHE-GRA	.89	.73	1.11	.67	.86
WHE-SUG	3.36	2.76	4.90	1.21	2.40
GRA-SUG	1.10	1.62	2.11	2.08	2.16

Tab.7 Inter-crops distances for June 26th HH 4-Look image
D=|M2-M1|/(STD1+STD2)

7. REFERENCES

[1] Vaillant,D., "Varan-S: An airborne syntetic aperture radar for research in microwave remote sensing" Proc. Earsel Symposium, Strasbourg 1985.

[2] Ulaby,F.T., Fung,R.K., "Microwave remote sensing", vol. I, II and III, Addison Wesley, 1981-82.

[3] Elachi C., "Spaceborne Radar Remote Sensing: Applications and Techniques", IEEE Press 1987.

[4] Albrizio,R., Mazzone,A., Veneziani,N., "Elaborazione di dati grezzi SAR airborne (AGRISAR-86)" (Italian), II Convegno A.I.T., Bolzano, November 1988.

[5] Albrizio,R., Ferrara,L., Mazzone,A., Pasquali,F., "Analisi e sperimentazione di tecniche multilooks per l'elaborazione di dati grezzi SAR" (Italian), II Convegno A.I.T., Bolzano, November 1988.

[6] Li,F.K., Croft,G., Held,D.N., "Comparision of several techniques to obtain multiple-look SAR imagery", IEEE Trans. on Geoscience and Remote Sensing, Vol. GE-21, No. 3, July 1983.

[7] Sieber A.J., "AGRISAR-86 Preliminary Investigators' Report", S.P./I.87.25, J.R.C. ISPRA Establishment, July 1987.

[8] Conenna,P., Zicarelli,M., "Airborne Varan-S X-Band SAR data radiometric analysis",Proc. of EARSel, Capri, June 1988.

[9] Lopes,A., Touzi,R., "Extraction of backscattering coefficient of agricultural fields from an airborne SAR image", Proc. of IGARSS'86.

[10] Raney,R.K., "SAR processing of partially coherent phenomena", Int. Jou. of Remote Sensing, vol.1, n.1, 1980.

Helicopter Radar Simulations from Fine Resolution Airborne SAR Imagery

by

R.T.Lowry ‡, G.Wessels*, and G.McAvoy‡

INTERA Technologies Ltd.,

‡ 2500 101 6th Ave S.W.
Calgary T2P 3P4

* 600 1525 Carling Ave
Ottawa K1Z 8R9

ABSTRACT

A simulation of different radar imagery of sea ice representing helicopter altitutes has been conducted, based on data collected from a high altitude fine resolution SAR. The purpose of the study was to provide a first level of imagery to be used in defining the best tactical sensor to be placed aboard Coast Guard helicopters. The first part of the study evolved an empirical relationships between ice reflectivity and incidence angle based on data from the SAR at 9100 and 1500 m altitude. The compromise between real aperture radars (SLAR) and SARs, when flown at low altitudes was then studied. It was found that an X-band SLAR with a 3 metre antenna produced imagery with a useful swath width of 15 km when an altitude of 1000 m was assumed. A SAR at the same altitude had an effective swath width that was approximately the same. Similarly a PPI type radar with an antenna length of 1metre was found to produce an optimum image, if a near angle cutoff of 35° was assumed. An altitude hole becomes bothersome at higher altitudes, limiting the image utility and the far range data have poor resolution with the PPI style, short antenna.

1.0 Introduction

Ice reconnaissance, using fine resolution airborne SAR systems has become the accepted standard for both operational and research needs. The all weather performance of radar, combined with the ability of SAR to match the resolution, swath width and data rate to the needs and means of the user, has made SAR the ideal ice reconnaissance sensor. SAR systems, such as the dual sided CIRS (Mercer et al, these proceedings) provide operational, wide area coverage, from high altitude aircraft. These data are useful for both strategic and tactical navigation purposes, but have a rather short useful lifetime for close tactical support in moving ice conditions. It is therefore desirable to provide ships, navigating in ice infested waters, with real time information, similar to that available from SAR for close tactical support.

In the past, close tactical information has been available to a ship's crew from the ship's radar, and, if available, from the ship's helicopter pilot or ice observer The ship's radar, at heights above ice typically less than 30 metres, is not able to image the ice beyond ranges of a few km. Further, the range and angular resolutions are not typically adequate for a high quality image of the ice Visual reporting from a helicopter, while very useful, is limited to conditions of good visibility, and provides no permanent record that can be used for interactive decision making on board the ship The solution that would seem to present itself to those operators that carry helicopter is to equip them with imaging radars that return the images to the ship, either as a record, or on a telemetry link.

Helicopters are not an ideal platform for an imaging radar, as they fly too low, and often fly with highly variable ground speeds and directions. SAR systems that require very controlled, linear flight lines, require extensive motion compensation to fly on helicopters. SARs are also heavy and expensive, even in the versions intended for fixed wing aircraft. The more serious problem is the rather low altitudes at which helicopters typically fly. This means, that for any significant swath width, the angle of incidence of the radar energy at the ice surface can be very shallow.

Provided the illumination angle of the radar is kept in the mid-angles, the reflectivity of sea ice is very nearly constant, and imagery is easily interpreted. At very steep or very shallow illumination angles, the reflectivity is a strong function of both the angle, and the type of ice. At shallow angles, ice that is not deformed, and therefore being viewed at what is effectively a steeper angle, becomes nearly a specular reflector, and the illuminating energy is reflected away, and not returned to the radar. As a result, undeformed ice becomes invisible to radar at very shallow angles. The angle at which ice starts to become a very poor reflector is a function of both the salinity and the details of the surface morphology of the ice. The significance of this for helicopter reconnaissance is that unless the craft is flown at very high altitudes, the energy returned from the far edge of the swath will come exclusively from deformed ice. Thus either the helicopter must be flown at very high altitudes or the swath widths kept to very narrow ranges to generate useful ice information.

Since the purpose of the helicopter radar is to provide close range, tactical information, keeping the swath width restricted is not a severe limit. If one accepts that the radar will be restricted to low altitude by its operating capabilities, and the swath width will be kept restricted by the physics of the radar-ice interaction, then the possibility of substituting a SLAR or a Search type radar for the SAR presents itself. The prime advantage of the SAR is the ability to maintain constant resolution over wide swaths. If there is a restricted swath due to the physics, this capability could well be of little advantage. The degradation at far ranges seen in real aperture radars may not be serious for narrow swath radars used for helicopter reconnaissance in a close tactical support role.

To investigate the compromises involved in choosing a helicopter radar, INTERA undertook to provide the Canadian Coast Guard with a series of simulated images that would have the 'look and feel' of the imagery that would be expected from a helicopter borne radar, flown at low altitudes (1000 to 10000 feet ASL). The basis for the simulation was SAR imagery that was filtered and processed to represent data collected from radars flown at lower altitude. Three types of radars were simulated, a low altitude SAR, a SLAR, and a Search or PPI type radar. The SAR simulation consisted of modifications to the data to simulate the lower altitude, and therefore the more shallow illumination that would be obtained from a helicopter. The SLAR simulation was

applied over the SAR simulation and consisted of a variable low pass spatial filter used to simulate the degraded resolution associated with a real aperture system. The PPI simulation combined the radiometric and geometric simulations of the SAR and SLAR simulations, but with the antenna assumed to be rotating above the image.

2.0 Data Set Used

In November 1988, the INTERA STAR-1 system was mapping ice in Lancaster Sound. Data were selected for further study,that contained a wide mix of ice types, including newly forming ice, thicker first year ice, and old ice that appeared to consist of both second year and multi year ice. The data were collected at 9100 m (30,000 feet) altitude, with the radar in the wide swath mode. In this mode the radar has an effective swath width of 46 km, with a range pixel size of approximately 12 metres. The incidence angle varied from 33° at the near edge of the swath to 8° at the far edge of the swath. The azimuth (along track) spacing of the data lines was reduced in processing from 4 to 12 metres to both give 'near square' pixels, and to reduce the effects of coherent fading or 'speckle'. A scene of 4096 lines by 4096 pixels was selected from the data to be studied and is shown in Figure 1.

Two days later, a second pass was made over the area of interest, using the same flight line, at an altitude of 1524 metres (5000 feet) with the radar operated in the fine resolution mode. This provided a calibration altitude image set to develop the empirical transfer function. In this mode, the range pixel size is 6 metres, giving an effective swath width of 23 km. The incidence angle varied from 7° at the near edge of the swath to 2.4° at the far edge of the swath. The data were studied with the line spacing of 4.2 metres.

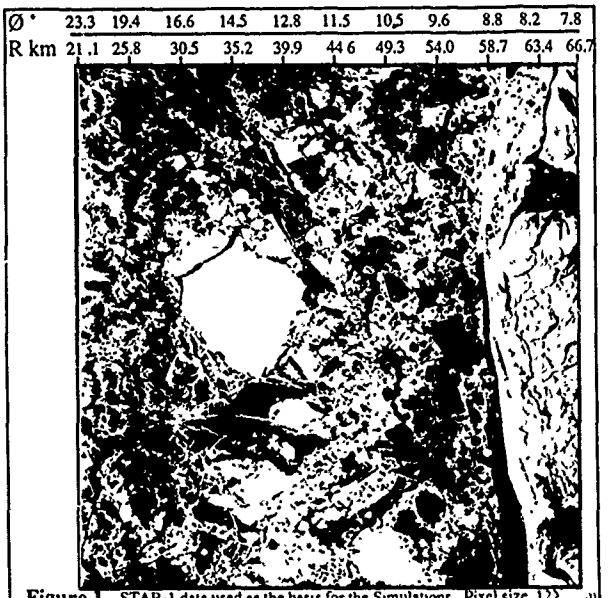


Figure 1 STAR-1 data used as the basis for the Simulations. Pixel size, 12? m
4096X4096 pixels shown, Actual Altitude = 9100 m

3.0 SAR Simulations

To simulate an airborne SAR flying at a lower altitude, it is necessary to have a model of the reflectivity of sea ice as a function of both ice type and angle. The radar reflectivity of sea ice is a strong function of the angle of illumination. This is particularly true at very shallow angles which are typical of marine radars (incidence angles between 0 and 3°). While considerable data are available on the reflectivity of sea ice at angles typical of satellite or airborne radars (Ulaby and Dobson, 1989), there have been very few experiments that have provided quantitative reflectivity data at these shallow angles.

To apply a reflectivity model to the different ice types require a means of classifying the data into ice types that exhibit different angular reflectivitiy charecteristcs. There have been several attempts to classify ice based on a single frequency and polarization, namely X-band, HH. These attempts have been largely unsuccessful, however because the ice types have significant overlap in their reflectivity. The data set shown in Figure 1 was examined to assess the separability of ice by category, based solely on X-HH reflectivity. It was noted that there was significant overlap in reflectivity between Old Ice and Second Year Ice as well as with younger ice forms such as ridges and rubble and even with new leads covered with frost flowers. As it was beyond the scope of this study to examine tonal-textural and other more complex classifiers, a very simple tonal classifier was used. It can be argued that the results of the simulation, while reduced in scientific validity by this simplification, will still serve their intended purpose, to show the type of imagery that such radar systems would generate.

The reflectivity model used was to classify the ice based on a simple density cut, and to treat each class with a separate roll-off curve with the incidence angle Ø. The curve took the form:

If $Ø < \text{Cutoff Angle}$ $\quad X_{out} = \text{Gain} * X_{in}$

If $Ø \geq \text{Cutoff Angle}$ $\quad X_{out} = \text{Gain} * X_{in} * (255 ** (-1*D*S))$

where $S = (Ø - Ø_{CO})/(90° - Ø_{CO})$
i.e. linear 0 to 1 for $90° \leq Ø \leq Ø_{CO}$

where for each intensity range, the Cutoff Angle ($Ø_{CO}$), Gain (G), and Decay Coefficient (D), were specified. The values were chosen by trial and error, based upon experience with SLAR, and Marine Radars, and by comparison of the data sets collected at an altitudes of 1524 metres and 9100 metres.

The rational behind the cutoff angle portion of this algorithm is as follows: It has been noted that the reflectivity of ice begins to fall off rapidly at some angle between 0 an 10° incidence, and that this exact angle is a function of the ice type. For first year ice and young ice forms such as leads covered in frost flowers, this angle tends to be between 5 and 7°. For older ice it is more shallow, occurring at 3 to 5°. For ridges and other 'deformed' ice types, the falloff occurs near to 0°. Assuming for the moment that simple reflectivity could distinguish between these categories, a cutoff angle is assigned to each intensity range. The simulation consists of calculating the applicable angle, and applying the calculated coefficient.

The Decay coefficient determines the speed of falloff of reflectivity after the cutoff angle has been reached. Again, this is based on observations of Marine radar data, more than on published statistics, and was arrived at in an empirical fashion. The values used were arrived at by comparing filtered versions of Figure 1with data from the low altitude pass.

The decision to apply Gain to certain categories is based again on observations of marine radar data. Features such as ridges, large rubble fields and icebergs often exhibit a marked increase in reflectivity at very shallow angles, relative to the average reflectivity of the scene. Again, the value of the gain used was calibrated against the data collected at 1500 and 9100 m altitude.

The values used in the simulation are relative to the data in Figure 1. The STAR-1 radar is not calibrated to an absolute radiometric standard, although it has been shown to be possible to do relative calibration. Therefore the values arrived at in this simulation have little external relevance but the coefficients used have been tabulated below for reference.

Reflectivity Range	Cutoff Angle	Gain	Decay
0-40	85.0°	1.0	1.0
41-60	86.5°	1.3	1.1
61-90	87.0°	1.5	1.1
91-255	88.5°	1.5	1.2

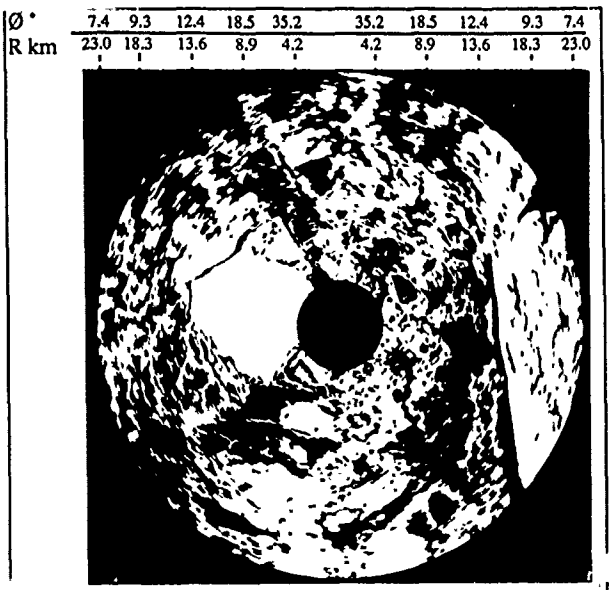


Figure 5 This simulated PPI image has an assumed antenna length of 1 metre, and a near range incidence angle of 35° Simulated Altitude = 3000 m

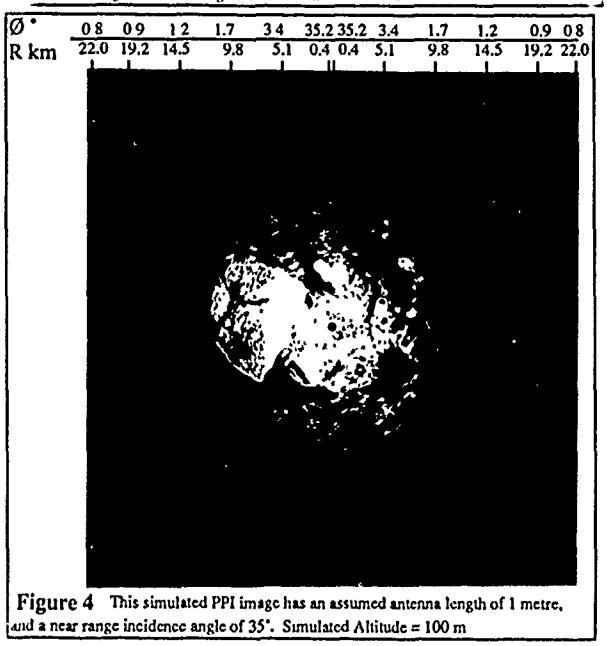


Figure 4 This simulated PPI image has an assumed antenna length of 1 metre, and a near range incidence angle of 35°. Simulated Altitude = 100 m

4.0 Observations on the SAR Simulations.

The simulations used are only approximate, and in the absence of a better ice classifier, there will be errors in the treatment of some ice features. However, the overall effects are quite reasonable, and provide a graphic display of the results of lower altitudes. By calibrating the simulation with the SAR data at 1500 m altitude, and with marine radar imagery (not of the same area), a good representative simulation has been obtained.The simulations therefore represent a suitable means of limiting the amount of flying to be done to conduct preliminary tests of the different styles of radars.

As the opportunities become present INTERA will be collecting data at other altitudes to further verify this empirical algorithm.

4.1 Observations on the SLAR Simulations

Before the simulations started, it was assumed that the fall-off in reflectivity with angle would dominate the utility of the SLAR imagery before the loss of along track resolution, for some antenna lengths, at some altitudes. The question was to establish the altitude, and antenna length compromise which was most appropriate for the helicopter application. Since both the altitude and the antenna length have distinct practical limits, only a subset was examined.

For an altitude of 100 m, the reflectivity fall-off clearly dominated, and even for an antenna length of 1 metre, the SLAR imagery was not seriously degraded in utility over the SAR data. For an altitude of 3000 metres, the azimuth smearing was always apparent. At an altitude of 1000 metres, both the 1 and 2 metre antennas showed azimuth smearing dominated before the loss of reflectivity. With a antenna 3 metres, as shown above, the two effects are both present, and a usable swath width of 10-15 km can be obtained. For a helicopter that will not likely be more than 100 km from its ship, this is not a serious limitation.

4.2 Observations on the PPI Simulations

With an altitude of 3000 m, it was observed that the azimuthal smearing always inhibited the image utility, unless the unwieldy 2 m antenna was used. At 1000 m altitude, the 1 m antenna produced a usable swath width of 5-10 km, but over 360°. This area coverage gives an enhanced impression of the image utility, by, in effect doubling the area coverage compared to a single sided SLAR. Further, a PPI type radar covers the 'hole' under the aircraft by looking fore and aft.

PPI style radars can be used on the helicopter for lice reconnaissance, and for a number of other tasks such as navigation and search and rescue. Since this is beyond the scope of the current study to address all these operational considerations, it is left to the CCG to consider which radar best suits their needs.

4.4 Conclusions

The simulations conducted would indicate that a SLAR or PPI type radar, with characteristics similar to currently available units could provide a local ice reconnaissance capability that would, for the altitudes typical of ship based helicopters, be of similar value to SAR imagery.

5.0 Acknowledgements

The work reported here was conducted by INTERA under contract to the Transportation Development Centre of Transport Canada. The wide swath data in Figure 1 were collected as part of the Ice Atlas study under contract to Canarctic Shipping Company Ltd.

6.0 References

Mercer, J.B. (1989) A New Airborne SAR for Ice Reconnaissance Operations. (These Proceedings)

Ulaby T.U. and M.C.Dobson (1989) Handbook of Radar Scattering Statistics for Terrain. ARTECH House Inc. Norwood MA U.S.A.

3.1 SLAR Simulations

The SLAR simulation is based on the same reflectivity assumptions as the SAR simulations. The difference in the along track resolution however was simulated by having a range dependent low pass filter reduce the effective resolution of the data. The beamwidth of the antenna was estimated from the length of the antenna, assuming a simple un-weighted antenna, using the formula;

$$\beta = \lambda \div L$$

where λ = Radar Wavelength,
L= Antenna Length, and
β= The effective Antenna Beamwidth (in Radians).

The length of the low pass filter used in the along track dimension was calculated based on the range and the beamwidth, using the simple formula;

$$W = (R * \beta) + \rho_a$$

where W is the filter width in pixels,
R is the range to the pixel being filtered, and
ρ_a is the SAR pixel size.

The data were then filtered with a simple box car low pass filter of this size. This resulted in a rather smeared image at far ranges, for a short antenna. For example, the beamwidth of a 1 metre antenna at X band (λ=3.2 cm) is 0.032 radians (1.8°) giving a resolution of 320 metres at a range of only 10 km.With the SAR pixel of 12 metres there is a smearing of more than 26 pixels.

The images generated for the SLAR simulation are based on the same data simulated for lower altitudes for the SAR simulations, but with the low pass filter applied. An example of the simulation is shown in Figure 2. Here the antenna length has been assumed at 3 metres, and the altitude at 1000 metres. This image shows that the resolution limit of the SLAR antenna causes considerable degradation at the far edge of the image. However, before the resolution limit becomes very serious, the loss of information due to the shallow angles begins to dominate. Figure 3 shows the opposite set of conditions: With an effective antenna length of 1 metre, and an altitude of 1000 metres, the data smearing due to the antenna becomes much more severe before the loss of information due to the reduced reflectivity at far range becomes a problem.

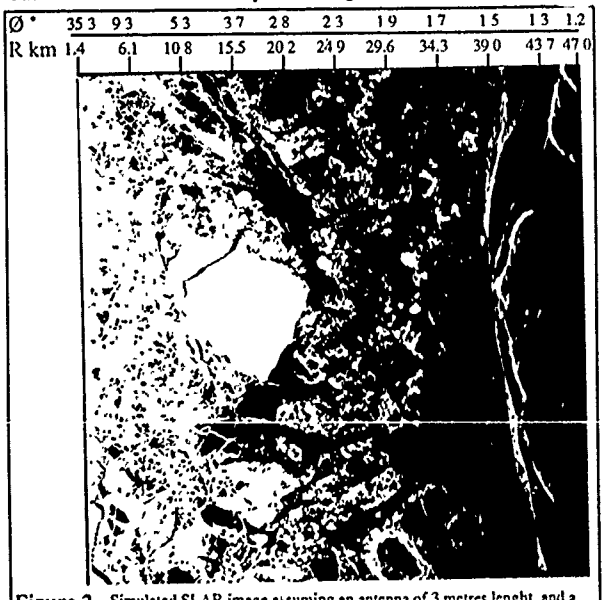


Figure 2 Simulated SLAR image assuming an antenna of 3 metres lenght, and a range pixel size of 12 metres. Simulated Altitude = 1000 m

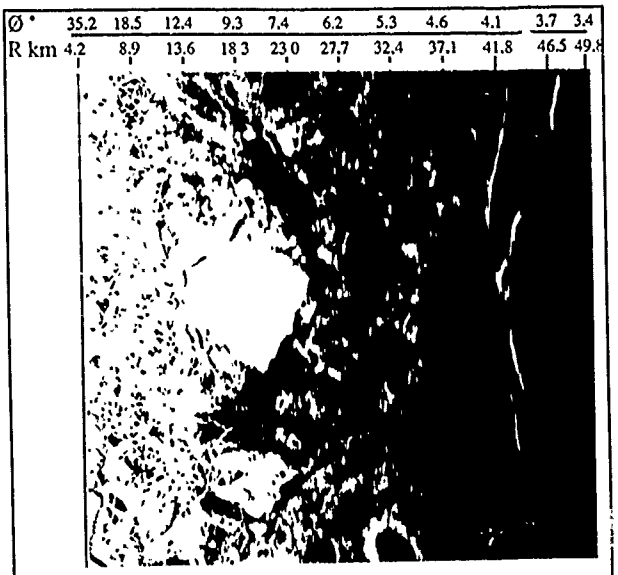


Figure 3 Simulated SLAR image assuming an antenna of 1 metre length, and a range pixel size of 12 metres Simulated Altitude = 1000 m

3.2 PPI Simulations

The Simulation of the PPI radar imagery involved the application of the same algorithms developed for the two previous simulations, but with the scene geometry much modified. A helicopter was assumed to be flown down the centre of the 46 km wide swath of SAR data, at the 4 altitudes selected for simulation, with both a 1 and a 2 metre antenna, for a total of 8 simulated scenes. The data from the SAR were first reduced, by simple averaging to 24 X 24 metre pixels, both to make the computations more tractable, and to simulate a more realistic resolution capability for the PPI radar.

As with the SAR simulations, the intensity of the returns was calculated on the basis of the depression angle, and the scene intensity. The depression angle to be used was calculated from the helicopter location, which was assumed to be at a point in the centre of the original SAR swath. The higher the intensity in the source SAR data, the slower the return was assumed to fall off with angle, as described in Section 3.0. The beamwidth smearing associated with the real aperture of the antenna was again assumed to be the result of a simple averaging, but not along a straight line. Instead, the averaging took place along an arc with centre at the 'helicopter', and with a radius that was the range being simulated. Figure 4, is a simulation of a PPI at an altitude of 100 m with a 1 metre antenna . As with the SAR and SLAR, it was assumed that the antenna illumination was adequate to cover the swath from approximately 30° to at least 5°. For operational radars, this is a very reasonable assumption, as a narrow beam, in the vertical dimension, is needed to make the system performance compatible with the helicopter operation. This angle assumption is relatively unimportant in a single sided SAR/SLAR configuration. However, for the PPI, it determines the size of the 'hole in the centre'. At the higher altitudes, this becomes significant. This is apparent in the image shown in Figure 5, where the altitude is assumed at 3000 m and the antenna is assumed to be 1 metre in length. Although the returns are there from a large area, by the time the near range angle criterion is met, the antenna smearing has already reduced image interpretability.

C-BAND SAR RESULTS OBTAINED BY AN EXPERIMENTAL AIRBORNE SAR SENSOR

R. Horn

Institut fuer Hochfrequenztechnik
German Aerospace Research Establishment (DLR)
Oberpfaffenhofen
D-8031 Wessling , Federal Republic of Germany

ABSTRACT

The experimental airborne synthetic aperture radar system (E-SAR) is under development at the Institut fuer Hochfrequenztechnik of the German Aerospace Research Establishment (DLR). It allows the study of the SAR method and its problems, such as motion error correction and overall system calibration. The sensor is designed to operate on board a Dornier DO 228 aircraft in either L- or C-band (Horn, 1988). A first series of flight experiments in L-band was completed successfully in spring 1988.
The C-band system installation on board the aircraft was completed in October 1988 and first flight experiments were carried out over local test areas in the south of Germany. The C-band front-end represents a first step towards an active array. The amount of quantization and saturation noise is minimized by adapting the received signal power variation to the dynamic range of the A/D converters. Platform attitude and navigation data are collected and recorded on high density tape (HDDT).

Keywords: Airborne SAR, C-band sensor, active array antennas, motion error correction, adaptive AGC/STC, SAR processing, test flight results

1. INTRODUCTION

The C-band SAR front-end development for the E-SAR system was started in 1986. The system goals were a new active array transmitter design to replace a conventional TWT-amplifier and the internal calibration of the sensor hardware. Two separate microstrip array antennas are used for transmit and receive to achieve low direct coupling of transmitted energy into the receiver chain. Both antennas are fixed in azimuth direction. Only the depression angle can be varied mechanically. The azimuth antenna beamwidth is wide to prevent unwanted modulation of the radar signal by yawing of the aircraft. This results in a lower antenna gain and a high pulse repetition frequency (PRF), requiring a high data rate to be stored on HDDT. The main advantage of the antenna design is the wide Doppler spectrum, which is necessary for the motion error correction method, we apply (J. Moreira, 1989).

The radar is presently a single channel system, which accommodates either a nadir or a narrow swath imaging mode. It is operated at rather low altitudes ranging from 800 m to 3500 m above ground. An adaptive automatic gain control (AGC)/sensitivity time control (STC) subsystem was developed to match the received signal power variation to the system's instantaneous dynamic range given by the ADC's (Moreira, Poetzsch, 1988).
Platform attitude and navigation data, necessary to support motion error correction and SAR processing on ground, are acquired by a new auxiliary data acquisition (ADA) unit, encoded as I/Q video data and stored on HDDT.
The system was installed on board the DO 228 aircraft in October 1988 for first test flights. It overflew different local test areas in the south of Germany. The SAR image presented in Fig. 3 shows the area around the town of Sinningen near Kellmuenz a.d. Iller about 30 km south of the city of Ulm.

2. E-SAR C-BAND SYSTEM DESCRIPTION

The system block diagram of the L-/C-band radar sensor is presented in Fig. 1. The RF electronics are shown in the upper part. The L-band front-end with 500 W transmit peak power is presently under development. The C-band front-end in its present configuration consists of eight GaAs-FET high power modules (HPA) each supplying up to 12 W peak and a single low noise amplifier (LNA) in the receiver chain. The transmitted power totals about 90 W peak in the far field of the antenna . The 3.5 dB receiver noise figure results from the combining network of the receiving antenna and the 1.5 dB noise figure of the LNA. Both antennas have identical patterns with 3 dB beamwidths of 17 deg in azimuth and 33 deg in elevation. The antenna gain was measured to be 16.5 dBi.
For internal calibration, the expanded radar pulse, generated by the SAW pulse compression subsystem, is used as a test signal. In C-band we test the up/down converter unit including the receiver. In a second test loop the high power modules are included one by one. The output of the selected module is fed into the receiver. As this requires low signal levels, we have built-in fixed attenuators and eight test probes on the transmitting antenna array, each providing a coupling factor of 20 dB.
The AGC/STC unit controls a slow 7 bit PIN-diode attenuator (0 to 63.5 dB, AGC) at the input of the pulse compression module and a fast voltage controlled GaAs-FET attenuator (0 to 15 dB, STC) at its output. AGC and STC generate the required fast gain variation in IF-band (300 MHz) to adapt the received signal to the dynamic range of the 6 bit ADC's without affecting the analogue pulse compression. This method is used to minimize the amount of saturation and quantization noise produced by the ADC's.
A high precision broadband I/Q-detector converts the

radar IF signal into baseband I and Q video signals.

The lower part of Fig. 1 shows the digital electronics containing two 6 bit flash ADC's, operated at a sampling frequency of 100 MHz, and echo buffer memories with 2560 words capacity providing a 25-times data rate reduction. This function is controlled by the system timing unit (STU), which generates all timing signals required by the radar. It is designed to provide a +/- 30% range of PRF variability for on board real time motion compensation. The PRF and the tape speed of the high density tape recorder (HDDR) can be slaved to the aircraft's forward velocity.
The digitized I and Q video signals are supplied to different digital subsystems.
A real time SAR processor unit correlates the range compressed radar data in azimuth direction in an unfocused SAR mode (A. Moreira, 1989). The processed images are displayed in real time on a monitor and are recorded digitally on video tape.
The radar raw data are formatted in the SAR 580 HDDT format and the LR-NS (low resolution-narrow swath) mode. They are recorded on a 14 track Bell& Howell M14LR high density tape recorder.
The AGC/STC unit requires I and Q video data to generate the optimum IF-gain function. It adaptively equalizes the signal intensity, which is a function of target type and range, incidence angle and antenna pattern.
Via the system control unit (SCU) the radar data are made available to the system processor unit (SPU). The SPU performs the main routines of the real time motion compensation, which are under test. The radar altimeter function is performed by the SCU. The method is based on the detection of the nadir echo and evaluates its time delay. The SCU is furthermore an interface of the SPU to control the radar system operational modes. It reads the data from the high precision barometric altimeter and is connected to the ADA unit.
The ADA unit is a microprocessor controlled interface, which acquires platform attitude and navigation data from a flight data recorder (FDR). The FDR is basically a VLF Omega navigation unit (GNS 1000), which belonges to the aircraft's standard equipment. As this unit does not supply the accuracy required by SAR, we are presently testing a Global Positioning System (GPS) receiver. An Inertial Navigation System is under investigation as well. The navigational data supplied by these units can be read also by the ADA unit. System related data, such as PRF, range delay, operating mode, etc., are acquired from the SCU. A block of auxiliary data (total amount 576 bit) is generated and encoded as I and Q video data. Via the ADA MUX the data are fed into the raw data stream to be formatted and recorded on HDDT. After the transcription to computer compatible tape (CCT) the auxiliary data form a block of 48 complex bytes at the end of each range line, which containes 2 k complex bytes in LR-NS mode.

3. ON GROUND DATA PROCESSING

The outputs of the E-SAR sensor are real time image data on video tape and the radar raw data combined with the aux data on HDDT.

3.1 Radar raw data transcription

A SAR 580 high density tape transcription system transfers the HDDT data onto CCT formatted in the SAR 580 video signal CCT format. We had to modify the system's LR-NS dual chan mode to accommodate a single channel transcrip n mode, which provides the data rate required for the processing of the C-band radar data. In this form th data are available to any SAR processor handling SAR 0 formats.

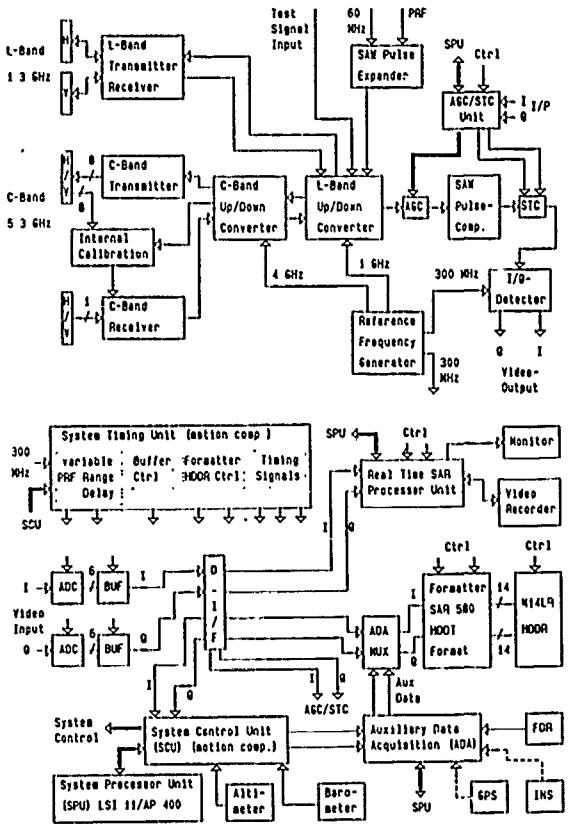


Figure 1: The E-SAR sensor system block diagram.

3.2 Auxiliary data processing

Auxiliary data extraction and processing are the next steps to follow. Attitude and navigation data are separated from the sensor related data. Errors are eliminated by median filtering and the amount of data is reduced. The data set of each parameter is written to a separate file and stored on disk or CCT to be available for off-line motion compensation.

3.3 The off-line motion compensation SAR processor

To find effective motion compensation routines, which can be applied in real time on board the aircraft, we decided to develop an off-line motion compensation SAR processor (J. Moreira, 1989). The block diagram is presented in Fig. 2. The applied motion compensation procedure extracts the motional errors from the range compressed raw data with additional auxiliary data information, especially the altimeter and barometer data. It corrects firstly geometric distortions in azimuth and range direction due to forward velocity variations and line-of-sight (LOS) displacement. After a 4-times azimuth presumming a motion error dependent phase correction is applied on each range line.
The focused multi-look processing of the corrected raw data is fairly straight forward. To generate the azimuth reference function one needs to know the average forward velocity given by the motion compensation procedure, the average range delay, some sensor related parameters, such as the transmitted wavelength, the desired azimuth resolution and the number of looks. The azimuth correlation is carried out in the frequency domain. After the look summation and detection the data rate is reduced by a factor of 8. The image data

in slant range coordinates are stored on disk.

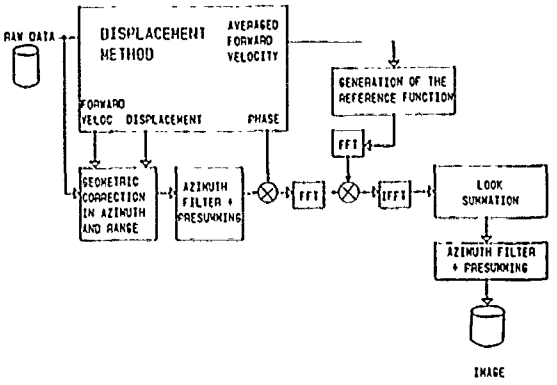


Figure 2: Block diagram of the off-line motion compensation SAR processor.

3.4 Image processing

The equalization of the mean image intensity, which is a function of target type and range, incidence angle and antenna pattern, is the first step in image processing. Although this is done already in real time on board by the AGC/STC unit, we have to repeat this operation, if an IF-gain function , which was evaluated over terrain with low backscatter, was applied on a scene with high backscatter. Figures 4 and 5 show the mean image intensity curve of the slant range image of the scene in Fig. 3 before and after this procedure, respectively.

As the E-SAR radar is operating at low altitudes, the incidence angle varies considerably across the swath (up to 50 deg). If we only corrected the range and antenna pattern dependence of the signal power, we would still get low image intensities in far range due to a low backscatter coefficient, which decreases with the angle of incidence. The operation described above effectively corrupts the radiometric accuracy, but is required to achieve images with uniform intensity. For the evaluation of the radiometric characteristics of the image this operation can be reversed. With the introduction of a range and antenna pattern model a correction curve can be derived and the radiometric accuracy restored.

A slant range to ground range conversion and geometric correction in azimuth direction is the next procedure. The final step is the formatting and recording on CCT for the reproduction on a film recorder FIRE 240.

These procedures were applied on the image shown in Fig. 3.

4. RESULTS

On October 18, 1988 the E-SAR system with the new C-band front-end on board a DO 228 aircraft overflew the Kellmuenz-test site at an altitude of about 2000 m above ground. We selected this scene, because of its well-known distinct features, such as the Iller river and canal. It shows agricultural areas as distributed targets and some tall pylons with high tension cables leading from the hydro-electric power station in near range as point target scatterers. The image was processed on the off-line motion compensation SAR processor with 8 looks and an azimuth resolution of 3 m to match the 3 m range resolution. The average ground speed was 79 m/s. The image signal-to-noise ratio (SNR) was calculated to be about 13 dB from the intensity curve shown in Fig. 5. The different fields can be easily discriminated by their different radar sig-

Figure 3: C-band SAR image of the Kellmuenz-test site processed on the off-line motion compensation SAR processor. The image has a constant mean intensity versus range.

natures. To demonstrate the geometric resolution capability of the system the signature of a pylon in near range next to the lake was selected. Fig. 6 shows a surface plot with the response of two scattering centres on the same pylon. The first peak in Fig. 7 was measured to have a 3 dB-width in slant range of 3.6 m and the second had 2.7 m. Both peaks are about 18 m apart, which corresponds to the hight of the pylon. From the azimuth cut through the first scatterer in Fig. 8 the 3 dB-width in azimuth was evaluated to be about 5 m. The radiometric resolution capability could not be evaluated, because the system was not calibrated.

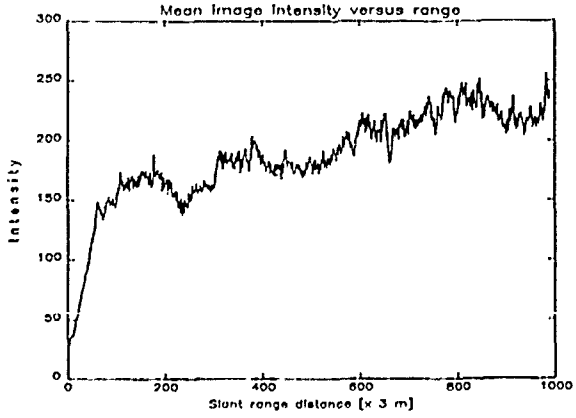


Figure 4: Mean image intensity versus slant range resulting from the AGC/STC real time operation on board.

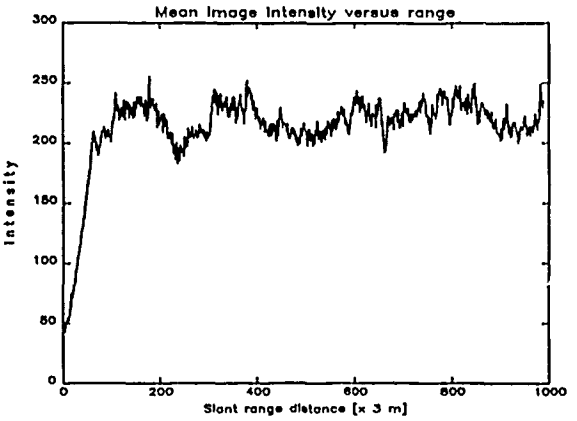


Figure 5: Mean image intensity versus slant range after the on ground applied intensity correction.

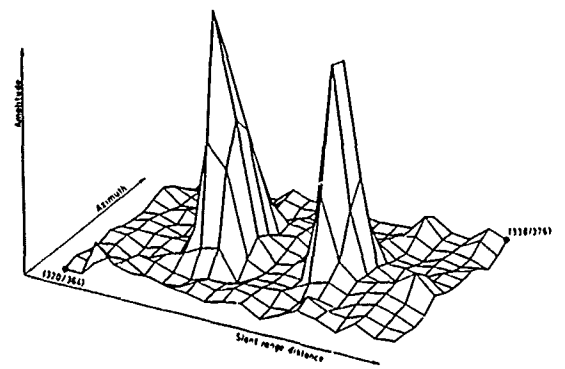


Figure 6: Three dimensional surface plot of the response of two scattering centres on a pylon. The pylon is placed in near range of the image (Fig. 3) near the Iller river at the upper end of the big lake.

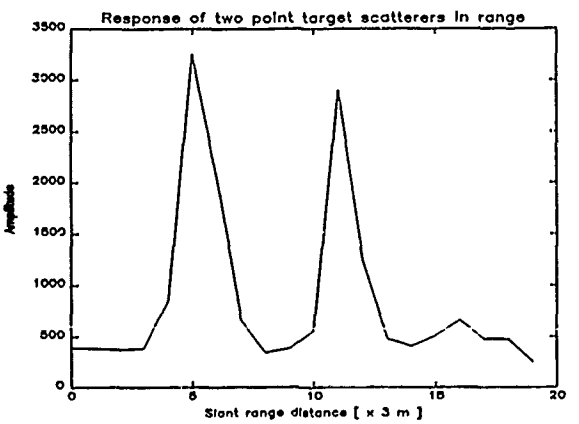


Figure 7: A cut through the response of the two scattering centres on the pylon in slant range.

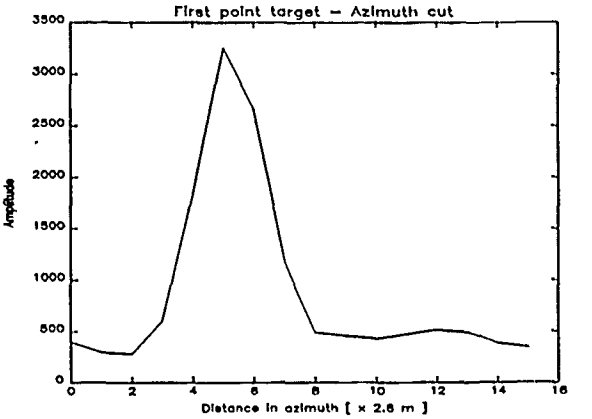


Figure 8: A cut in azimuth direction through the first scatterer on the pylon.

5. CONCLUSIONS

The main objectives of the October 1988 flight experiments were general system test on the integrated C-band subsystems, first system performance tests and to obtain first C-band radar raw data for off-line SAR processing. The results presented in this paper demonstrate the excellent performance of the DLR E-SAR C-band radar system. All image quality requirements aimed at were achieved, although full calibration of the system has not yet been completed. The present goal is to realise a fully calibrated system working in two channels, i.e. either two frequency bands or two polarisations for underflying the ERS-1 and SIR-C/X-SAR instruments as well as for calibration and target classification experiments.

6. REFERENCES

1. Horn, R.,"E-SAR The experimental airborne L/C-band SAR system of DFVLR",Proc. of IGARSS'88 Symposium, Edinburgh, Scotland, Sept. 1988.

2. Moreira J., Poetzsch W.,"Real-time adaptive radiometric correction for imaging radar systems", Proc. of IGARSS'88 Symposium, Edinburgh, Scotland, Sept. 1988.

3. Moreira A.,"Concept and results of the DLR realtime SAR processor", AGARD'89 symposium, The Hague, Netherlands, May 1989.

4. Moreira J.,"A new method of aircraft motion error extraction from radar raw data for real time SAR motion compensation", to be presented on IGARSS'89 symposium, Vancouver, Canada, July 1989.

7. ACKNOWLEDGEMENTS

The author would like to thank the many colleagues who worked with him on this project, but who are too numerous to name individually, with the exception of the the project team members, J. Moreira, A. Moreira and S. Buckreuss.

A NEW METHOD OF AIRCRAFT MOTION ERROR EXTRACTION FROM RADAR RAW DATA FOR REAL TIME SAR MOTION COMPENSATION

Joao R. Moreira

German Aerospace Research Establishment (DLR)

Institute for Radio Frequency Technology

8031 Wessling, FRG

ABSTRACT

This paper reports a new solution for real time motion compensation. The main idea is to extract all the necessary motions of the aircraft from the radar backscatter signal using a new radar configuration and new methods for evaluating the azimuth spectra of the radar signal. Hence an inertial navigation system becomes unnecessary for a lot of applications. The motion compensation parameters for real time motion error correction are the range delay, the range dependent phase shift and the pulse repetition frequency. The motions of the aircraft to be extracted are the displacement in line of sight (LOS) direction, the aircraft's yaw and drift angle and forward velocity. Results show that a three look image with an azimuth resolution of 3m in L-band using a small aircraft is achievable and the implementation of this method in real time using an array processor is feasable.

1. INTRODUCTION

Synthetic aperture radars (SAR) synthesize a long antenna by transmitting electromagnetic energy and coherently adding the successive reflected and received pulses in order to obtain high resolution in flight (azimuth) direction. The resolution in range direction is achieved by transmitting very short pulses or by using pulse compression. To achieve a coherent integration, called azimuth compression, it is necessary that phase errors, resulting from spurious platform motion errors, are compensated. The platform motion error is defined as the error between the actual flight path and the nominal one. For SAR systems mounted on small aircrafts the motion errors are considerably high due to atmospheric turbulence and aircraft properties. Obtaining the motion errors of the aircraft, motion compensation can be realized adjusting the pulse repetition frequency (PRF), applying a range dependent phase shift to each received pulse and delaying it. By adjusting the PRF one compensates for the aircraft forward velocity variations, so that the emissions will accur at constantly spaced intervals. Adjusting the phase and range delay one compensates for the displacement in LOS-direction.

This paper will report a method to extract the displacement in LOS-direction, the aircraft velocity, the yaw and drift angle from the radar raw data. This method is based on the analysis of the azimuth spectrum of the radar raw data. The primary condition to implement this method is the use of a wide azimuth antenna beam. This is obtained using a short fixed mounted pencil beam antenna rather than an usual long stabilized antenna (Boesswetter et al,1983). Thus both the complex gimbaling system and the clutterlock loop is avoided.

2. PROPERTIES OF THE AZIMUTH SPECTRUM

Consider the radar geometry of a strip mapping SAR, where $V_v(t)$ is the aircraft forward velocity, $V_b(t)$ the velocity error in LOS-direction, t is the time, PRF is the pulse repetition frequency, ϑ the angle of the cross-flight direction to the direction of a point on the ground and R the range. The antenna points at a right angle to the nominal flight path illuminating a swath to one side of the aircraft. Due to $V_v(t)$ and $V_b(t)$ the transmitted pulse will suffer a frequency or doppler displacement of:

$$F_{doppler} = \frac{2 \cdot V_v(t) \cdot \sin \vartheta}{\lambda} + \frac{2 \cdot V_b(t) \cdot \cos \vartheta}{\lambda} \qquad (1)$$

where λ is the wavelenght of the transmitted pulse.

Using the radar equation the azimuth power spectrum P(f) can be expressed as the product:

$$S(f) = k \cdot G^4(f) \cdot B^2(f) \qquad (2)$$

where f is the frequency, $G(f)$ the one-way antenna gain in azimuth direction and $B(f)$ the ground reflectivity, considered real. $G^4(f)$, that represents the antenna pattern, can be shifted in frequency by the combination of yaw and drift angle φ and the velocity in LOS-direction $V_b(t)$ as follows:

$$f_a \cong \frac{2 \cdot V_v(t) \cdot \sin \varphi(t)}{\lambda} + \frac{2 \cdot V_b(t)}{\lambda} \qquad (3)$$

where f_a is the frequency shift of $G^4(f)$, $\varphi(t) = \alpha(t) + \beta(t)$, $\alpha(t)$ is the yaw and $\beta(t)$ is the drift angle of the aircraft.

$B^2(f)$, that represents the ground reflectivity, can be shifted in frequency by the velocity in LOS-direction $V_b(t)$ as follows.

$$f_r \cong \frac{2 \cdot V_b(t)}{\lambda} \qquad (4)$$

where f_r is the frequency shift of $B^2(f)$.

The extraction of the motion errors of the aircraft is based on two methods. The first method analyses only the ground reflectivity part of the azimuth spectrum and is called Reflectivity Displacement Method (RDM). The second method analyses only the antenna pattern part of the azimuth spectrum and is called Spectrum Centroid Method (SCM).

3. THE REFLECTIVITY DISPLACEMENT METHOD (RDM)

The reflectivity displacement method analyses the frequency shift between two ground reflectivity functions of adjacent and strong overlapping azimuth spectra. It is considered, that the ground reflectivity function $B(f)$ has a high contrast defined as:

$$K = \frac{\frac{1}{PRF} \cdot \int B^2(f)df}{\left[\frac{1}{PRF} \cdot \int B(f)df\right]^2} - 1 \qquad (5)$$

so that the antenna pattern, that has a wide beam in azimuth direction, can be considered approximately independent of frequency. The frequency shift $V(t)$ between two ground reflectivity functions can be derived from eq. 1 and 4 and expressed as:

$$V(t) \cong -\frac{2 \cdot V_v^2(t) \cdot \Delta t}{\lambda \cdot R} + \frac{2 \cdot \dot{V}_b(t) \cdot \Delta t}{\lambda} \qquad (6)$$

where R the range, Δt the time interval between the two adjacent azimuth spectra. The maximum bandwidth of the frequency shift function $V(t)$ is dependent on Δt and is expressed, considering the sampling theorem, as:

$$BW_v = \frac{1}{2 \cdot \Delta t} \qquad (7)$$

The frequency shift V(t) can be determined from the position of the maximum of the correlation between two adjacent azimuth spectra. This way gives a very accurate frequency shift due to the fact that the two adjacent azimuth spectra are strongly correlated.

As we get V(t) by this correlation, the acceleration in LOS-direction $V_b(t)$ must be separated from the forward velocity $V_v(t)$. This is done using a low and a high pass filter. The forward velocity has a very low bandwidth, for example 0 to 0.05Hz assuming a turbulence with a standard deviation of about 1m/s and a small aircraft as the Dornier Do228. The acceleration in LOS-direction has a much higher bandwidth, where only the higher frequencies are important for motion compensation, for example frequencies higher than 0.1Hz. Considering the power spectral density (PSD) of the acceleration in LOS-direction and the PSD of the forward velocity of the Do228 aircraft, we get an overlap between the two terms of eq. 6 of less than -15dB, so that both the forward velocity and the acceleration in LOS-direction can be well separated from each other. The displacement in LOS-direction is obtained integrating two times the acceleration in LOS-direction. Figure 1 shows the block diagram of the reflectivity displacement method.

The accuracy of RDM depends on the contrast K of the reflectivity function B(f). By imaging land surface that is not homogeneous we get always a high contrast K, that implies a high accuracy of the RD method. By imaging an homogeneous source as sea or desert we get a low contrast K, that implies reduced accuracy of the RD method. The degradation of the accuracy of the displacement in LOS-direction with the decrease of the contrast K is greater than the degradation of the accuracy of the forward velocity.

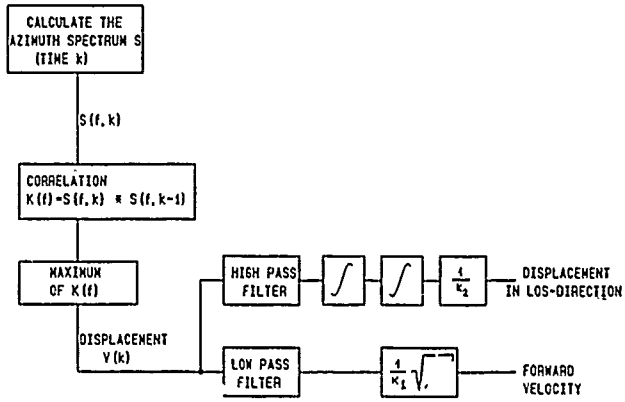


Figure 1. Block diagram of the Reflectivity Displacement Method (RDM)

4. THE SPECTRUM CENTROID METHOD

This method basically determines the doppler centroid of the antenna pattern part $G^4(f)$ of the azimuth spectrum. As the beamwidth of the antenna in azimuth direction is very wide, the method using the azimuth frequency spectrum of the processed complex imagery (Li et al, 1985) has a very low bandwidth, due to the illumination time of the antenna that is longer than 30s in X-band.

As the illumination time of the antenna is long the ground reflectivity can be determined very precisely using a Kalman filter, if the reflectivity does not change with the time very much. Extracting the ground reflectivity from the azimuth spectrum we can determine the doppler centroid with high accuracy and bandwidth. The accuracy of the ground reflectivity prediction depends on its contrast K. If K is high the ground reflectivity varies strongly with the time due to bright targets that always have a variable and not predictable reflectivity. If K is low, the ground reflectivity can be easily determined.

The doppler centroid of the calculated antenna pattern is shifted by $\varphi(t)$ and $V_b(t)$ according to the eq. 3. $\varphi(t)$ is a combination of the yaw $\alpha(t)$ and the drift angle $\beta(t)$ of the aircraft. Yaw motion is caused by turbulence and aircraft instability and the drift angle is caused by wind. The yaw motion has a high bandwidth and the drift angle has a very low bandwidth, so that $\varphi(t)$ has almost the same bandwidth as the velocity in LOS-direction. Thus, they have to be separated considering the geometry of the aircraft. The displacement in LOS-direction, yaw and drift angles are calculated by estimating the doppler centroid of the calculated antenna pattern in near ($C_n(t)$) and far range ($C_f(t)$) and evaluating the data of a high precision barometer ($B_r(t)$) supported by a radar altimeter ($H(t)$). The yaw motion is separated from the drift angle via a high and a low pass filter. Figure 2 shows the block diagram of the spectrum centroid method (SCM).

Finally, the SCM works with high accuracy, when the contrast K of the ground reflectivity is low. When the contrast K increases, the degradation of the accuracy of the displacement in LOS-direction and of the yaw angle is greater than the degradation of the accuracy of the drift angle.

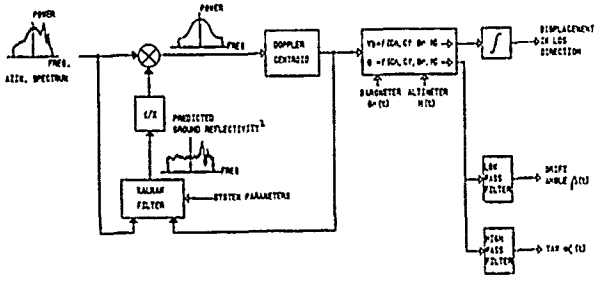


Figure 2. Block diagram of the Spectrum Centroid Method (SCM)

5. COMBINATION OF THE MOTION EXTRACTION METHODS

The combination of the two methods is implemented to increase the accuracy of the displacement in LOS-direction. Thus the accuracy of the displacement in LOS-direction will be almost independent of the contrast K of the ground reflectivity function. This is implemented giving different weightings to both methods according to K.

6. RESULTS

This motion compensation system has been successfully implemented off-line with the raw data of the Experimental SAR System of DLR (Horn,1988). The main features, that this system has, are a high PRF and an azimuth antenna width of 46° in L-band and 17° in C-band. Figure 3 shows an example of motion compensation where the displacement in LOS-direction in near range is represented on the bottom. The main parameters for motion compensation were:

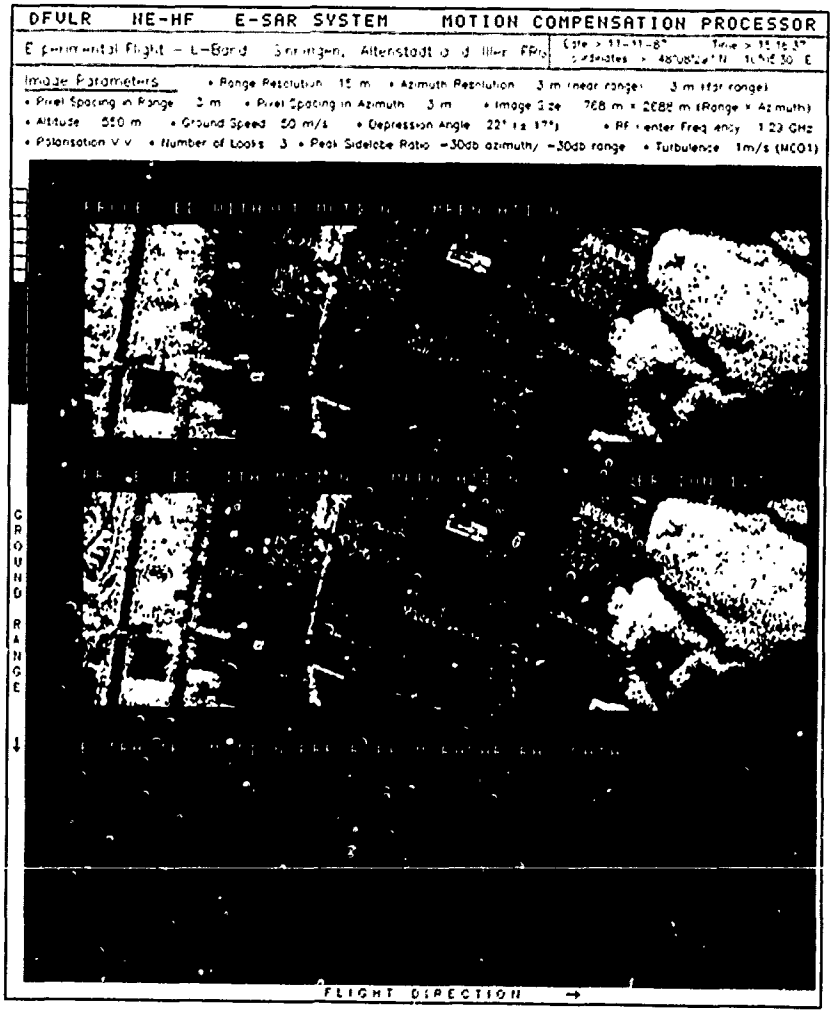


Figure 3. Example of Motion Compensation with the RD-method

- Illumination time > 30s.
- Implementation only of the RD-method.
- Bandwidth in LOS-direction ≅ 0.1 up to 1Hz.
- Bandwidth of forward velocity ≅ 0 up to 0.05Hz.
- Number of range bins for the averaged azimuth spectrum : 32.
- Number of averaged azimuth spectra used : 20.
- Estimated residual error of displ. in LOS : 8mm (1σ).
- Estimated residual error of forw. velocity: 0.1m/s (1σ).

7. REFERENCES

[1] Boesswetter, C. and Wolframm A.,"The design of an experimental synthetic aperture radar (SAR) for small aircrafts", 13th Europ. Microwave Conf.,1983.

[2] Li F.K. et al,"Doppler parameter estimation for spaceborne synthetic aperture radars",IEEE Trans. on Geosc. and Rem. Sens.,Vol.GE-23,47-56,1985.

[3] Horn, R.,"E-SAR - The experimental airborne L/C-band SAR system of DFVLR", IGARSS Symposium, Edinburgh, pp1025-1026, 1988.

8. ACKNOWLEDGEMENTS

The author wishes to thank Mr. C. Dahme, Dr. K. Gruener, Mr. R. Horn, Dr. W. Keydel, Mr. A. Moreira, R. Schmid, U. Spies, M. Werner and F. Witte of DLR, Dr. C. Boesswetter, Mr. T. Pike and A. Wolframm of MBB-Ottobrunn and Dr. T. Freeman of JPL, California, whose suggestions and contributions made this work possible.

The Development of a Polarimetric SAR for the Detection of Man-Made Objects

Ron Saperi, Maria T. Reyii, Rick Wintjesi. C.E. Livingstoneiii
Malcolm R. Vantii, Robert Austiniii

i Atlantis Scientific Systems Group Inc., ii Department of National Defence iii Canada Centre for Remote Sensing,
iiii Canadian Astronautics Limited.

Abstract

Polarimetric Synthetic Aperture Radar (SAR) has gained in popularity in recent years because of its ability to measure the complete backscattering characteristics of a resolution cell This data may be used to synthesize images which would be obtained under various transmit/receive polarization states. The work of Zebker et al has shown that the contrast of man-made objects against a clutter background can vary dramatically under changing synthetic polarizations

The Canadian Department of National Defence, the Canada Centre for Remote Sensing, Canadian Astronautics Limited and Atlantis Scientific Systems Group Incorporated have developed a multifrequency polarimetric SAR acquisition and analysis system and are currently evaluating its use for the detection of small man-made objects in the presence of clutter. The current paper outlines the architecture of the SAR polarimeter, calibration findings, analysis software and results to date.

Key Words: SAR, Polarimetry, Man-Made Objects, Calibration

1. Introduction

Polarimetric Synthetic Aperture Radar has been the subject of intense scrutiny in recent years because of its ability to measure the complete backscattering characteristics of a resolution cell. In an ideal polarimetric SAR these measurements are obtained by acquiring co- registered data at orthogonal polarization states, typically the HH, VV, HV and VH linear polarization combinations. One application of polarimetric data is to calculate the image which would result from arbitrary combinations of transmit and receive polarization states (Zebker,1986).

The Defence Research Establishment Ottawa (DREO), in conjunction with the Canada Centre for Remote Sensing (CCRS), Canadian Astronautics Limited and Atlantis Scientific Systems Group Incorporated have developed a multi-frequency polarimetric SAR instrument and data analysis system During the jointly funded DREO-CCRS polarimetry experiment (POLEX) which took place April to August 1988, equipment built by Canadian Astronautics Limited was successfully integrated with the Canada Centre for Remote Sensing (CCRS) X/C band SAR and placed on the CCRS Convair 580 aircraft Data were collected at X band over suburban, urban and rural terrain near Petawawa, Ontario, with the intent of evaluating the usefulness of polarimetric data and polarization synthesis for the enhanced detection of small man-made targets in the presence of clutter. The raw data for each of the four orthogonal linear polarizations were processed to complex image form by DREO These were then used to synthesize real images of any desired polarization, using software developed by Atlantis Scientific Systems Group Inc.

2. Structure of the Polarimetric SAR

The radar system for the polarimetry experiment consists of two synthetic aperture radars, one normally operated at 5.31 GHz (C Band) and the other normally operated at 9.25 GHz (X band). The two radars share a common architecture which allows the system to be modified for operation as a SAR polarimeter at either of the two frequencies The two receivers are fully phase coherent and synchronized via a single master oscillator A simplified block diagram of the SAR polarimeter is shown in Figure 1

The antennas use horizontally and vertically polarized feeds to drive a single radiating aperture The isolation between the horizontal and vertical polarizations is greater than 40 dB at the antennas The radar is operated at twice the normal PRF so that horizontally and vertically polarized signals are transmitted alternately when power is routed through a 4 port ferrite circulator switch to the appropriate antenna feed During the receive window the circulator routes returned energy to the appropriate receiver. Receiver A is the like-channel receiver and processes HH and VV channels multiplexed in time Receiver B is the cross-channel receiver and processes HV and VH channels multiplexed in time. The transmit circulator shares the same port as Receiver A. For the experimental polarimeter described in this paper the like and cross polarization isolations are limited to about 20dB for the cross-over switch used.

The baseband output of the receivers is routed through the polarimetry interface unit which buffers and splits the I and Q outputs to the two dual channel processor/control units (PCU) which perform real-time motion compensation by phase-rotating the complex returns The Analog to Digital Converters (ADC's) of the PCUs operate alternately such that PCU1 processes VV & VH data and PCU2 processes HH & HV data. The PCUs perform motion compensation on the data using identical algorithms driven from the same inertial reference.

An advantage of this architecture is that it permits different gains to be applied to the like and cross returns in the RF signal path and thus makes best use of the ADC dynamic range windows

3. Calibration Issues

Accurate polarization synthesis requires accurate *relative* magnitude and phase measurements of the scattering matrix components The synthetically polarized image picture element (pixel) is formed as.

$$V = K_{xx}S_{xx} + K_{cross}S_{cross} + K_{yy}S_{yy} \qquad (1)$$

where the K's are the factors resulting from multiplying out the transmit and receive polarization state coefficients and the S s are the scattering matrix coefficients. (Note that S_{xy} and S_{yx} have been replaced with S_{cross} as

per the reciprocity assumption (Raney,1987) In this work, target contrast against a clutter background under changing synthetic polarizations is of primary interest, and absolute calibration is not that important When even *relative* calibration is not attained the effect may be considered equivalent to having incorrect values of K_{xx}, K_{cross} and K_{yy} Changes in target contrast may still be observed, but the correspondence between the specified and effective polarizations will not be known

Figs. 2a, 2b are signal path diagrams that can be used to model the system Switch leakages and insertion losses, phase path lengths, scattering coefficients and complex admittances are all modelled using these diagrams. In principle, all the scattering matrix coefficients contribute to the phase and amplitude observed by each of the four channels.

From the signal path diagrams it is a simple matter to obtain expressions for the phase and amplitude seen by the four channels. These quantities have arbitrary units since we are concerned only with relative amplitudes and phases Eqn 2 is the channel equation for the HV signal path as illustrated in figure _ The VH, HH and VV channel equations are similar Our approach is similar to, though different in detail from the work of Derryberry et al (Derryberry,1988).

$$
\begin{aligned}
V_{hv} = K\{ \; & I_{3-2}\, e^{j\Phi_{2-h}}\, G_h\, S_{xx}\, G_h\, e^{j\Phi_{2-h}}\, L_{2-1}\, R_{cross}\, P_{hv}\, e^{j\Phi PCU2} \; + \\
& I_{3-2}\, e^{j\Phi_{2-h}}\, G_h\, S_{xy}\, G_v\, e^{j\Phi_{4-v}}\, I_{4-1}\, R_{cross}\, P_{ht}\, e^{j\Phi PCU2} \; + \\
& L_{3-4}\, e^{j\Phi_{4-v}}\, G_v\, S_{yx}\, G_h\, e^{j\Phi_{2-h}}\, L_{2-1}\, R_{cross}\, P_{hv}\, e^{j\Phi PCU2} \; + \\
& L_{3-4}\, e^{j\Phi_{4-v}}\, G_v\, S_{yy}\, G_v\, e^{j\Phi_{4-v}}\, I_{4-1}\, R_{cross}\, P_{ht}\, e^{j\Phi PCU2} \} \quad (2)
\end{aligned}
$$

where

V_{hv}	=	The complex HV image value.
K	=	A constant which includes range and elevation angle effects for a given resolution cell.
S_{xx}	=	The scattering coefficient for x incident, x scattered
S_{yy}	=	The scattering coefficient for y incident, y scattered
S_{xy}	=	The scattering coefficient for x incident, y scattered
S_{yx}	=	The scattering coefficient for y incident, x scattered
I_{m-n}	=	Switch insertion loss and phase, port m to n path
L_{m-n}	=	Switch leakage and phase for port m to n path.
R_{cross}	=	Cross RF path complex gain.
P_{hv}	=	Polarimetry interface gain for HV path.
G_h	=	Scalar H antenna gain.
G_v	=	Scalar V antenna gain.
Φ_{2-h}	=	Phase path length: switch port 2 to H antenna.
Φ_{4-v}	=	Phase path length: switch port 4 to V antenna.
Φ_{PCU2}	=	Phase error due to phase correction unit 2.

The switch leakage and insertion losses, antenna waveguide phase path lengths and antenna gains are self-explanatory, though it should be kept in mind that the leakages and admittances must be complex numbers to account for both attenuation and path length. The R_{cross} (and the corresponding R_{like}) admittance is a lumped quantity which includes the effects of the microwave plumbing from the switch to the cross (like) receiver as well as the coherent receiver itself.

P_{hv} represents the gain introduced by the polarimetry interface unit This unit buffers and splits the I and Q outputs to the PCUs

The Φ_{PCU2} phase change (and the corresponding Φ_{PCU1} phase change for the VV and VH channels) is a phase offset introduced by PCU 1 as part of the motion compensation algorithm. Details may be found in (Living,1988)

Accurate losses and approximate phase path lengths have been measured for the cross-over switch. More precise measurements are in progress. On the basis of the switch isolation measurements alone we may simplify the channel equations considerably.

The isolation from ports 3-2 and 3-4 was at least -28 dB, and thus leakage terms in the channel equations which include L_{3-2} or L_{3-4} can be considered negligible for the target detection problem. For the simple proof-of-concept experiment performed all leakage terms less than -20dB may be ignored.

Further terms may be dropped from the VV and HH channel equations where lesser switch isolations (measured to be on the order of -17dB) appear in the terms with S_{xy} or S_{yx}. It is well known that the returns from cross components are nominally 8 dB below the like component returns The combined effect of switch isolation and the lower expected cross pol return when multiplied together in equation 2 exceeds the -20dB criteria for rejection we have set.

For the like channels we have only a single dominant term, while the cross channels require two terms. In equation 3 we show the simplified two-term HV channel equation in a form where the scattering matrix coefficient S_{xy} has been isolated on the left hand side of the equal sign Similar expressions have been developed for the other three channel equations.

$$
\begin{aligned}
S_{xy} \; = \; & \frac{V_{hv}}{K\, I_{3-2}\, I_{4-1}\, G_h\, G_v\, e^{j(\Phi_{2-h}+\Phi_{4-v})}\, R_{cross}\, e^{j\Phi PCU2}\, P_{hv}} \quad (3) \\
& - \; S_{xx}\, e^{j(\Phi_{2-h}-\Phi_{4-v})}\, \frac{G_h\, L_{2-1}}{G_v\, I_{4-1}}
\end{aligned}
$$

Equation 3 can be simplified combining all of the system parameters as in equation 6 below. For the purposes of end-to-end, in-the-field calibration we may simplify all four channel equations by combining the system parameters into six complex quantities which may be deduced empirically.

$$
\begin{aligned}
S_{xx} &= Q_{hh}\, V_{hh} && (4) \\
S_{yy} &= Q_{yy}\, V_{vv} && (5) \\
S_{xy} &= Q_{hv}\, V_{hv} + R_{xx}\, S_{xx} && (6) \\
S_{yx} &= Q_{vh}\, V_{vh} + R_{yy}\, S_{yy} && (7)
\end{aligned}
$$

For an experiment with end-to-end calibration, the correct approach for determining the lumped parameters in eqns 4 through 7 is to deploy calibration targets with known scattering properties, observe the polarimetric image set and solve for the lumped Qs and Rs. We have four complex equations per calibration target and six complex unknowns, therefore we need at the very least two good calibration targets.

There is a subtle difficulty in solving these equations. The equations form a consistent set (i.e. simultaneous equations) for *only a single resolution cell.* This is because the parameter K in eqn. 2 contains a magnitude and phase factor which is common to all the channel equations only for a given resolution cell. While the scaling varies little across adjacent range cells, the phase offset, which is caused by the fine-scale detail in range, is essentially random from cell to cell.

The polarimetric measurements from a pair of calibration targets must be aligned in phase before they may be used to solve equations 4 through 7. The phase correction will express the phase of the measurements relative to a like channel which is non-zero for both targets. A like channel is used because the like channels are less prone to leakage terms. Targets should be selected which have a strong common return in one of the like channels in order to permit phase-alignment of the resulting equations. This is essentially the same result which Sheen et al have found for the P3-SAR (Sheen,1988).

4. Polarization Workstation Software

PSAR was developed to provide a versatile polarization synthesis capability using recently introduced micro-computer technology

PSAR was developed on a 80386/25 based micro-computer with a Weitek mW3167 math co-processor, 13 megabytes of RAM, a VGA graphics adapter and a 300 megabyte hard disk. The software was designed to get the maximum benefit from the 386 processor, numeric coprocessor and large memory. All image data is directly accessed from memory during processing which eliminates disk accesses or virtual disk accesses from random access memory Rapid animation of polarization sequences is realized using this system.

The PSAR user interface allows one to specify initial transmit and receive antenna polarization states plus incremental changes in these states In this way images can be synthesized at regularly spaced intervals along the surface of a Poincare sphere for any number of iterations

Image display is accomplished using a VGA graphics adapter in the standard 640x480 pixel, 16 colour graphics mode The time required to display a 512x480 image, without synthesis, is less than 3 5 seconds The time required to synthesize a 512 by 480 image is approximately 6 seconds

Status of the polarization configuration is supplied to the user via a report window to the right of the image. Information reported to the user includes a colour bar, current iteration number, maximum allowable iterations, and a graphical display of transmit and receive polarization ellipses (see the right hand side of the display screen in Figs 3a and 3b) When a small window of interest is specified the polarization animation speed becomes very rapid.

Rough calibration parameters were determined to account for the physical characteristics of the SAR recording instrumentation Based on these calibration parameters a small portion of the SAR data acquired over the city of Petawawa was prepared for a preliminary viewing To date a number of points of interest have been identified. A sample of these results is discussed below

5. Preliminary Results

Data collected over Ottawa and Petawawa, Canada, were processed by the Department of National Defence using their SAR-II processor (Vant, 1984). The data for the four linear polarizations was processed to 5 6m cross range by 6m slant range Sensitivity time control was turned off during acquisition to eliminate possible sources of unwanted modulation.

A 150 by 150 pixel region which is an enlarged display from the PSAR program is shown in Fig. 3a. The bright object in the centre of the image is the return from a transmission tower near the town of Petawawa The image is the result of polarization synthesis with a linear transmit polarization at 135° and a linear receive polarization of 45° This polarization state was observed to produce the brightest spot for the tower return

Fig. 3b shows the same data but with the transmit polarization rotated to 210°. The target is lost in clutter. The like polarizations produce similar results. In the cross polarized image the target is visible, but is not as bright as in Fig 3a. This result is not unique — contrast variation of points is observed throughout the data set.

Since calibrated targets will not be imaged until May of 1989, the calibration of the present data set is only approximate and the polarization views do not correspond exactly to those specified by the polarization state vectors The polarization synthesis experiments show results which are qualitatively as expected — small objects have been observed which exhibit varying contrast under the changing but imprecisely defined polarization states

6. Conclusions and Future Work

The modified SAR has been shown to be capable of producing data sets in which scene elements behave qualitatively as we would expect. An analysis of calibration issues suggests that corrected, quantitatively accurate data will be obtainable given successful imaging of a properly laid-out calibration field. The approach will be to empirically estimate a small number of lumped system parameters identified from theoretical considerations as described in this paper

Preliminary results which show variable contrast of man-made targets under changing synthetic polarization states have been encouraging. Since precise calibration has not yet been achieved one may not draw quantitative conclusions about the causes of this contrast variation. The observation of these variations shows, however, that polarization synthesis can indeed enhance the detectability of man-made objects in our data sets.

It should be noted that X band frequencies impose limitations upon the types of terrain in which small man-made targets may be detected. If a small target is within a region of dense forest canopy, detection at X band is doubtful because of the severe attenuation of X band signals by vegetation Longer wavelengths would be more useful if canopy penetration is required.

References

[1] Derryberry, B; Blanchard, A and Kaupp, V, *Examining the Polarisation Phase Retrieval of a Dual Channel Multipolarisation Radar*, Proceedings of IGARSS '88, Edinburgh Scotland, 13-16 September, 1988.

[2] Livingstone, C E , Gray, A L , Hawkins, R.K. and Olsen, R.B., *CCRS C/X Airborne Synthetic Aperture Radar An R & D Tool for the ERS-1 Timeframe*, IEEE Aerospace and Electronics Systems Magazine, V3, No. 10, October 1988

[3] Raney, R.K., *A Free 3dB in Cross Polarized SAR Data*, Proceedings of IGARSS '87 Symposium, Ann Arbor, 18-21 May 1987.

[4] Sheen, D R, Kasischke, E S and Shuchman, R A, *Phase Calibration of Polarimetric SAR*, Proceedings of IGARSS '88, Edinburgh Scotland, 13-16 September, 1988.

[5] Vant, M.R and Wu, K H., *A Digital SAR Processor Based on the Coherent Sub-Aperture Addition Technique*, Proceedings of Int'l Conference on Radar, Versailles, France, May 1984.

[6] Zebker, H A, Van Zyl, J.J & Held, D N., *Imaging Radar Polarimetry from Wave Synthesis* Proceedings of IGARSS '86 Symposium, Zurich, 8-11 September, 1986.

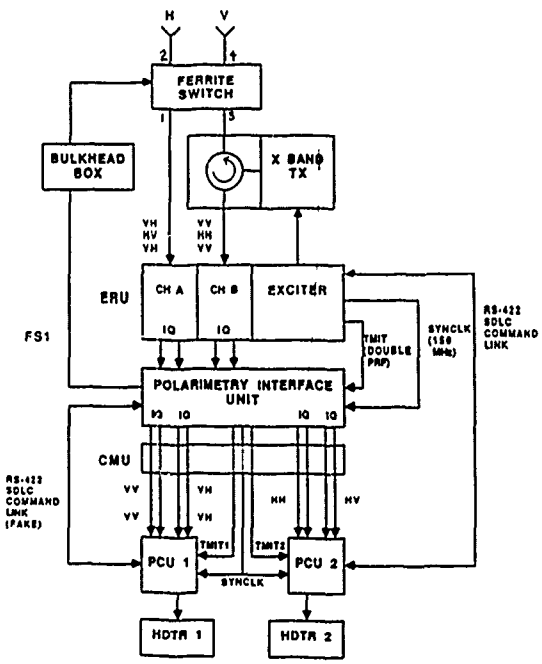
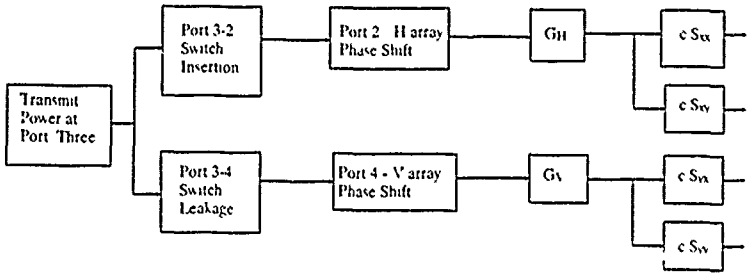


Fig 1. Structure of the Polarimetric SAR



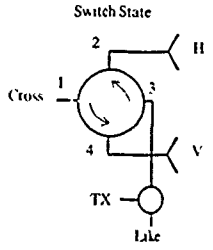
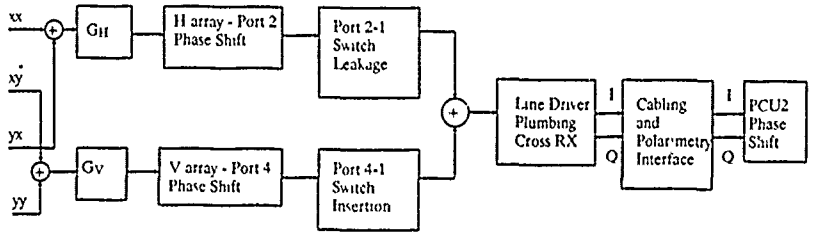


Fig 2a. H Transmit Signal Path



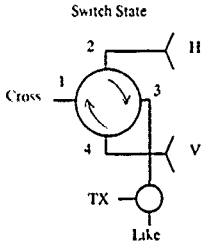


Fig 2b. HV Receive Signal Path

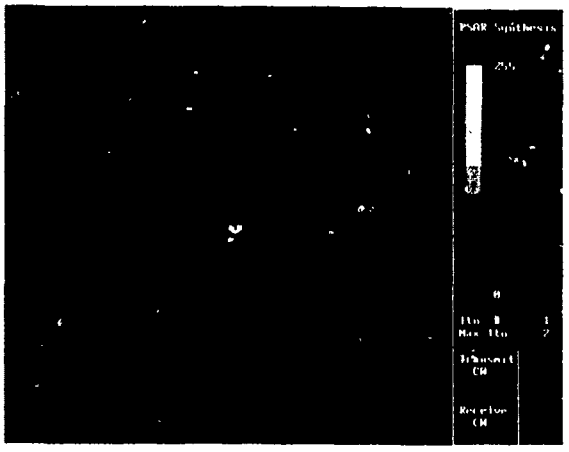


Fig 3a. High Contrast Polarization View of Tower

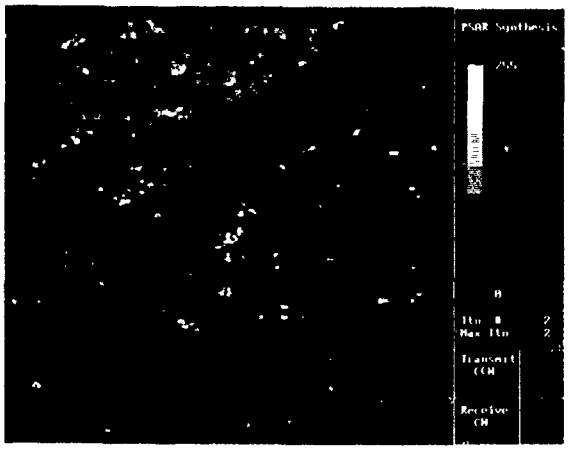


Fig 3b. Low Contrast PolarizationView of Tower

A REALTIME QUICK LOOK PROCESSOR FOR AIRBORNE SAR

Alberto Moreira

German Aerospace Research Establishment (DLR)
Institute for Radio Frequency Technology
D-8031 Wessling, Federal Republic of Germany

ABSTRACT

A Realtime Azimuth Processor was developed for the airborne E-SAR System (Experimental Synthetic Aperture Radar) of DLR. The processor works with a modified unfocussed compression method. This method greatly simplifies the data processing and is easily implemented by a moving average approach. Images processed in realtime are presented and have good contrast and strong suppression of the sidelobes. The processor hardware can be implemented with reduced costs in small aircrafts.

Keywords: Synthetic Aperture Radar (SAR), Signal Processing, Realtime Operation, Algorithms.

1. INTRODUCTION

The use of synthetic aperture radar for remote sensing has increased largely in recent years due to the all-weather, wide swath and high resolution capabilities. Image formation from SAR data is well known and involves coherent processing between the received signal and the matched impulse response function, which is also called the reference function (Fitch, 1988; Beasley, 1982).

The E-SAR System of DLR (Horn, 1988) works experimenteily in L-band without pulse compression (100 ns pulse duration) and in C-band with a FM pulse. In the latter case, range compression is performed by a SAW device. An azimuth processor was developed for the E-SAR System to produce realtime images. It offers some important advantages:

- the complete SAR sensor hardware can be tested in realtime,
- areas illuminated by the radar are monitored in realtime, facilitating for example the detection of oil spills at sea,
- the processed image contains only a reduced amount of data due to mulitlooks or to a smaller processed bandwidth. Hence the storage can be reduced in capacity and realised by a conventional video recorder,
- the realtime images can be used as a reference for the high resolution off-line data processing.

2. CONCEPT OF THE REALTIME PROCESSOR

For small apertures ($\varphi < 20°$ in figure 1), the phase history of a point target is a quadratic function of the transmitted pulse nummer N and is given by (no squint angle is considered):

$$(2.1)\quad \Phi(N) \cong \frac{2\cdot\pi\cdot V^2\cdot(PRI\cdot N)^2}{\lambda\cdot R_0}\ ,$$

where PRI is the pulse repetition interval and R_0 is the minimum range between aircraft and point target. Differentiating the phase history yields, for a sufficiently small PRI, doppler frequency as a function of pulse number, as follows:

$$(2.2)\quad f(N) \cong \frac{2\cdot V^2\cdot PRI\cdot N}{\lambda\cdot R_0} \cong \frac{2\cdot V\cdot\varphi}{\lambda}\ .$$

The assumed doppler bandwidth for processing can then be expressed as a function of the aperture φ:

$$(2.3)\quad B = \frac{4\cdot V\cdot|\varphi_{max}|}{\lambda}\ ,$$

$$|\varphi_{max}| \le \frac{\text{azimuth 3 dB beamwidth of the antenna}}{2}.$$

The developed azimuth processor works with an unfocused processing method. In this case it processes only a small bandwidth of the backscattered azimuth signal. The azimuth signal is correlated with a simple rectangular function which does not correct its quadratic phase history. The correlation at doppler frequency zero is performed with an acceptable maximum phase error of 90°. This means a maximum two way range variation, $2\cdot[R(\pm N) - R_0]$, of $\lambda/4$.

The unfocused processing method is based on the fact that the accepted phase error of 90° leads to sufficient image quality. For the determination of $|\varphi_{max}|$, the one way range variation is limited by $\lambda/8$.

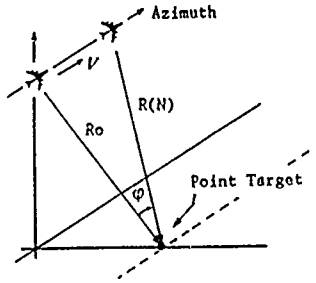
$$(2.4)\quad |\varphi_{max}|_{unfoc} = \frac{1}{2}\cdot\sqrt{\frac{\lambda}{R_0}}\ .$$

The azimuth spatial resolution, ρ, is indirectly proportional to the processed doppler bandwidth and is

obtained by substituting (2.4) in (2.3).

$$\rho = \frac{\sqrt{\lambda \cdot R_0}}{2} \qquad (2.5)$$

The azimuth resolution increases with the square root of range and is not constant as in the case of the focused processing method.



where $\varphi \leq \frac{\alpha}{2}$, $\alpha = \frac{\lambda}{D}$,

$2 \cdot \varphi$ = *processed aperture*,

α = *azimuth beamwidth* of *the antenna*,

D = *azimuth length* of *the antenna*,

λ = *wavelength* and V = *aircraft speed.*

Figure 1: Range variation between aircraft and point target for a SAR.

Since the reference function is rectangular, the correlation process can be carried out very easily by the moving average approach, where the most recent value is added and the oldest value is subtracted for each correlated point. The computational requirements for this operation are greatly reduced and the processor realisation is very much simplified in comparison with the focused case.

2.1 The Point Target Response

Figure 2 left shows a simulated point target response (L-band, R_0 = 1500 m) that was obtained with the unfocused correlation method. The response has high sidelobes with a peak sidelobe ratio (PSLR) of -9 dB and an integrated sidelobe ratio (ISLR) of -7 dB. An image processed with the unfocused method has low contrast and gives a blurred impression.

A new algorithm was developed, so that a triangular amplitude weighting could be incorporated in the unfocused processing method without additional complications. The algorithm consists basically of correlating the azimuth signal twice with a single moving average approach. This corresponds to a single correlation with a triangular function, which is the weighting function. Several simulations were done to optimize the processing parameters. The critical parameter to optimize is the duration of the reference function. A shorter reference function reduces the phase error of the processing, but the resolution becomes worse due to the lower time bandwidth product. An extremely long reference funtion causes a degradation of the image quality due to the greater phase error. The length of the reference function was chosen to achieve a good compromise between PSLR, ISLR, resolution and processing loss. Good results were obtained due to the fact that the triangular reference function gives little weight to the correlation with the largest phase errors. The optimised point target response is shown in figure 2 right. The processing loss due to weighting and phase errors totals about 2.5 dB. The weighted response has a PSLR of -28 dB and a ISLR of -13.9 dB and is superior to the non-weighted response.

The realtime azimuth processor is composed of units performing the following functions: corner turning (Transposing of the received data), correlation with the reference function (moving average approach), resampling, weighting, image detection, output interface and test pattern generation. There is no presumming of the data before azimuth processing due to the fact that the moving average itself is a filter with a $\sin(x)/x$ frequency response. The realtime azimuth processor was developed providing two modes of operation: 1. traditional unfocused processing and 2. unfocused processing with triangular weighting (Moreira, 1989). Comparisons can be made in realtime to show the improved image quality of the new approach. The power consumption totals only 18 W due to the use of modern CMOS circuits.

3. RESULTS

Interesting results were obtained with the realtime processor. The images were postprocessed on a Micro Vax II computer in oder to reduce the speckle noise

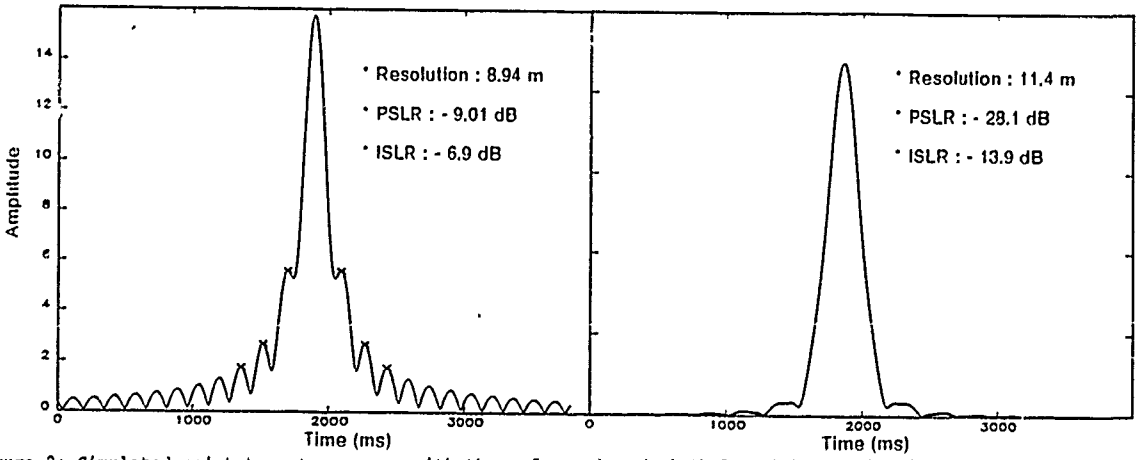


Figure 2: Simulated point target response with the unfocused method (left: without weighting; right: with weighting).

(local statistics filter by Lee, 1986) and to perform geometric and radiometric correction (conversion of slant range to ground range and compensation of the reduction in backscattered signal intensity with range, respectively).

The lake, Ammersee, near Munich is shown in figure 4. The upper image is the continuation of the lower image. Together, the image dimensions total about 3x12 km. It is possible to see some sailing boats as well as different structures on the water surface caused by

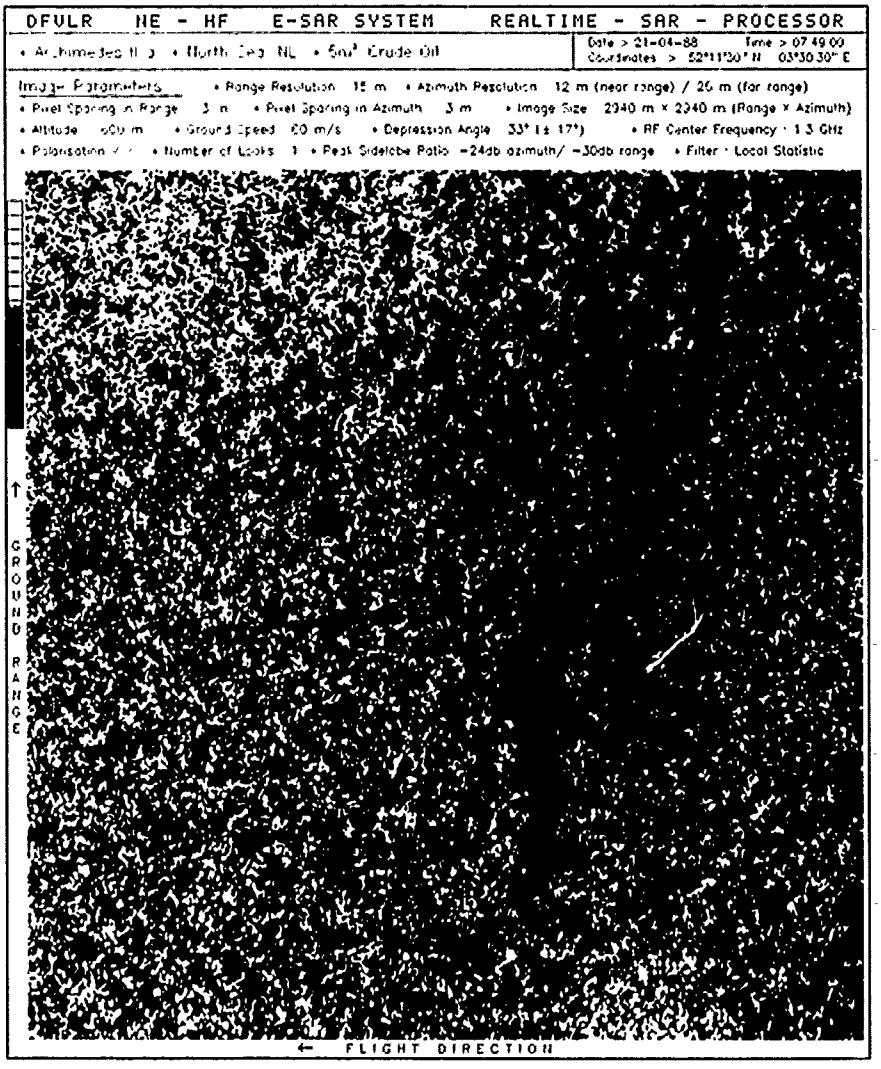


Figure 3: Image processed in realtime: oil pollution on the sea.

Figure 3 shows a 5 m^3 oil spill (dark region) that was detected during the Archimedes IIa campaign at the Dutch North Sea Coast (Moreira, Horn 1988). Radar is a able to detect oil pollution on the sea surface by virtue of the attenuation of capillary waves by the oil spill. The ship, which discharged the oil intentionally for this experiment, can be seen at the top of this figure as a large bright spot due to its strong radar reflectivity. There ist up to 9 dB dynamic range between sea and oil surface intensity in the final image. Although the transmitted peak power was only 50 W with 100 ns pulse width in L-band it was possible to achieve a good image contrast. This contrast arises from the partially specular behaviour of the backscattered signal from the sea which becomes a gain in the correlation process. The signal to noise ratio is 0 dB in mid range for this image.

the wind. The bright area at far range corresponds mostly to forests and small towns.

4. CONCLULSION

The processing of SAR data with the proposed algorithm can be performed in realtime with relatively simple hardware. Although the supression of th sidelobes of this modified, unfocused, processing method is comparable to the focused case, the azimuth resolution is limited because only a small part of the doppler bandwidth is processed. The azimuth resolution can be improved by flying at low altitudes and by the use of small radar wavelengths.

The implementation of multi-looks is also possible with this algorithm but we have opted for a single look processing and off-line filtering of speckle noise with the local statistics algorithm.

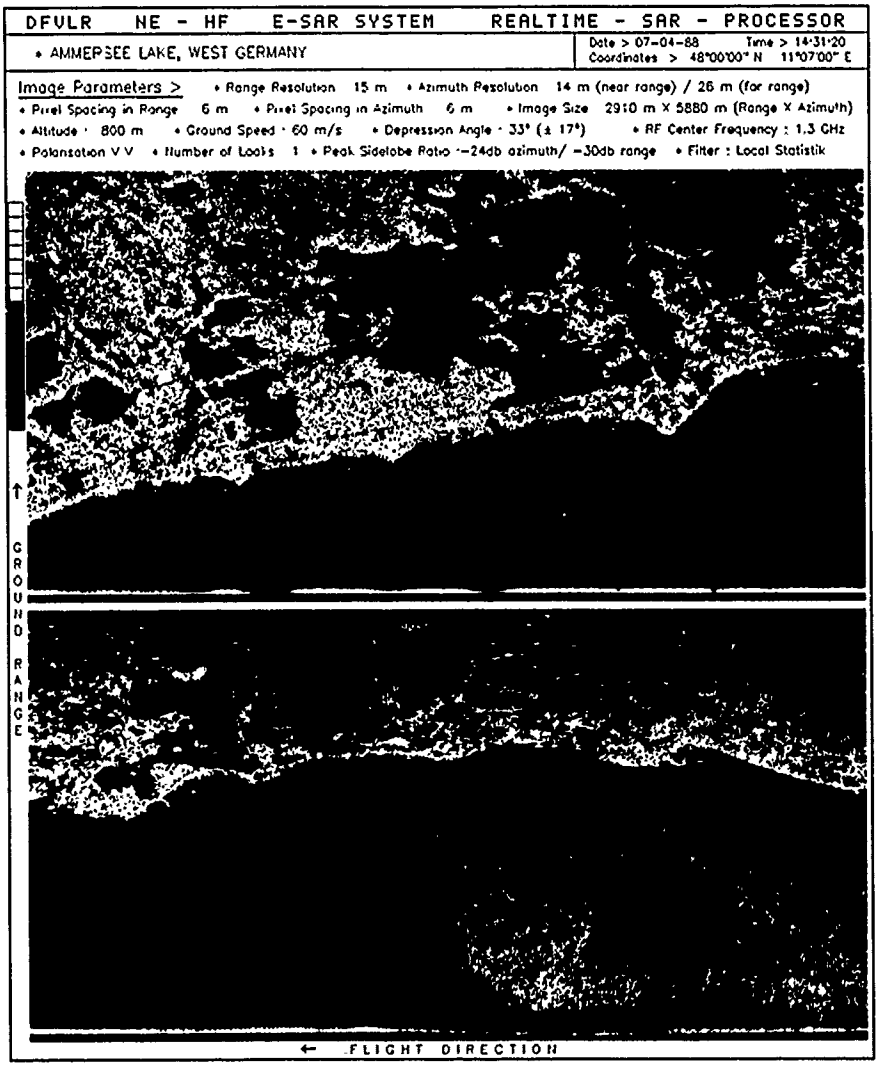


Figure 4: Image processed in realtime: Ammersee lake near Munich. (released by "Regierung von Oberbayern" Nr. GS 300/272/88).

5. REFERENCES

[1] Fitch, J.P.: Synthetic Aperture Radar. Springer, 1988.

[2] Beasley, A.R.: Synthetic Aperture Radar Signal Processing for Airborne Applications, The Marconi Review, vol. XLV, N. 225, Apr. 1982.

[3] Horn, R.: E-SAR - The Experiment Airborne L/C-band SAR System of DLR, Proc. of IGARSS, Edinburgh, Sept. 1988.

[4] Moreira, A.: Konzept und Ergebnisse des DLR Echtzeit-Azimutprozessors für das E-SAR System. DLR Research Report, Apr. 1989.

[5] Moreira, A.; Horn, R.: The Archimedes IIa Experiment on Oil Slick Detection over the North Sea - Measurement Results obtained by the E-SAR System of DLR. DLR Internal Report, Nov. 1988.

[6] Lee, J.S.: Speckle Suppression and Analysis for Synthetic Aperture Radar Images. Optical Engineering, vol. 25, May 1986.

AN INTEGRATED ICE SURVEILLANCE SYSTEM FOR SUPPORT OF OFFSHORE OIL AND GAS EXPLORATION

Philip Rudkin Herbert Ripley Scott Gillis Ken Ludlow

ATLANTIC AIRWAYS LIMITED

Abstract

Over the past four years, Atlantic Airways Limited has developed and is operating an integrated airborne ice surveillance and data management system, designed specifically for the marginal ice zone off Canada's East coast. This system utilizes a Litton 360 degree surveillance radar interfaced with a mini computer. The entre system is mounted in a business class turboprop aircraft. The following paper will describe the various system components and will detail actual results obtained over the past few years in regional and site specific ice management.

Unique conditions, Different approach, Iceberg detection, Operational data

Introduction

Atlantic Airways Limited is a Canadian owned firm based in St. John's, Newfoundland and has been active in the field of remote sensing for several years. Included in these activities is offshore ice surveillance off the Canadian East coast. This area is within the marginal ice zone and is frequented by pack ice from January to late April and year round by icebergs. The main influx is between February and late May.

Visual reconnaissance was severely hampered by fog and the difficulties associated with locating ice not on or near the flight path. While SLAR gave greater area coverage it was hampered by it's operational altitude and its inability to distinguish between ships and icebergs (Rossiter 1985) (Thayer 1985).

It became obvious that the unique conditions found in the marginal ice zone required a totally different approach to ice surveillance. In early 1985 Atlantic Airways concluded that the good points of both radar and visual surveillance should be combined into one multi-mission aircraft capable of providing detailed, accurate and cost effective ice data.

Performance Considerations

1. The radar would need to reliably detect point targets as small as four (4) sq.m the equiivalent to the average RCS of a bergy bit (Klein 1986)
2. The radar would need to be able to detect and map the leading edge of the pack ice, (usually 2-6 tenths of new to first year ice).
3. A suitable method of data collection and data management for passing information to the client.
4. A method of communication for graphic and tabulated data from aircraft to ground.

Equipment Selection

Radar. Several radars were tested under actual operational ice conditions. The radar selected as most suitable was the Litton APS 504 (V)5. This radar is an ASW radar (anti-submarine warefare). It is a 360 degree rotating X band radar with multi processing levels. Radar pulse lengths are user selectable and feature a 500:1 pulse compression. Radar imagery is displayed on a hight resolution 875 line video display which utilizes 16 gray levels for display purposes. Flight tests conducted on a two meter calibrated radar target determined that this unit exceded the manufacture's of a 2 sq.m target at 20 miles in a sea state four, and more than covered the performance specifications in 1 and 2 above.

Data Management. To accomplish this task a PC computer was installed in the aircraft and interface and manipulation software was custom designed by Atlantic Airways staff. The Litton radar was equipped with a custom interface board allowing digital data to be transfered to the PC. This data is then used by additional sub programs to display in both tabulated and graphic form all targets and pertinent flight data. A user interface via multi-level softkeys was designed to allow for a complete inflight data management package.

Communications. Several Options were considered before going with a low tech

solution. A line scanning fax was selected as the means of communication, this fax gave the advantage of being able to transfer data both tabulated and graphic to any ground station equiped with a simple low cost weather fax receiver.

Operational Evaluation

Our operational evaluation comprises three re (3) distinct subcategories:
- glacial ice in open water
- pack ice edge and interior featur
- glacial ice within pack ice

Glacial Ice in Open Water

Iceberg detection is dependent on an iceberg's RCS, and in the case of small icebergs and below is also dependent on sea state. Unlike SLAR which only views the target in its side looking beam pattern for a very short time and from only one angle, surveillance radar will literally get hundreds of hits from numerous angles. This will therefore reduce the effect of losing the target in the sea return. (Rudkin, 1988). This is one of the prime reasons that this type of radar was selected over SLAR or SAR. Operationally, small icebergs and above are normally detected with a high level of confidence within the confines of a 100 mile range. Medium and large icebergs have been detected out to 160 mile ranges. Within a 60 mile radius of the rig the radar is operated in either a 25 or 50 mile range with full processing employed. Priority is placed on detection of bergy bits and growlers, which are typically detect at ranges out to 20 miles.

Sea Ice Edge and Interior Feature

Operational data collection is performed primarly on a 50 to 100 mile range with minimal processing. With 16 gray levels it is possible to clearly see interior features. With experience it is possible to type and classify the different ice type returns by the gray levels.

Glacial Ice Within Sea Ice

From high altitude levels, icebergs that are within the ice pack show returns similar to that of free floating bergs in a heavy sea state. Therefore, medium to large bergs stand out well from the surrounding sea ice returns and present no problem. However, to reliably detect small icebergs it is necessary that CFAR processing be employed to remove the surrounding pack returns. With this processing engaged small icebergs are normally detectable at ranges up to 30 or 40 miles.

Conclusions

Radar

Overall, the system has produced better results than initially expected, especially in the area of sea ice features. With four ice seasons completed it has been proven to be a very reliable system and client satisfaction with the data product has been very favorable. The APS 504(V)5 radar has proven to be an exellent tool for handling the diverse conditions found in the marginal ice zone.

On the negative side, this radar requires considerable operator skill to achieve the best results. The combinations of pulse lengths and various processing, when used correctly, well result in impressive imagery of ice conditions. However, when used incorrectly will result in poor imagery. Despite this fact and some initial development problems, the system would seem to be the best compromise for operations in the marginal ice zone.

Computer

The computer system has proven to have enormous potential and development of additional software is carried out on a continuous basis. The data management package, while providing a product that meets our various clients requirements, has room for major expansion in the area of sea ice. The amount of information that is passed via the radar/computer data line is limited. We are at present testing a new software package that utilizes a video image digitizer to grab complete radar images which can then be manipulated via a limited airborne image analysis package. We are hopeful that this package will be operational by the next ice season.

The system as a whole has proven to be extremely versatile, cost effective and reliable. This fact has be realized by numerous users of ice data, and is born out by the fact that Atlantic Airways Limited is now the largest private supplier of ice surveillance data in Eastern Canada.

Table 1. Typical size and Radar Cross Sections for icebergs.

TYPE	HEIGHT	LENGTH	MASS	RADAR CROSS SECTION	
	METERS	METERS	TONNES	AVE	RANGE
LARGE	46 - 100	120 - 220	10,000,000	180	
MEDIUM	15 - 46	60 - 120	2,000,000	132	49 - 211
SMALL	5 - 15	15 - 60	100,000	17	0.12 - 73
BERGY BIT	1.5 - 5	5 - 15	10,000	4.5	0.01 - 5.1
GROWLER	< 1.5	< - 5	1,000	.07	< - 1.0

Table 2. Iceberg Detection Range vs Sea State.

TYPE	DETECTION RANGE	MAXIMUM SEA STATE
LARGE	200 nm	Unaffected
MEDIUM	175 nm	Unaffected
SMALL	100 nm	Sea State 4 <
BERGY BIT	50 nm	Sea State 3 <
GROWLER	20 nm	Sea State 2 <

NOTE: Growler detection is the maximum under ideal conditiions and is representative of a detectable target. ie. There is some reflective surface.

Bibliography

Rositer, J.R. et al "Assessment of Airborn Imaging Radar for Detection of Icebergs", Environmental studies revolving fund report # 016, 1985

Rudkin, P. et al, "Surveillance Radar a New Tool for Ice Surveillance", Proceedings or The 21st International Symposium of Remote Sensing, 1987

Rudkin, P. et al, " Intergrated Aerial Surveillance System Designed of Ice Management in the Marginal Ice Zone". Proceedings of POLAR TEC 88. Trondheim, Norway 1988

RESULTS OF THE REAL-TIME ADAPTIVE RADIOMETRIC CORRECTION IMPLEMENTED IN THE DFVLR L/C-BAND SAR

Joao R. Moreira & Winfried Poetzsch

German Aerospace Research Establishment (DLR)

Institute for Radio Frequency Technology

8031 Wessling, FRG

ABSTRACT

This paper presents the results of the real time radiometric image correction (RIC) implemented in the DLR-experimental SAR-system. The RIC system is based on a control system, that analyses the analogue-to-digital (A/D)-converted raw data and controls the gain of the IF-section with range. Therewith the quantization and saturation noise introduced by the process of A/D conversion of raw data can be minimized for the whole swath. A short introduction of the RIC system and a report of its performance will be given.

1. INTRODUCTION

In a conventional imaging radar system the backscattered signal is received, downconverted, analogue-to-digital (A/D) converted, formatted and recorded. If this radar system is flown at low altitudes, 1000m for example, the backscattered signal intensity will vary too widely within the range swath, so that a high signal dynamic range of the radar system is necessary.

The dynamic range, normally limited by the A/D-converter, can be expanded by varying the gain of the IF-section, so that the variance of the signal before the A/D-conversion remains constant and is independent of range. The choice for the value of the signal variance for a quantization with a minimum distortion has been studied by (Max,1960) and (Gray et al,1971).

The variation of the IF-section gain compensates the range dependent decrease of the backscattered signal power. This decrease, described by the radar equation, depends mainly on the target range, type and incidence angle and on the antenna pattern. A typical gain curve for this compensation begins with a small value and increases with range.

As this gain curve cannot be exactly determined in advance, due to inaccurate information about the terrain characteristics, a method is used to generate it precisely in real-time. This method is called radiometric image correction (RIC).

2. THE RIC-SYSTEM

Figure 1 shows the implementation of the RIC-system in a Synthetic Aperture Radar (SAR) with I/Q-detection. The backscattered signal is received, downconverted and range compressed. At the IF-section the signal amplitude will be modified by the gain or correction curve. The correction curve is realized by two devices: a high speed and a low speed attenuator.

By a high speed attenuator the fast variation of the correction curve is generated (normally called sensitivity time control or STC).

By the low speed attenuator we can adjust the optimum operating point or attenuation bias for the high speed attenuator. The low speed attenuator is normally called automatic gain control or AGC.

The IF-signal is then demodulated, by the I/Q-detector for example, and A/D-converted. The RIC-system, that controls the AGC- and STC-attenuators, reads the converted raw data and calculates the signal intensity as a function of range over a given integration time. The integration time can be varied from several seconds up to some minutes. Using an adaptive non-linear control system the STC-curve and the AGC-attenuation is then calculated (Moreira,1988).

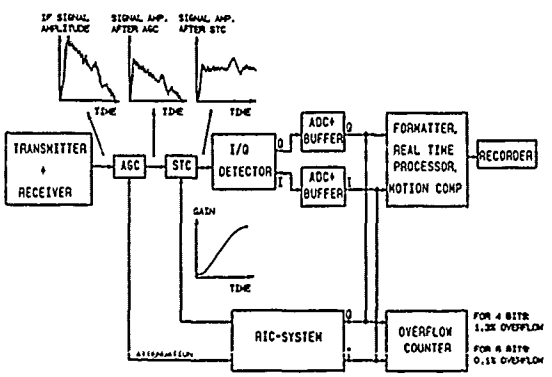


Figure 1. Implementation of the real-time radiometric image correction (RIC)

3. RESULTS

The implementation of the RIC-System has been successfully performed with the experimental C-Band SAR-system during flights in October 1988. The SAR-system uses an I/Q-detection with two 6-bit A/D-converters which each run at 100MHz and have a dynamic range of 25dB at 40MHz or an effective bit number of 4 (Moreira, 1987). As a system dynamic range of more than 40dB was required, a STC-attenuator with aprox. 15dB range was employed. The hardware of the RIC-system consists of an interface to the A/D-converter buffers, a microprocessor system that calculates the STC- and AGC- attenuations, an 8 bit-D/A-converter for controlling the analogue STC-attenuator and a digital 7 bit-attenuator for the AGC-function (Moreira,1986, Spies,1987, Poetzsch,1989).

Figure 2 shows an example of a STC-curve generation. The aircraft (Do-228) flew in the up-down direction and on the left side of the image. Hence the left side of the image corresponds to near range and the right side corresponds to far range.

The upper part of the image is processed without STC, this means with a constant STC curve. In this case the A/D converters have a high saturation rate in near range and produce a high quantization noise in far range. In mid range the A/D converters work with minimum conversion noise. The averaged signal amplitude is shown in the upper curve as a function of slant range.

In the middle of the image the generation of the STC-curve with 7 iterations was started. Here the RIC-system has averaged 32 range lines within 3 seconds for each iteration that corresponds to a strip in the image. The lighter strip represents the initialisation and the next strips correspond to the first, second, ... iterations.

After carrying out the seven STC iterations one obtains the corrected image shown in the bottom part of the image. The lower curve shows that the averaged signal amplitude remains constant with range. The residual signal variance error of less than 1.0dB at the inputs of the A/D converters was also obtained.

As the A/D-converters used have an effective number of 4 bits, we get from (Gray,1971) the following conclusions:

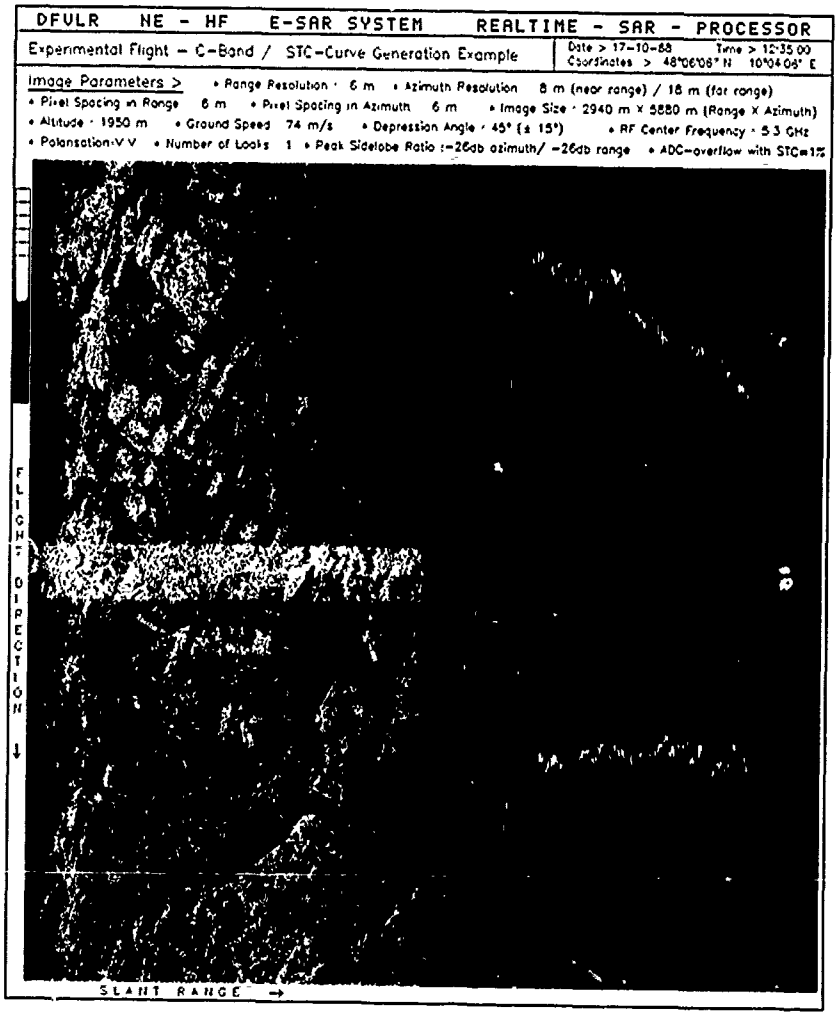


Figure 2. An example of the STC-curve generation

- Optimum K. K is defined as the saturation level of the A/D-converter to the input rms voltage ratio. For a 4 bit A/D-converter we have the minimum A/D-conversion noise at K = 2.5.
- Saturation ratio. With K = 2.5 we obtain a normalized quantization noise power of 1.05E-2, a saturation noise power of 0.24E-2 and a total noise power of 1.29E-2. As the probability density function of the radar signal before the A/D-conversion is a normal distribution, we get for K = 2.5 a saturation ratio of 1.3%.

Figure 3 shows the histogram in far range of the image without STC (upper part of figure 4). The histogram was calculated from 1024 samples of the A/D-converter. The standard deviation is 7.5 and the saturation level of the A/D-converter is 32 considering 64 steps of the A/D-converter. The overflow frequency is 1 sample that corresponds to 0.1%. The ratio K for this case is 4.3. The A/D-converter has a small excursion, works below its optimum operating point and causes a high conversion noise due to high quantization noise.

Figure 4 shows the histogram in near range of the image without STC (upper part of figure 4). The histogram was also calculated from 1024 samples of the A/D-converter. The standard deviation is 23.1 considering 64 steps of the A/D-converter. The overflow frequency is 329 samples that corresponds to 32%. The ratio K for this case is 1.4. The A/D-converter has a large excursion, works over its optimum operating point and causes a high conversion noise due to high saturation noise.

Figure 5 shows the histogram in near range of the image with STC (bottom part of figure 4). The histogram for the whole range swath is practically the same, so that we can take this example as a general one. The histogram was also calculated from 1024 samples of the A/D-converter. The standard deviation is 12.7 considering 64 steps of the A/D-converter. The overflow frequency is 14 samples that corresponds to 1.4%. The ratio K for this case is 2.5. This result matches with (Gray,1971), so that the A/D-converter works at its optimum operating point and causes minimum conversion noise.

4. REFERENCES

[1] Max, J.,"Quantizing for Minimum Distortion",IRE Transactions on Information Theory, vol. PGIT-6, 1960.

[2] Gray, G. et al,"Quantization and Saturation Noise Due to Analogue-to-Digital Conversion".IEEE Transactions on Aerospace and Electronic Systems, Jan. 1971.

[3] Moreira, J.,"Real-Time Radiometric Correction for Imaging Radar Systems".IGARSS Symposium, Edinburgh,23-24,1988

[4] Moreira, J.: Error Analysis of SAR Hardware. European Space Agency, Technical Translation ESA-TT-1024, March 1987.

[5] Moreira, J.: Experimentelles SAR-System; Das STC/AGC-System. DFVLR MEMO, 12.06.1986.

[6] Spies, U.: Hard- und Software fuer das AGC/STC-System des experimentellen DFVLR SAR. DFVLR IB Nr.:551-5/87, 24.04.87.

[7] Poetzsch, W.: Software fuer das AGC/STC-System des experimentellen DFVLR SAR. DFVLR IB in preparation.

5. ACKNOWLEDGEMENTS

The author wishes to thank Mr. U. Spies for the hardware design and Mr. R. Horn for the interfacing with the experimental SAR.

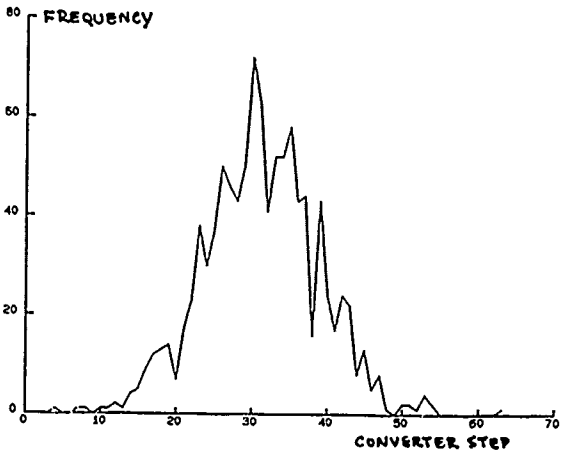


Figure 3. Histogram of the A/D-converted signal without STC in far range.

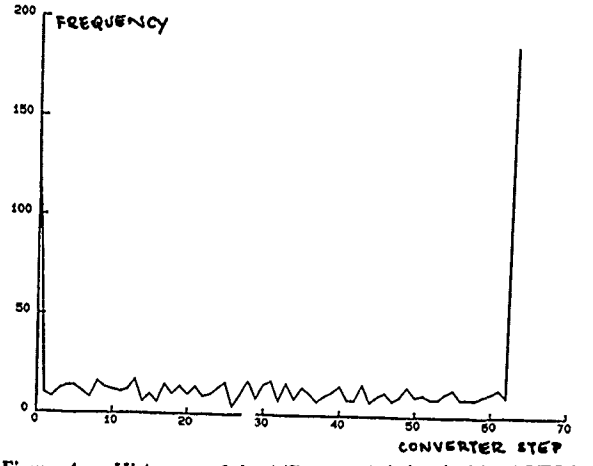


Figure 4. Histogram of the A/D-converted signal without STC in near range.

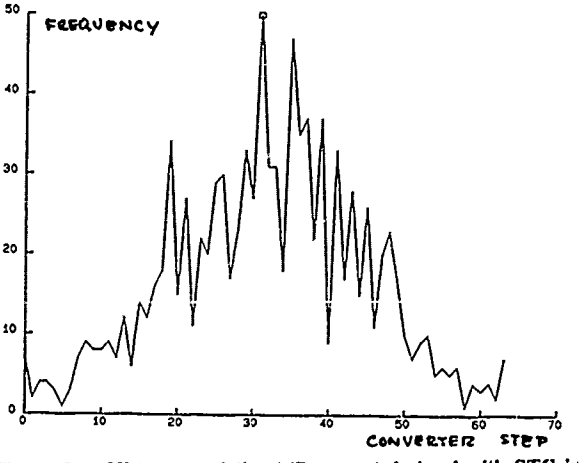


Figure 5. Histogram of the A/D-converted signal with STC in near range.

APPLICATIONS OF THE SIMD MPP TO PROCESSING AND ANALYZING REMOTELY-SENSED IMAGERY

James C. Tilton

Mail Code 636
NASA Goddard Space Flight Center
Greenbelt, MD 20771

ABSTRACT

The Massively Parallel Processor, or MPP, is a Single Instruction, Multiple Data Stream (SIMD) computer containing 16,384 bit serial processing elements (PE's) arranged in a 128-by-128 mesh connected array. It can be programmed in its assembly language (MCL), MPP Pascal, and MPP Forth. While the MPP was originally designed primarily for image processing and pattern recognition algorithms, our experience has shown that it can be used very effectively in applications covering a wide range of general physical and mathematical sciences. This paper concentrates on image processing and analysis applications.

Keywords: Data parallel analysis, Remote sensing, Image Analysis, Image processing.

INTRODUCTION

The Massively Parallel Processor was conceived in 1975 at the NASA Goddard Space Flight Center (GSFC). It was developed and built, under contract, by the Goodyear Aerospace Corporation (now Loral Aerospace). When the MPP was delivered to NASA GSFC by Goodyear in May 1983, it was the world's largest operational massively parallel computing system. The MPP is a Single Instruction, Multiple Data Stream (SIMD) computer containing 16,384 bit serial processing elements (PE's). The PE's are arranged in a 128-by-128 array with connections from each PE to its four nearest neighboring PE's.

Most applications are programmed on the MPP using MPP Pascal (an alternate language is MPP Forth). The current version of MPP Pascal supports most standard Pascal features plus several extensions which were added for programming convenience and to allow easy use of the massively parallel architecture. Programming the MPP in MPP Pascal is often more straightforward than programming a conventional serial computer in standard Pascal, since arrays of data are manipulated as single entities.

The MPP was originally designed primarily for image processing and pattern recognition algorithms that involve local neighborhood operations on image data. However, an MPP working group, formed in 1985 and consisting of researchers in various disciplines scattered throughout the U. S. A., has shown that it can be used very effectively in applications covering a wide range of general physical and mathematical sciences. For example, Earth science applications include not only the processing and analysis of remotely-sensed image data, but also Earth system modeling.

This presentation will concentrate on image processing and analysis applications including synthetic aperture radar (SAR) data processing, stereo image analysis for generating topographic maps, sea ice-floe motion detection from time-varying SAR imagery, contextual classification, spatially constrained clustering, and neural network clustering of image data, and image data compression. It will also briefly discuss future plans for the use of massively parallel computation in remote sensing applications at Goddard.

The following bibliography lists earlier surveys of the application of the MPP to processing and analyzing remotely sensed imagery, along with the proceedings from two symposiums focusing on massively parallel computation on the MPP and other massively parallel machines.

BIBLIOGRAPHY

1. Fischer, J. R. (ed.), "Proc. of the First Symposium on Frontiers of Massively Parallel Scientific Computation", NASA Conference Publication 2478, Greenbelt, MD, Sept. 24-25, 1986.
2. Fischer, J. R. (ed.), "Proc. of the Second Symposium on the Frontiers of Massively Parallel Computation", IEEE Computer Society Pub. # 892, Fairfax, VA, Oct. 10-12, 1988.
3. Halem, M., "Applications of remote image sensing with the NASA Massively Parallel Processor", Phil. Trans. R. Soc. Lond., Vol. A-324, pp. 365-372, 1988.
4. Ramapriyan, H. K. and J. P. Strong, "Applications of array processors in the analysis of remote sensing images", Proc. of the Seventeenth International Symposium on Remote Sensing of Environment, Ann Arbor, MI, May 9-13, 1983.
5. Ramapriyan, H. K., J. P. Strong and J. C. Tilton, "The Massively Parallel Processor - programming and applications", Proc. of the Pecora X Symposium, Fort Collins, CO, Aug. 20-22, 1985.
6. Tilton, J. C. and J. P. Strong, "Analyzing remotely sensed data on the Massively Parallel Processor", Proc. of IGARSS'84 Symposium, Strasbourg, France, pp667-672, Aug. 27-30, 1984.

Applications of MIMD Transputer Arrays to Automated Image Understanding of SPOT

Jan-Peter Muller,
Department of Photogrammetry & Surveying,
University College London,
Gower Street, London WC1E 6BT, UK
ARPANET:jpmuller@cs.ucl.ac.uk

The Alvey MMI-137* project is concerned with the automated derivation of extremely dense and very accurate Digital Elevation Models (DEMs) from multiple SPOT images. The computational complexity of the image matching and space intersection algorithms means that CPU times in excess of a few days are not uncommon for processing full SPOT scenes on serial computers. The algorithms have recently been ported onto a PARSYS™ Supernode array of 32 T800 transputers and results on mapping the algorithms and computational speed-ups will be given.

The routine production of DEMs and associated orthoimages means that a large variety of image understanding techniques can now be applied to the resultant data-sets to attempt to build a future fully automated satellite interpretation system.

Two components of such a MIMD system will be demonstrated, viz. the application of matching to automated change detection in off-nadir SPOT imagery and the application of a Monte Carlo ray-tracing system for the simulation of satellite images. Results will be shown for the actual vs. predicted speed-ups and the difficulties and advantages of using transputers in such a processing system.

*In collaboration with Thorn EMI Central Research Laboratories, Royal Signals and Radar Establishment Malvern, Laser-Scan Laboratories and the Department of Computer Science, University College London.

Applications of MIMD Hypercubes to Imaging Spectrometry Data Analysis

Jerry E. Solomon
Jet Propulsion Laboratory
California Institute of Technology
Pasadena, CA 91109

ABSTRACT

While imaging spectrometry is in the process of revolutionizing the world of remote sensing imagery, users of this data are faced with the problem of processing and analyzing this complex data in a timely fashion. These instruments acquire multispectral image data in hundreds of narrow, contiguous spectral bands covering the visible and short-wave infrared portion of the spectrum. Such high resolution spectral data is, in principal, capable of providing extremely detailed information regarding the surface composition of the earth (or planetary bodies for space exploration applications) through spectral signature analysis. Unlike traditional multispectral imaging systems, which use a few broad spectral bands, imaging spectrometer instruments provide spectral sampling sufficient for direct surface materials identification rather than simple class discrimination. Given the complexity and volume of this data it is clear that there is a strong need for development of efficient analysis algorithms and computing architectures in order to fully realize the tremendous potential of these instruments for the remote sensing science community.

This paper describes some early results in development work aimed at building a prototype high-performance science analysis workstation environment based on relatively small MIMD Hypercube concurrent processors. The current implementation uses an 8-node JPL/Caltech Mark III hypercube tightly coupled to a commercial Sun-4 workstation. A major portion of this effort has been development of a concurrent image processing system which provides interfaces for both the science user and the applications programmer. The Mark III hypercube is a local RAM MIMD machine, and this paper discusses some of the issues involved in utilization of this architecture for imaging spectrometry and image processing in general. In particular, the question of problem decomposition on the hypercube is disussed along with issues relating to the fact that it is operated as an attached processor on a host machine. Finally, results of automatic materials identification and spectral mixing decompostion in imaging spectrometry data are shown which illustrate the power and flexibility of this system for remote sensing science analysis.

FREQUENCY DOMAIN FILTERING OF REMOTELY SENSED SATELLITE DATA USING PARALLEL PROCESSING

B. Oldfield, J.E. Robinson, G.A. Betzos[2].

Geology Department, Syracuse University, Syracuse, NY 13244
[2]School of Computer and Information Science, Syracuse University, Syracuse, NY 13244

ABSTRACT

Physical structures and facies variations in bedrock and sedimentary horizons may not be confined to a specific spectral reflectance domain. These features may be part of one or more reflectance or topographic boundaries that may be revealed only by subtle alterations that outline a continuous physical shape or texture. These anomalous features are best displayed by procedures that emphasize the continuity of shapes and textures rather than the discrimination between spectral reflectance channels. Band pass filters that enhance features of specific size and display them without the distraction of conflicting larger or smaller features are effective for image analysis. A sequence of spatial filters, each designed to retain a relatively small size range of features, can be applied to the image and results displayed for interpretation.

Band pass filters can be applied in the spatial domain by convolution or by its operational transform in the frequency domain using multiplication. Frequency domain filtering, with its added advantage of permitting the recognition of preferred orientations in the frequency spectrum, filtering the features of interest, then converting the image back to the spatial domain using the Inverse Fourier transform, enables the design of filters for image enhancement. Parallel Fast Fourier transform software allows for the display of the Fourier spectrum, the design of any number of narrow band, radially symmetrical zero phase filters, and the calculation and display of the resulting image in the spatial domain.

Filtered Thematic Mapper data over the Patrick Draw region of southwestern Wyoming show the presence of features that were not distinguishable in the original image. Standard enhancement techniques improve the original image but do not cause subtle features to stand out. Band pass filtering has proven very effective in the display and analysis of subtle features that may be masked by unrelated reflectances in the original data. This ability is significant because these subtle features may have economic and environmental importance.

Keywords: remote sensing, frequency domain filtering, parallel processing.

I. INTRODUCTION

A remotely sensed image represents the sum of electromagnetic energy reflected or emitted by the Earth's surface and the atmosphere that is measured and recorded by the sensor. It includes the reflectance from all physical features that make up the scene, ranging from large regional geologic structures to small agricultural areas and man-made structures. Usually the smaller and brighter features dominate the appearance of the image and mask medium sized, and larger features. Spatial filtering, using precisely designed narrow band filters, permits the extraction and display of specific features that are present in the image. Frequency domain filtering, utilizing Fast Fourier Transform software and powerful parallel processors, allows for almost real time design and application of any desired spatial filters and rapid viewing of filtered results. The extracted features can be displayed singularly or added back to the original image so that selected features appear enhanced against the original background.

II. THE FOURIER TRANSFORM

In its original form the image, in the spatial domain, is defined by rows and columns, in which all features are displayed as a function of distance. The frequency domain displays the information in a satellite image in terms of sinusoidal components of varying amplitude, frequency and orientation. Fourier transforms relate the spatial and frequency domains. The information is identical in either domain, however, the frequency function consists of two complex spectra. One spectrum displays the amplitude of the component frequencies, and the other, their phase. Frequency domain filtering is the operational transform of distance domain filtering and normally involves the altering of the amplitudes of selected frequency components. The phase spectra relates to the position of the physical features in the image and is not altered.

As Cooley and Tukey (1965) and Gentleman and Sande, (1966) described, the two-dimensional Fourier transform of a function g(x,y) is:

$$G(u,v) = \int_{-\infty}^{+\infty} \int_{-\infty}^{+\infty} g(x,y) e^{-2\pi i(ux+vy)} \, dxdy$$

The Inverse transform is defined by:

$$g(x,y) = \int_{-\infty}^{+\infty} \int_{-\infty}^{+\infty} G(u,v) e^{2\pi i(xu+yv)} \, dudv$$

The digital implementation of the Fourier transform is:

$$F(m,n) = \frac{1}{MN} \sum_{j=0}^{M-1} \sum_{k=0}^{N-1} f(j,k) e^{-2\pi i\left(\frac{jm}{M} + \frac{kn}{N}\right)}$$

$m = 0,1, \ldots, M-1$
$n = 0,1, \ldots, N-1$

The Inverse Fourier transform is given by:

$$f(j,k) = \frac{1}{MN} \sum_{m=0}^{M-1} \sum_{n=0}^{N-1} F(m,n) e^{2\pi i\left(\frac{jm}{M} + \frac{kn}{N}\right)}$$

$j = 0,1, \ldots, M-1$
$k = 0,1, \ldots, N-1$

The Fast Fourier transform (FFT) has replaced the standard digital Fourier transform because of its computational speed (Cooley and Tukey, 1965; Bergland, 1969). In the FFT the number of operations is proportional to N^2 where N is the number of sample points, whereas the number of operations required for the standard Fourier transform is $N\log_2 N$.

III. FILTERING

Filtering in the spatial domain requires that the filter operator be convolved with the image. Filtering in the frequency domain is the operational transform of convolution and consists of multiplying the amplitude spectrum of the filter by the amplitude component of the image and adding their phases.

The image in the frequency domain is displayed as an amplitude spectrum or a phase in which the components are related to their specific frequencies. The area around the zero origin, the center of the amplitude spectrum, corresponds to the low spatial frequencies. The area near the outside edges of the spectrum relates to the high spatial frequencies, and the intermediate frequencies lie in the intervening area.

The filter consists of frequency components that are to be retained. If the filter has radial symmetry about the zero frequency origin then it has zero phase characteristics and the phase spectra can be ignored. All retained features in the filtered output will be in their correct positions and orientations. If the frequency components in the filter are made equal to one, they will be increased. If the filter components are made equal to zero then the component they represent in the output will be deleted.

A simple but useful category of filters either allows or blocks a band of frequencies from passing through. There are three general types of filters: low pass, high pass and band pass. Large, regional features in the image are made up primarily of low, long wavelength spatial frequencies. Features of intermediate size are composed of medium scale spatial frequencies, while small features dominate the high frequencies.

Filter band widths should be at least one octave to prevent side lobe oscillation or ringing. Because steep boundaries require high frequencies for their definition, the output from low and intermediate features will tend to be rounded but will be correct in terms of shape and position. The maximum frequency, or shortest wavelength that can be measured other than the zero frequency, is the Nyquist frequency with a wavelength of two samples. The minimum, or lowest, that can be detected has a wavelength equal to one-half the number of samples in the transformed image. Filters can be selected anywhere between these two end members.

Where the size of a feature is known it can be extracted by selecting a center frequency of the one octave pass band that has a wavelength of twice the width of the feature. The band pass filter should include at least one-half octave on either side of this center frequency. Where it is necessary to test for a number of features of different sizes the parallel processing system enables the use of a large number of filters ranging from low to high pass; thus the analyst can selectively enhance and display all contained features. Just as selected features can be enhanced on the basis of their size, so can unwanted features be eliminated from the image by the use of notch filters. These filters, which eliminate a selected range of frequencies, can delete regular patterns such as agricultural field patterns to improve the overall image appearance and aid interpretation. Notch filters can be very narrow to eliminate regular patterns such as stripping and not effect any other features in the image.

IV. COMPUTER PROCESSING

A two-dimensional Fast Fourier transform program, running on a Connection Machine, was utilized to process the satellite data. The Connection Machine is an SMID (single-instruction, multiple-data), fine-grained parallel computer consisting of up to 64,000 processors connected to a general communication network (Hillis, 1985; Treleaven, 1988). Many algorithms that process digital images perform computations that can be applied to each pixel simultaneously (Rosenfeld and Kak, 1982). These algorithms are well-suited to implementation on architectures like the Connection Machine. Each pixel in the image is assigned to one processor, and processor communication is kept to a minimum, resulting in a significant speed-up over sequential algorithms that

process one pixel at a time. With the parallelism of the Connection Machine, all the rows in an image are computed at the same time, and then all the columns.

The program computes the FFT of an image band, saves the amplitude spectrum, applies a user-specified circular function to filter desired frequencies, and computes the Inverse FFT to convert the image back to the spatial domain. All stages of processing are viewed on a networked color Sun workstation.

V. EXAMPLE

A 512x512 pixel subset of a Thematic Mapper (TM) band 5, 1.55-1.75µm, of the Patrick Draw area of southwestern Wyoming, containing a major oil field, was selected for analysis. A sequence of narrow band filters was applied to determine if the oil field could be displayed by selective analysis. Figure 1 is the original image, and Figure 2 is the amplitude spectrum. Processing the data on the Connection Machine permitted the almost-immediate display of a sequence of filters and their resulting filtered images. The low pass filtered image illustrated in Figure 3, with its corresponding filter in Figure 4, shows an anomaly that corresponds with the outline of the Patrick Draw oil field. The high frequency image, Figure 5, resulting from the application of the filter illustrated in Figure 6, shows closely spaced detail over the oil field and appears to be related to roads and field development. Drainage features are more prominent. The application of a band pass filter, Figure 7, designed to remove the zero frequency and high frequencies, enhances the outline of the oil field (Figure 8). The edges are not as well-defined as in the low pass filtered image however: the true margin of the field is not accurately known.

VI. CONCLUSION

Frequency domain filtering using parallel processing is a fast, effective method of image analysis that allows for the extraction, removal and enhancement of features in almost real-time mode.

ACKNOWLEDGEMENTS

Thanks to the Advanced Graphics Research Lab and the Northeast Parallel Architectures Center for providing research computing facilities. ARCO Oil and Gas Company contributed the data used in this project.

REFERENCES

1. Cooley, J.W., and J.W. Tukey, "An algorithm for the machine calculation of complex Fourier series", Math. Comput., Vol. 19, pp297-301, 1965.
2. Gentleman, W.M., and G. Sande, "Fast Fourier transforms for fun and profit", Fall Joint Computer Conference, AFIPS Proc., Vol.29, pp563-578, 1966.
3. Bergland, G.D., "A guided tour of the Fast Fourier transform", IEEE Spectrum, Vol. 6, No. 7, pp41-52, 1969.
4. Hillis, W.D., *The Connection Machine*, MIT Press, Cambridge, MA, 1985.
5. Treleaven, P.C., "Parallel architectures overview", Parallel Computing, Vol. 8, pp59-70, 1988.
6. Rosenfeld, A., and A.C. Kak, *Digital Picture Processing*, Vol. 1, Academic Press, New York, 1982.

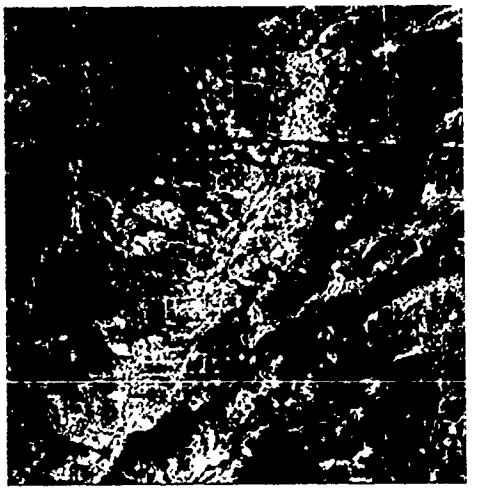


Figure 1: Original TM Band 5 image of the Patrick Draw area of southwestern Wyoming, Patrick Draw oil field trends southwest-northeast.

Figure 2: Amplitude spectrum of TM band 5.

Figure 3: Low pass filtered image.

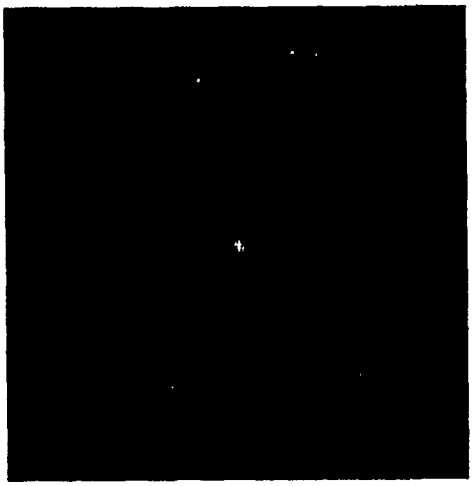


Figure 4: Low pass filter.

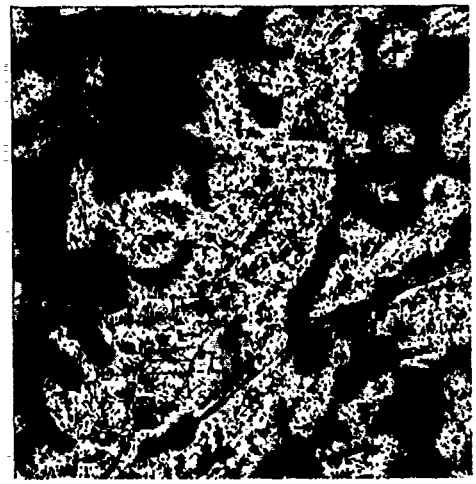


Figure 5: High pass filtered image.

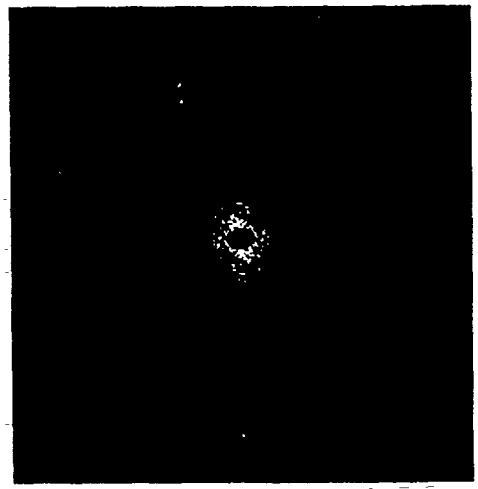


Figure 6: High pass filter.

Figure 7: Band pass filtered image.

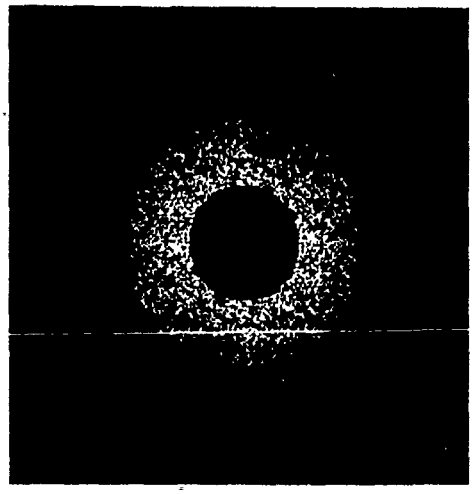


Figure 8: Band pass filter.

DIGITAL ENHANCEMENT OF STAR-1 SAR IMAGERY FOR LINEAR FEATURE EXTRACTION

E.R. Sudibjo, G.D. Lodwick and S.H. Paine

Department of Surveying Engineering
The University of Calgary
Calgary, Alberta T2N 1N4
Telephone : (403) 220 - 5826
Facsimile : (403) 284 - 3697

ABSTRACT

Edges can be defined as the transition zones between areas in an image, which have different properties. The problem of edge detection is of great interest in image processing with many different techniques being proposed, most of which are based on differencing or gradients. This study seeks to assess the feasibility of using airborne Synthetic Aperture Radar (SAR) images in linear feature detection and identification. The first stage involves implementation and assessment of speckle reduction using five spatial filters. The next stage of the study considers the implementation of statistical texture analysis procedures prior to the application of linear feature extraction techniques. The last step of this research attempts to find an optimal segmentation technique which utilizes texture for linear feature detection. The data used are from an image of the Brazeau Range area NW of Calgary, Alberta collected with the Intera Technologies Ltd STAR-1 SAR system. The overall objective is to develop techniques that will be relevant to solution of the mapping problems faced by tropical areas, due to cloud cover.

INTRODUCTION

SAR images contain a large number of different features, both natural and man-made. For the intensive use of the images, it is necessary to detect a wide range of clearly identified objects. The digital analysis involves spatial information, with the specific process dependent on the type of imagery and target objects. Radar imagery is often disturbed by noise (speckle) and careful elimination of the speckle is necessary for reliable object detection. Reduction of speckle in SAR imagery needs to be performed in such a way that the objects to be detected are retained or enhanced while the random nc e is removed or reduced. The speckle in radar imagery requires flexible techniques as different applications and target objects will affect the choice procedures. For example, the choice of the type and spatial extent of the speckle reduction filter is dependent on the type and size of the linear feature of interest.

In passive sensor images, such as Landsat and SPOT, brightness differences can be attributed to variations in the target reflectivity. However, in a radar scene, the brightness differences are due, not only to the variation in reflectivity, but also are the result of differential grazing angles, which affect the scattering coefficient. The backscatter from a radar scene may change by several orders of magnitude with relatively minor variations in the viewing angle. In radar imagery the electromagnetic interaction with the targets is also modified by the object's dielectric constant. Thus a global classification of targets based on spectral properties is not a reliable technique.

The spatial variability in scattering properties from targets gives them an intrinsic texture which can be represented as a texture random variable (Ulaby et al., 1986). Thus, it is expected that valuable information can be determined by a measure of textural variance within a radar scene. An auto-correlation function can be used to measure the spatial variability (texture) in an image, and to discriminate different land cover types. The intent of this study is to optimize this texture analysis process to segment for linear feature detection.

TEXTURE

Texture is an important aspect in image analysis as it involves a measure of both the spectral and spatial variation in the scene. The primitive of the image texture is a collection of pixels which share a common property and are geometrically connected. These pixels are related to the texture with a relationship which may be structural, probabilistic or both. According to Haralick (1979), there have been several statistical approaches to the measurement and characterization of image texture. Some of these are spatial gray tone co-occurrence probabilities, gray tone run lengths, and textural edgeness. The properties of image texture can be described as fineness, coarseness, randomness and regularity, which can be translated to tonal primitives and relations between tonal variations.

Textural information in an image is contained in the spatial relationship of the gray levels and is therefore a

function of distance and direction of the relations (Pultz and Brown, 1987). However, in this study, these spatial gray tone co-occurrences will be used to detect linear features in the imagery based on their implicit directional information. Co-occurrence is a modelling of the texture based on the distribution of the gray tones in an area, using a frequency of occurrence in a specific spatial relation, which has both directional and distance parameters.

FEATURE ANALYSIS

The techniques discussed in the previous sections are implemented to quantify the spatial relationships between pixels in the imagery to either reduce or enhance certain effects based mainly on the texture. The research in progress is using the STAR-1 data of the Brazeau area NW of Calgary, Alberta. The data were resampled to a pixel size of 6m. The speckle reduction involves evaluation of five different spatial filters. The filters, mean, median, mode, nearest values and minimum variance, are applied to the radar imagery with a fixed kernel size. These are low pass filters that remove the high frequency component, which is assumed to be noise, by smoothing out isolated points of high variance (Paine, 1987).

First Order Statistic

The first phase of the research is to analyze the imagery using a first order statistic, which is simply concerned with the frequency of the gray levels in the scene. The first order statistic can be described by the distribution of the gray levels for a measure of the domain and range of a fixed area. Examples of these are the calculation of the mean for location, the standard deviation to measure the dispersion, the skewness to describe the asymmetry, and the kurtosis to represent the shape of the distribution. These statistics provide a more robust measure for feature identification, as they involve local distributions rather than simple absolute values.

It is possible to use these measures for relative comparisons within each local area. An example of this would be the measure of difference between the actual central pixel value and the expected value based on the measured distribution. However, if the kernel used to calculate the statistics covers more than one target then these measures are unreliable. To avoid the multitarget problem, it is necessary to analyze the relationship between the target pixel and its neighbours using a second order measure.

Co-ocurrence: Spatial Gray Level Dependence

Co-occurrence measures are an estimate of the frequency of gray tones relationship, which is a specific spatial relation with distance and directional parameters. This method computes the second order statistic based on the probability of a pixel with value i having a spatially related pixel with value j (Dube et al., 1986). In this study the directional parameter used to calculate the co-occurrence matrix was specified for the four directions of 0° (horizontal), 45° (right-diagonal), 90° (vertical) and 135° (left-diagonal). The distance parameter was defined as one pixel (adjacent pixel), two pixels and four pixels. For example in Figure 1a, the value 20 for the element in the (2,2)th position of the co-occurrence matrix, is calculated from Figure 1b as the total number of times gray tones of value 2 and 2 occurred with one pixel and two pixel distances for the four directions.

(a)

GRAY TONE	0	1	2
0	0	0	0
1	0	12	12
2	0	12	20

(b)

1	2	1
2	2	2
1	2	1

Figure 1. (a) Co-occurence matrix calculation from (b) 3 x 3 window

Windows of 5 x 5 to 7 x 7 pixel kernel are used to derive the co-occurrence matrix. The kernel size was chosen so that there would be a meaningful distribution of gray tones from each application. To process the co-occurrence matrices in the real image, which has very many gray levels, is not always possible. It is often necessary to reduce this range of gray levels by a transformation to a uniform distribution with fewer values. This was achieved with the use of histogram equalization techniques.

From the co-occurrence matrix, a variety of measures can be employed to extract textural information. The study is utilizing four measures of texture selected from Haralick (1979) who described seven different measures that can be used for this kind of work. The four chosen are :

1) Contrast : $\sum_i \sum_j (i - j)^2 \, p(i,j)$

2) Uniformity: $\sum_i \sum_j [p(i, j)]^2$

3) Correlation: $\sum_i \sum_j \frac{(i - u_i)(j - u_j) \, p(i,j)}{\sigma_i \, \sigma_j}$

4) Entropy : $\sum_i \sum_j p(i,j) \log p(i,j)$

where :

p(i,j): (i,j) of the co-occurrence matrix
i: gray level at row
j: gray level at column
u: mean value
σ: standard deviation

Contrast measures the amount of variation in the image segment, uniformity is a measure of homogeneity, correlation is a measure of gray tone linear dependencies and entropy is a measure of the complexity.

The co-occurrence matrix is used for segmentation according to texture measures in the image. If the co-occurrence matrices of a region (image segment) are similar, then the region is considered to be "uniform" in texture and may be classed as a single target. After a number of "uniform" regions have been identified, then the parameters (attributes) of the region can be used for further segmentation as described in the process diagramed in Figure 2.

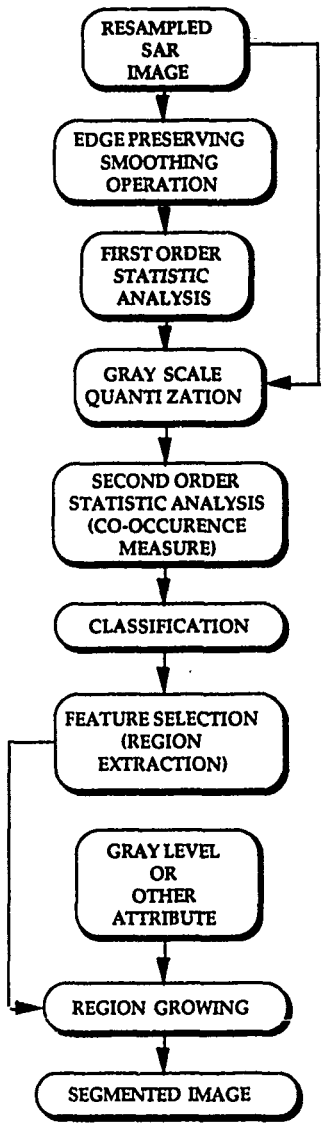


Figure 2. Segmentation processing flow chart

RESULTS

Five edge preserving filters were used to reduce the speckle effect of a single data set. The data used in this study were collected in April 1986. The image covers approximate 6.5 x 6.5 km with a 6 x 6 m pixel (see Figure 3). This segment of the SAR image comprises an area of the Rocky Mountains, which contains relatively homogeneous surface conditions. The image data used were digitally preprocessed by Intera Technologies Ltd, Calgary, Alberta and resampled and rectified by Vecxel Corporation, Boulder, Colorado.

A good edge preserving filter should maintain edges and textures. Therefore, it is expected that after filtering the images will have more contrast compared to the original. Therefore, for linear feature extraction, the edges in the filtered images should be true in position, but more apparent and distinct than their surroundings. This is necessary for the next steps such as enhancement and segmentation. To assess this, the filtered images were evaluated by visual interpretation, by comparing the original and filtered images. In this test, in fact, all five filters smoothed out speckle but overall the linear features are not distinct enough to be recognized digitally.

The mean filtered image is the result of averaging the nine pixel values in the kernel. The averaged value was assigned to the central pixel position in the new filtered data set. The image is smoothed, however, this filter moved the position of linear features. According to Paine and Mepham (1986), this filter causes several classes of similar values to merge, and thus changes the shape of areas.

The median filter (Figure 4) maintained the edges and the positions. This filter basically assigns only the middle value of the range, from the minimum and maximum values in the kernel, to the central pixel. Therefore well defined linear features in the original are maintained, and weak linear features in the original are faded. This filter gave a good result.

The mode filter (Figure 5) eliminated some of the linear features. A nearest values filtered image, also reduced the linear features, similar to the mode filter. The last filter, minimum variance, was used with a 3 x 3 box and 1 x 3 bars. Both methods, box and bars, gave poor results in maintaining the linear features. Overall, of the five filters used, the median filter gave the best result.

CONCLUSIONS

The aim of this study is to assess the feasibility of using SAR images in linear feature detection without the help of other remote sensing images or land register. This paper reports on the first part of this work which is the implementation of the first order statistic. This involves filtering, which is performed to enhance the visual appearance. In this test the evaluation of the filtered images was done visually. However, in reality, a visual assessment of quality is a subjective measure. Therefore

this criterion can only be applied for evident distortions. A more quantitative comparison could be determined by computing the ratio between the mean values and standard deviations for all the filtered images, i.e. the signal-to-noise ratio.

Each of the five filters applied gave a unique result which can lead to the selection of a particular type of filter to obtain a certain objective. However, from the results of this filtering, the linear features were not statisfactorily enhanced to be recognized automatically (digitally). This shows that the five filters used in this work are not appropriate for the linear feature extraction using radar imagery. The reason is that, in these filters, the noise is assumed to have a normal distribution, whereas the noise (speckle) in radar imagery is uniform. Therefore linear feature detection for radar imagery should use another approach. In this study future work will involve the spatial gray tone co-occurrence to detect linear features based on their implicit directional information.

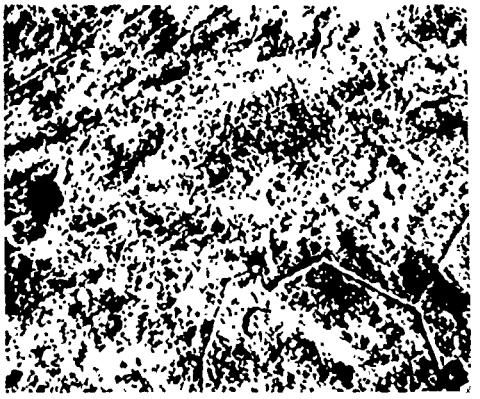


Figure 3. Original SAR image

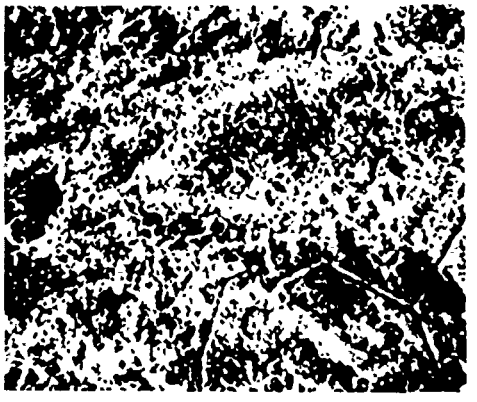


Figure 4. Median filtered image

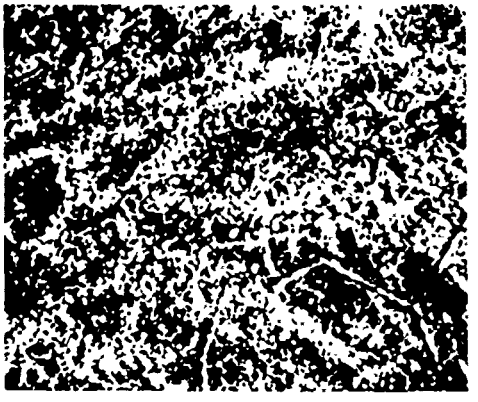


Figure 5. Mode filtered image

REFERENCES

Dube, C., X.H. Proul and K.P.B. Thompson (1986) Analysis of Spatial Structure of Synthetic Aperture Radar (SAR) Imagery For Better Separability of Cereal Crops, Wheat and Barley, Proceedings of IGARSS '86 Symposium, Zurich, Switzerland, September 1986, Ref. ESA sp-254, pp. 745 - 750.

Haralick, R.M. (1979) Statistical and Structural Approaches to Texture, Proceedings of IEEE, Vol. 67, No. 5, pp. 786-804.

Paine, S.H. (1987) Information Extraction from Digital Landsat Imagery for Integration into an LRIS, Ph.D. Thesis, Department of Surveying Engineering, The University of Calgary, 208 pp.

Paine, S.H. and M.P. Mepham (1986) Spatial Filtering of Digital Landsat Data for the Extraction of Mapping Information, Proceedings of Tenth Canadian Symposium on Remote Sensing, Edmonton, Alberta, Canada, May 1986, pp. 27-39.

Pultz, T.J. and R.J. Brown (1987) SAR Image Classification of Agricultural Targets using First and Second Order Statistics, Canadian Journal of Remote Sensing, Vol. 13, No. 2, pp. 85-91.

Ulaby, F.T., F. Kouyate, B. Brisco and T.H.L. Williams (1986) Textural Information in SAR Images, IEEE Transaction on Geoscience and Remote Sensing, Vol. GE.24, No. 2, pp. 235-245.

A TECHNIQUE FOR ROAD DETECTION FROM HIGH RESOLUTION SATELLITE IMAGES

L.LALITHA
NNRMS-ISRO
ANTARIKSH BHAVAN
NEW BEL ROAD
BANGALORE 560034
INDIA

ABSTRACT

Available enhancement techniques are applied to a Landsat TM urban scene to ascertain which of the techniques is effective in improving the contrast of the road features in the image. Templates designed to respond to linear features in an image are applied to detect roads from a satellite image.

KEYWORDS Linear feature, image enhancement, band ratio.

1. INTRODUCTION

Extraction of road networks from high resolution satellite images of Landsat thematic mapper (TM),SPOT and LISS II of the Indian Remote Sensing satellite is a challenging task in digital image processing. Specially in developing countries where infrastructure development is an ongoing process, satellite images are useful to monitor and assess the need for additional roads as well as design the road network. Road materials generally have a high reflectance in the visible bands (Bajcsy and Tavakoli, 1976; Raja Rao and Mahabala, 1983). Although the width of the roads is less than the spatial resolution of the sensor, the feature is picked up whenever its spectral contrast is significant with respect to its background. A road may be considered as a linear feature wherein the grey value distributions on both sides of the road are similar. Thus the problem of detection of roads from satellite images can be posed as a problem of detecting linear features from these images.

Fischler et.al.(1981) used a Duda-road operator together with edge detection operators to detect roads in an aerial imagery. Recently Ton et.al.(1987)proposed a three stage approach to detect and label roads seen in Band3 of Landsat TM image. Wang and Newkirk(1988) demonstarted that a k-means clustering algorithm implemented on the Laplace transformed images of bands 1,2,3 from Landsat TM can extract the highway network from a subscene. In this paper we will discuss the effect of different enhancement techniques in sharpening the roads in a satellite image. A methodology to detect the road pixels is then described. This is illustrated for a small part of the subscene.

2.ENHANCEMENT OF LINEAR FEATURES

Linear features from an image can be extracted using local methods or global methods. However, due to the lack of contrast between the road pixels and its background, enhancement of the raw data becomes necessary prior to the application of local methods. The type of enhancement to be applied will depend on the background in which the road is observed..And yet,even with similar background, it is difficult to decide which enhancement technique should be applied.

A subscene of 400 X 400 pixels from Landsat image containing an urban setup is used for illustration. It was noted that the visual clarity of roads was the best in bands 3,4,5 of the data. A false colour composite (FCC)of these three bands resulted in an image (Fig.1) with enhanced road features. This is found to be in good agreement with the observation made by Hord(1986) that the conventional FCC obtained with bands 2,3,4 does not improve the overall contrast of the image. A principal component analysis of bands 2,3,4,5, and 7 was then carried out. It was observed that linear features were not enhanced in any of the component images. A FCC of the first three principal component images (Fig.2) however, resulted in an image with an improved contrast for the urban features. Although edge detectors have been used to detect linear features, it was found that edge enhancement techniques such as the Robert's operator did not in any way enhance the linear feature in the image under study. Of the various band ratios tried, i. was found that the ratio image generated using bands 1 and 4 significantly highlihted the road(Fig.3). It must be emphasised that this particular band ratio may not enhance the road features in all the urban scenes.

3.METHODOLOGY FOR ROAD DETECTION

Local methods of linear feature extraction involve detection of pixels that belong to linear features and then building up the linear features. Detection of line pixels is done by operating a series of templates that are expected to respond to linear features. Usually the templates operate on the first order neighbours of a pixel. We have designed templates that operate on the second order neighbours of a pixel, since it is expected that the contrast between a line pixel and its first order neighbours will be rather low. One pair of templates (Fig.4(a),(b)) operate in the first stage and the maximum of the responses is thresholded to pick up the candidate line pixels. If the maximum response is in the horizontal or vertical direction, a second pair of templates (Fig.4(c),(d)) is operated on these candidate pixels and maximum of the responses is thresholded to confirm whether the pixel is a line pixel or not. If the maximum response in the first stage is in the diagonal direction, then the pair of templates shown in Fig.4(e),(f), are operated on the candidate pixels in the second stage. The threshold setting is guided by the modes of the image grey value histogram and can be set apriori. These templates can detect both dark lines in a light background as well as light lines in a dark background. This characteristic is particularly useful to detect roads from enhanced satellite images.

4.RESULTS

In Fig.5(a) is shown a line printer output of a portion of the urban scene of Landsat TM (Band4/Band1) showing a road. The proposed line detector was applied to this scene. The image grey value histogram suggested a threshold value of 38. However, it was found that with this value considerabe non-road pixels were picked up. By systematically increasing the threshold value it was found that for threshold at 45, most of the road pixels could be identified (fig.5(b)). Once the road pixels are identified any of the linking procedures could be employed to build up the roads.

5.CONCLUSION

Road detection from satellite images require considerable pre-processing efforts to obtain a high contrast image. It is not possible to decide apriori the best technique for the enhancement of linear features. A FCC of bands 3,4, and5 provide the best visual clarity of the road network. For the subscene under investigation it was found that a ratioed image (Band4/Band1) of Landsat TM data gave the best results.

A simple technique to detect linear features from satellite images has been applied to detect road pixels in a subscene. The advantage of the technique lies in its ability to detect roads irrespective of its background (dark or light), ease of implementation, and a guideline for threshold setting.

REFERENCES

1.Fischler M.A, J.M.Tenenbaum and H.C.Wolf, Detection of roads and linear structres in low resoltion aerial imagery using a multisource knowledge integration technique, Comp. Grph. Image. Proc., V15, 201-223, 1981.

2.Hord R.M, Remote sensing methods and applications, John Wiley & sons, New york, 1986

3.Raja Rao T and H.N.Mahabala, Extraction of road maps of townships from aerial imagery, Proc. IEEE Conf. on Sys. Man and Cybernetics, V2,1225-1229, 1983.

4.Ton J. A.K.Jain, W.R.Enslin and W.D.Hudson, Automatic road detection on Landsat 4 TM images, Proc. 21st Symposium on Rem. Sens. Env., 26-30, 1987.

5.Wang F and R.Newkirk, A knowledge based system for highway network extraction, IEEE Trans. Geosci. Rem. Sen., V26, 525-530, 1988.

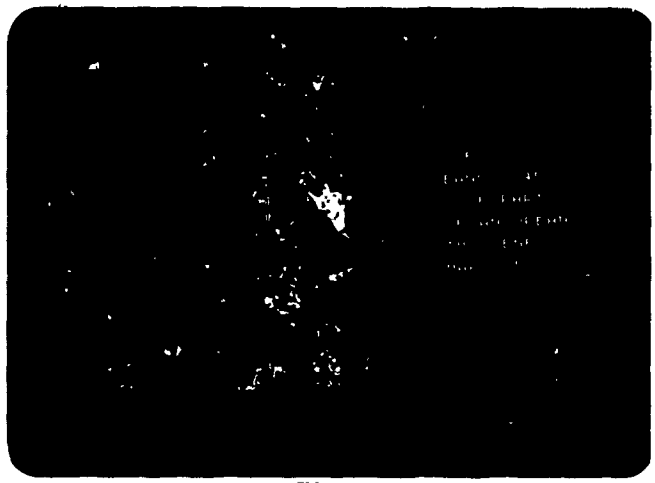


Fig 1

Fig 2

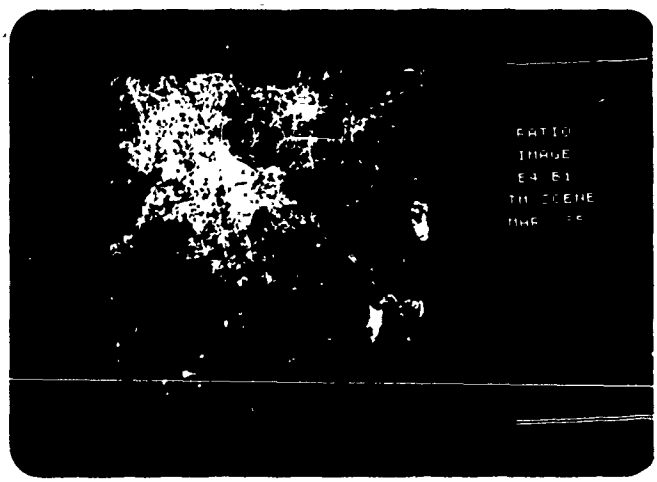


Fig 3

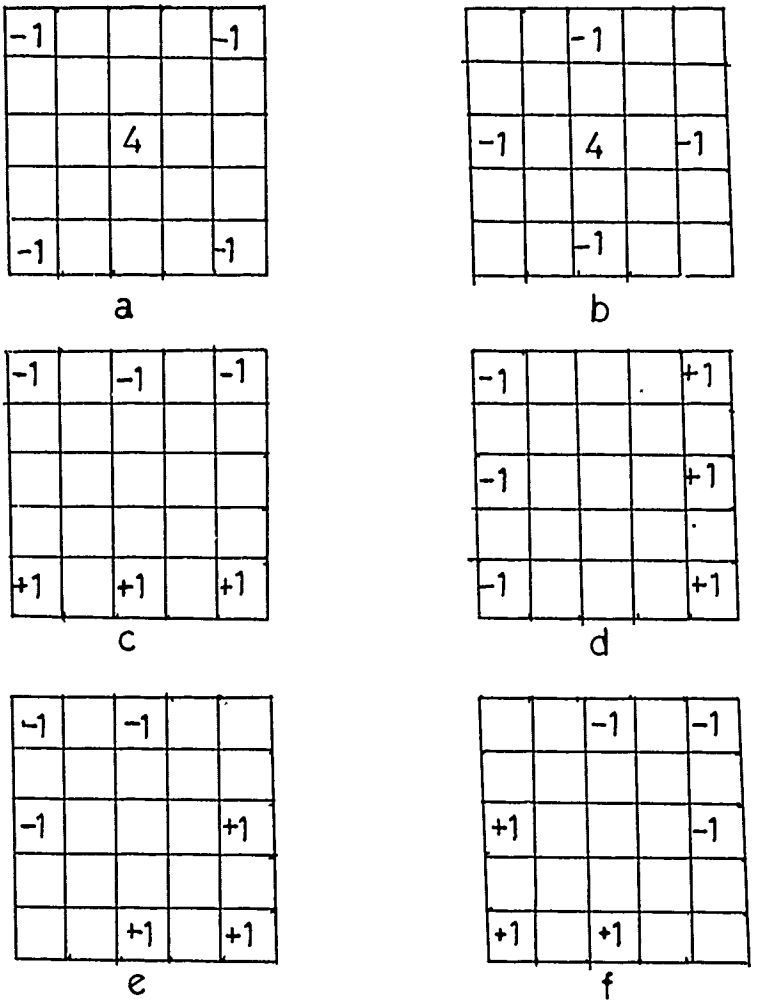


Fig 4 TEMPLATES FOR LINE DETECTION

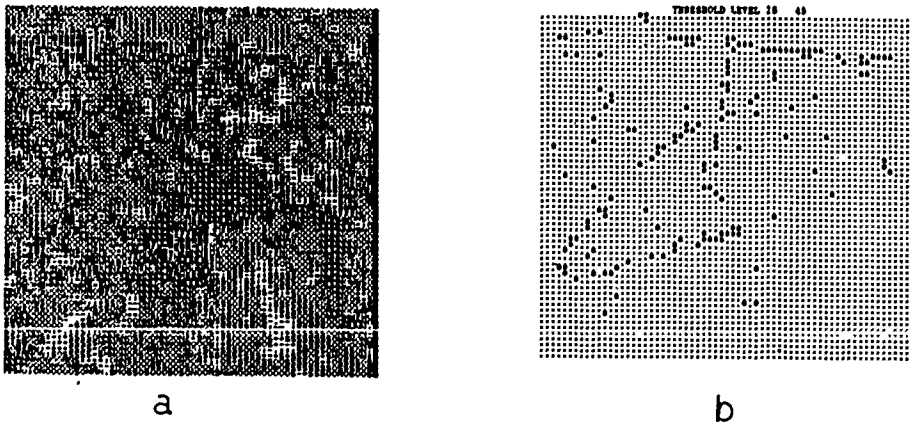


Fig 5 A SUBSCENE AND THE DETECTED ROAD PIXELS

COMPUTER-AIDED LINEAR PLANIMETRIC FEATURE EXTRACTION

John S. Zelek

MacDonald Dettwiler & Associates Ltd.
13800 Commerce Parkway, Richmond, B.C.
Canada V6V 2J3

Abstract

Shape and spectral information can be used as cues to identify features of interest in digital satellite and aerial imagery. This paper discusses and shows how shape structure primitives can be extracted. Shape structure primitives can be selected based on their geometric and spectral properties and spatial location. A toolbox of operations allows the user to identify the locations of domain independent primitives which have certain characteristics. The photointerpreter makes use of global, contextural and textural information to classify the feature type.

Keywords: Remote Sensing, SPOT, Image Understanding, Geographic Information.

1 Introduction

Linear planimetric features include such geographic features as roads, shorelines and streams. Rectangular features (ie. roads) can be defined by their major center axis and treated as linears Humans are very good in identifying the feature of interest, but digitizing is a very tedious and time-consuming operation. Technology in the area of computer vision has not advanced enough to produce a robust automated feature extraction system, but research in this area has showed how various low level operators can assist the photo-interpreter in recognizing the features of interest. The human element is still required in the process of image interpretation, however photointerpretation could be made faster if the computer provided certain tools that assist the user in the task of recognizing, identifying and extracting planimetric features.

Photointerpretation makes use of structural and spectral information in the imagery. The photo-interpreter interprets the input imagery information by utilizing domain specific knowledge (pertaining to a particular geographic feature) such as context, texture, spectral intensity and structure. The process of photo-interpretation is an iterative cyclic process involving various viewpoints. global, local, domain independent, domain dependent, contextural and textural. Useful *domain independent* shape information can be automatically extracted from the imagery.

To speed up the process of photo-interpretation what is proposed is to provide the user with a set of tools (a toolkit) that would assist in the feature extraction process. The tools would extract domain independent features (primitives). Primitives are primarily defined by their shape and secondarily defined by their spectral characteristics. The mapping of a primitive onto a geographic feature is done by the user, once the user has enough accumulated evidence to make that decision. The extraction of geometric structures is based upon the extraction of edges in the imagery.

2 Extraction of Shape Information

2.1 Edge Detection

An edge is a point which indicates the presence of an intensity change in a certain direction. Sub-pixel precision edges in the imagery are detected by convolving one band of imagery with a Difference of Gaussian (DOG) filter and the identification of zero-crossings [2]. The user selects the appropriate band of imagery and the size of the filter. The appropriate image band is determined by what features are desired and how recognizable they are in that image band. The size of the gaussian filter used depends on the resolution of the imagery and the size of the desired geographic features.

A collection of consecutive collinear edges are linked and approximated by a set of straight line segments (polyline). Such lines have radiometric statistics computed for the left and right sides to give relative (ie. bright or dark) or absolute measures. Straight line segments can be paired into *antiparallel line pairs* (apars). An apar is a pair of elongated parallel lines which have opposite contrast [7].

2.2 Apar Primitives

Apars are detected by pairing up straight line segments which approximate the edge boundary. Straight lines are paired based on constraints on the following properties. line parallelism, line pair geometry and radiometric contrast [8] [7].

Lines are perceptually grouped together by applying certain geometric and radiometric constraints. Neighbouring parallel

lines are paired together into narrow linears and can represent an areal feature such as a road or building. Narrow linears are radiometrically classified as being either bright, dark or undetermined with respect to their immediate neighbourhood. Other possible geometric perceptual groupings of linears include collinearity and perpendicularity (corners). Each narrow linear (and other perceptually grouped lines) has a set of attributes (geometrical and radiometric) associated with it.

2.3 Apar Grouping Primitives

Apars provide local information. To make inferences with regards to the type of feature, it may be desirable to obtain global information. Certain disjoint apars may belong to the same feature type. Apars can be disjoint because:

- The edge detection process may have failed to detect the relevant feature boundary,
- The linearization of the edge boundary may have caused a break between two apars, or
- There are no edges in the gap region. The feature of interest has the same reflectance as the neighbouring region and the local information is not sufficient to establish a boundary.

Certain constraints are applied to group apars that are of the same feature type, namely:

- Proximity constraint. The closer a pair of apars are, the more likely they are of the same feature type,
- Gap radiometric similarity. The radiometric statistics of the region between two apars is approximately the same as the radiometric statistics of the two apars.
- Neighbouring apar radiometric similarity. The radiometric statistics of the two apars are approximately similar.

3 Shape and Radiometric Classification

The extracted edges present too much information to be useful as a guide for digitizing features. For the edge information to be useful, the edges have to be pruned into a useful subset. The user has control over the parameters that define the constraints in the edge detection and perceptual grouping processes. The edges can be grouped into geometric structures (i.e, apars). Associated with the edges and geometric structures is a set of attribute information such as length, width and radiometric statistics. The set of geometric structures can be pruned into a manageable working subset by selection based on suitable attribute information. If roads are the feature of interest in SPOT PLA imagery, then bright long apars are selected. Short and bright apars are selected if the feature of interest are buildings in SPOT PLA imagery. A binary mask can be used to select geometric structures within a defined spatial area. Flexibility within the toolbox of operations is provided by allowing the user to interactively constrain what to observe and use for digitizing based on a certain primitive set (i.e, apars).

The following is an example of using the toolbox for locating road features:

- The user selects to display the centerlines of the extracted apars. Only the bright centerlines are chosen. A water region mask can be used to prune apars in a predefined region (Roads are not located in water). If the user knows the radiometric distribution of the roads, then this information can be used to further select which roads to keep. By interactively constraining what to view, the user reduces the set of apars to a set that are highly likely road segments. Typically it was found that the extracted bright apars represented 80 % of the road features in a typical representative SPOT PLA image.

4 Results

Figure 1 is a digitized aerial photograph at 0.75 m resolution of downtown Vancouver. The extracted edges of the Vancouver aerial image are illustrated as a graphics overlay in figure 2. At this type of resolution it is possible to identify such features as cars and ships. The extracted edges outline the boundary of such features. The shape and spectral statistics of the primitive features can be used as constraints for extraction of the desired real-world features.

Figure 1. Digitized Aerial Photograph of Downtown Vancouver

Figure 2: Extracted Edges From Digitized Aerial Photograph of Downtown Vancouver

Figure 3 is a SPOT PLA image of the Madelaine Islands, Quebec with a graphics overlay. The graphics overlay illustrates the extracted apars in the imagery. The apars in the water region (left portion of the image) were pruned by using a binary mask obtained by classifying water and land. The roads and shoreline can be separated by looking at the radiometric statistics of the relevant apars.

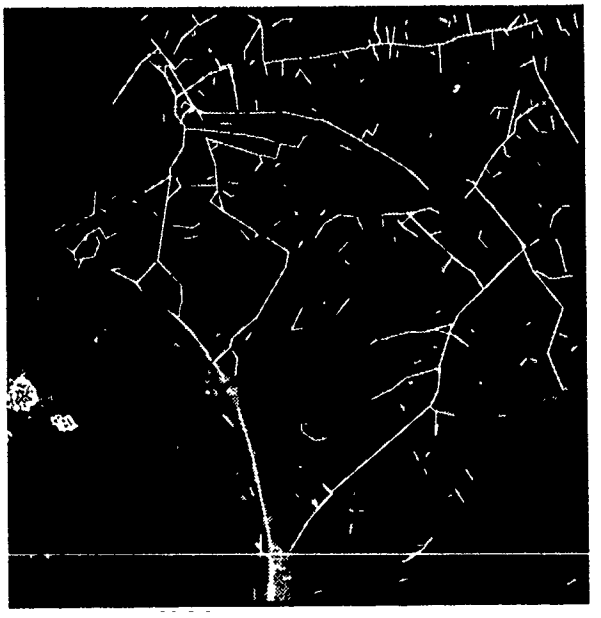


Figure 3: Madelaine Island SPOT PLA Scene With Extracted Apars Overlayed

Figure 4: Sherbrooke SPOT PLA Scene

Figure 4 is a SPOT PLA scene of Sherbrooke, Quebec. The extracted apars are shown as a graphics overlay in figure 5. The apars that are of a similar feature type are grouped together as shown by the graphics overlay in figure 6.

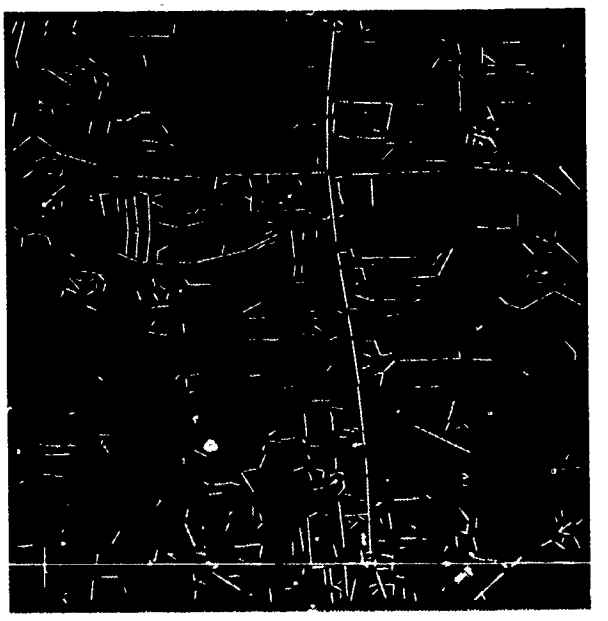


Figure 5: Sherbrooke SPOT PLA Scene With Extracted Apars Overlayed

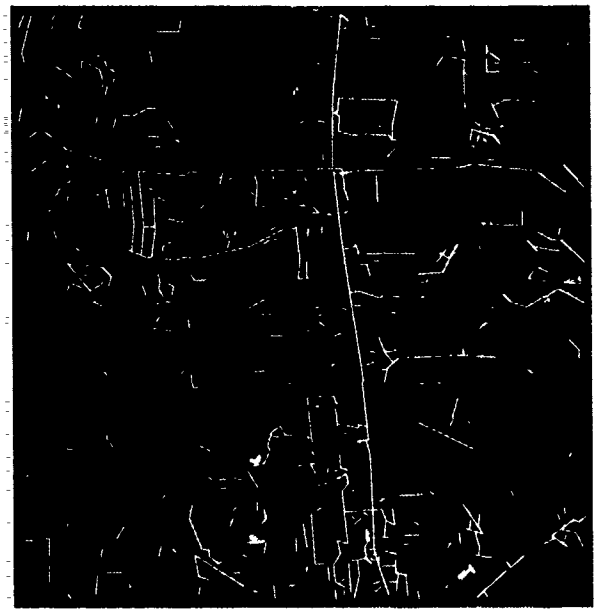


Figure 6: Sherbrooke SPOT PLA Scene With Connected Apars Overlayed

The user utilizes domain specific knowledge to determine what geometric and spectral constraints are used to reduce the set of primitives to a set that are highly likely to map onto the feature of interest. Bit masks that can be obtained from traditional pixel-based classification techniques can prune regions or keep regions of primitives. Additional information about a primitive can also be incorporated - for example, given an apar: interior statistics from another spectral band can be obtained; DEM (digital elevation model) information can allow the determination of a rate of elevation change across the length or width of the apar. DEM information can constrain what primitives to keep (Streams flow down elevation grades).

5 Conclusions

The toolbox provides the user with a method of interactively constraining what primitives to view in order to define a mapping from the primitives onto a geographic feature. The photointerpreter makes use of high level domain dependent knowledge to reason about the low level information which the toolbox algorithms provide. The results show the applicability of the tools with SPOT PLA and digitized aerial imagery. An interactive visualization and control of the primitives can make the process of photointerpretation faster.

References

[1] Michael Boldt and Richard Weiss, "Token-Based Extraction of Straight Lines". Technical Report COINS technical report 87-104, Dept. of Computer Science, University of Massachusetts, October 1987.

[2] Paul R. Cooper, Daniel Friedmann, and Scott A. Wood, "The automatic generation of Digital Terrain Models from satellite imagery by stereo". In *Proceedings of the 36'th International Astronautical Congress*, Stockholm, Sweden, October 1985.

[3] A. Huertas, W. Cole, and R. Nevatia, "Detecting Runways in Aerial Imagery". *Proceedings DARPA Image Understanding Workshop*, pp 272-297, Feb. 1987.

[4] D.C. Mason, D.G. Corr, A. Cross, D. Hogg, and A. Tailor, "Knowledge-based Segmentation of Remotely-sensed Imagery". *International Symposium on Mapping from Modern Imagery, Edinburgh UK*, Sept. 8-12 1986.

[5] Sridhar Vasudevan, Robert L. Cannon, and James C. Bezdek, "Heuristics for Intermediate Level Road Finding Algorithms". *Computer Vision, Graphics, and Image Processing*, 44:pp 175-190, 1988.

[6] Jian-Kang Wu, Dou-Shen Cheng, Wen-Tao Wang, and Deng-Lin Cai, "Model-based Remotely-sensed Imagery Interpretation". *International Journal of Remote Sensing*, 9(8):pp 1347-1356, 1988.

[7] John S. Zelek, "FEX: A Knowledge-based System for Planimetric Feature Extraction". *Proc. SPIE: Recent Advances in Sensors, Radiometry and Data Processed for Remote Sensing*, pp 292-296, April 1988.

[8] John S. Zelek, Steve Paine, Peter MacDonald, and Dave Hawkins, "FEX: Automatic Planimetric Feature Extraction". *Kyoto, Japan ISPRS Conference 1988*, 7.p, 746-755, July 1988.

[9] Miao-Liang Zhu and Pen-Shu Yeh, "Automatic Road Network Detection on Aerial Photographs". *IEEE*, 1986.

INNOVATIVE TECHNIQUES FOR THE SPACEBORNE IMAGING RADAR-C INSTRUMENT

Bryan L. Huneycutt

Jet Propulsion Laboratory
California Institute of Technology
4800 Oak Grove Drive, Pasadena, California 91109

Abstract - The Spaceborne Imaging Radar-C (SIR-C) is the first radar in the series of spaceborne radar experiments, that began with Seasat and continued with SIR-A and SIR-B, to use some very innovative techniques. The SIR-C instrument has been designed to obtain simultaneous multifrequency and simultaneous multipolarization radar images from a low earth orbit. It is a multiparameter imaging radar which will be flown during two different seasons. The instrument has been designed to operate in innovative modes such as the squint alignment mode, the extended aperture mode, the scansar mode, and the interferometry mode. The instrument has been designed to demonstrate innovative engineering techniques such as beam nulling for echo tracking, pulse repetition frequency hopping for Doppler centroid tracking, frequency step chirp generating for radar parameter flexibility, block floating point quantizing for data compression, and elevation beamwidth broadening for increasing the swath illumination.
Key Words - synthetic aperture radar, scansar, block floating point quantization, Spaceborne Imaging Radar-C

I. Introduction

The Spaceborne Imaging Radar-C (SIR-C) has been designed to fly on the Space Shuttle in a low Earth orbit, and to operate from a stable platform located in the Orbiter payload bay. SIR-C is the next critical step in the series of spaceborne radar experiments that began with Sea satellite (Seasat) in 1978 [1] and continued with SIR-A in 1981 [2] and SIR-B in 1984 [3, 4]. SIR-C is the next step in the imaging radar series, eventually leading to the Earth Observation System (EOS) imaging radar [5]. Three flights are currently planned for SIR-C in the early 1990's.

The important SIR-C instrument parameters are given in Table 1. SIR-C is a distributed synthetic aperture radar (SAR), using active phased array technology. SIR-C is a dual-frequency, quad-polarization radar operating at L-band and C-Band frequencies. The SIR-C antenna uses a partitioned aperture design for frequency diversity, whereby the L-band aperture is juxtaposed to the C-band aperture. However, for polarization diversity, each radiation patch of the planar array is fed both orthogonally polarized signals, alternately pulsed. The SAR uses the amplitude and phase of successive returns to synthetically generate a fine resolution antenna.

A low-power linear frequency modulated (LFM) pulse is produced in the radio frequency (RF) exciter at intermediate frequencies, up-converted to L-Band and C-band frequencies, and driven to the antenna at a medium power level. The RF signal is coupled via an RF distribution network in the antenna subsystem to the panels and each of the transmit/receive (T/R) modules on the panels. The coherent RF

Table 1. SIR-C Instrument Characteristics

PARAMETER	L-BAND	C-BAND
AZIMUTH CO-ALIGNMENT	±8 arc min	±2 arc min
AZIMUTH BEAM STEERING	±2 deg	±1 deg
EXTENDED APERTURE BEAMS	127	127
EXTENDED APERTURE DWELL TIME	20 msec	20 msec
SCANSAR BEAMS	4	4
SCANSAR DWELL TIME	30 msec	30 msec
INTERFEROMETRY PHASE CENTER OFFSET	N/A	3 m
BEAM NULL RATIO	1/PRF	1/PRF
PRF HOPPING DWELL TIME	1 sec	1 sec
NUMBER OF SELECTABLE PRFs	16	16
ELEVATION BEAMWIDTHS	5 to 18 deg	5 to 18 deg
NUMBER OF SELECTABLE BEAMWIDTHS	8	8
DIGITAL CHIRP PULSEWIDTH	8.4, 16.9, 33.8 µs	8.4, 16.9, 33.8 µs
DIGITAL CHIRP BANDWIDTH	10,20 MHz	10,20 MHz
QUANTIZATION	8 bits	8 bits
BFPQ SUBSET	4 bits	4 bits
BFPQ BLOCK LENGTH	128 samples	128 samples
PRF	1240 to 2160 Hz	1240 to 2160 Hz
STALO FREQUENCY	89.994240 MHz	89.994240 MHz
DATA STEERING	Offset Video	Offset Video
NUMBER OF HIGH RATE RECORD CHANNELS	4	4

signals are phase-shifted, amplified by the high-power amplifiers in the T/R modules, and coupled to the subarray of radiating patches for transmission to the surface of the Earth. Reflected echoes are captured by the subarray of patches, amplified by the low-noise amplifiers in the T/R modules, and coupled to the phase shifters. The sub-array contributions are combined via the distribution network, coupled to the receivers where the signals are amplified, and further downconverted to offset video frequencies. There are four receiver channels, two for L-band and two for C-band frequencies.

The four offset video signals are steered to one or more assemblies which digitize the offset video signals and format the data. The four digitized data streams are routed to payload high rate recorders for on-board storage, and one of these data streams can simultaneously be steered to the Shuttle high rate downlink to be recorded on the ground. The data, which were previously stored on-board, can be played back later to the ground for data quality monitoring.

II. Innovative SIR-C Operating Modes

Seasat, SIR-A, and SIR-B instruments operated at the single L-band frequency and single horizontal polarization. SIR-C has been designed to operate simultaneously at both L-band and C-band frequencies, and to utilize quad-polarization returns at each frequency. SIR-C effectively has four separate radars: L-band horizontal L_H, L-band vertical L_V, C-band horizontal C_H, and C-band vertical C_V. In the nominal mode, SIR-C operates each of its four radars by radiating at broadside, utilizing a fully focused antenna aperture, much like the previous experiments, except that now there

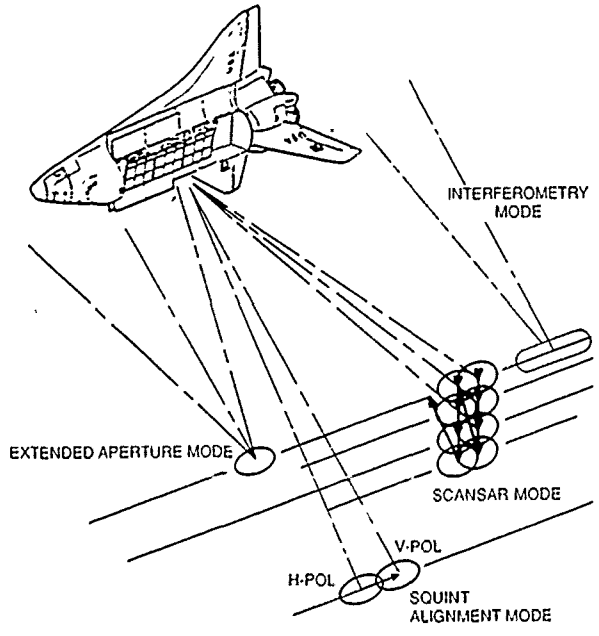


Figure 1. SIR-C Target Illumination by Modes.

are really four radars operating simultaneously. In addition to the nominal mode, SIR-C uses innovative techniques to operate in other modes, which were previously not possible.

A. Squint Alignment Mode

Electronic beam steering in azimuth may be needed to align the H- and V-polarization swath illuminations, even though the antenna design calls for an azimuth electrical-to-mechanical boresight error of less than 4 arc minutes. Azimuth steering allows SIR-C to take data at a given squint angle, whereby the azimuth angle is fixed to one side or the other with respect to broadside. As illustrated in Figure 1, if the H-polarization and V-polarization illuminations are offset for either L-band or C-band, the beams can be commanded to different azimuth squint angles until they are co-aligned to within 4 arc minutes. Azimuth steering also allows experimenters to collect SAR data at selected fixed squint angles. Several observations at different squint angles may give information on the azimuth angle dependence of the image return for that target.

Because of the limited area for locating the electronics on each panel, a single T/R module and phase shifter pair feeds a subarray, and not individual elements. This restricts the azimuth steering to only a few degrees. Electronic beam steering can be accomplished for only ± 1 degree in azimuth for C-band, and ± 2 degrees in azimuth for L-band; beyond these limits, the sidelobes increase, the mainlobe broadens, and grating lobes occur at unacceptable levels.

B. Extended Aperture Mode

Azimuth steering allows SIR-C to take data in the extended aperture mode, whereby the same scene on the ground is illuminated as the Shuttle passes the scene, and the effective aperture is thereby extended. This technique increases the available number of azimuth looks in order to improve the image quality. Figure 2 shows typical beam-steered azimuth patterns. To implement this extended aperture mode, as an example, the antenna beam is first squinted ahead at the full-range positive squint angle, and then decremented in small uniform steps of squint angle at uniform time increments, until the antenna beam is looking behind at the full-range negative angle.

C. Scansar Mode

Electronic beam steering in elevation allows SIR-C to image in a scansar mode, whereby the antenna beam is steered to as many as four previously selected elevation angles during each synthetic aperture interval. The beam is electronically steered in elevation as in azimuth by varying the phase-front across the aperture via the microprocessor-controlled phase shifters. The SAR processor must treat the data from each elevation angle as burst data. The azimuth resolution of the processed bursts of data is degraded by a factor of 4; however, the swath width illumination is increased by the same factor. Since the synthetic aperture interval increases with the look

angle off nadir, the dwell time at each of the four elevation angles increases with look angle.

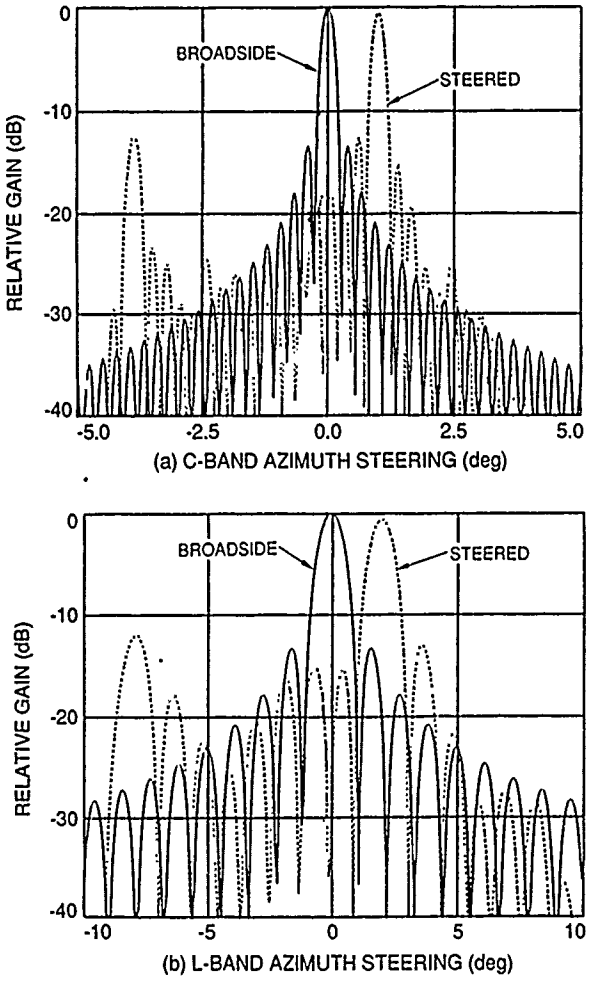


Figure 2. Electronically Steered Antenna Patterns in Azimuth.

D. Interferometry Mode

The L-band and C-band apertures are subdivided into three equal area sections called leaves. The received echo signals from the two extreme C-Band leaves can be coupled into two separate receiver channels in order to accommodate a special interferometry mode. This requires couplers in the C-Band RF feed system at the initial three-way power combiner, such that the inputs from the two extreme leaves can be routed to two separate receiver channels. This technique provides returns from two separate apertures, whose phase centers are separated by the distance of one leaf in the horizontal direction.

III. SIR-C Engineering Techniques

Some innovative engineering techniques were required by SIR-C in order to obtain calibrated data and to be much more flexible in radar parameter selection than in previous imaging radar experiments in this series.

A. Beam Nulling

The phase shifters in one half the array can be shifted 180 degrees in elevation, thereby nulling the elevation beam pattern. This is an echo tracking technique to determine the roll angle, and thereby account for the elevation antenna pattern in the raw SAR data during amplitude calibration. Results from simulations indicate that beam nulling one receive interval during each one second interval is sufficient to determine the roll angle and yet not degrade the image significantly. The roll angle can be determined with an accuracy of a few tenths of a degree. Figure 3 shows a typical beam null pattern.

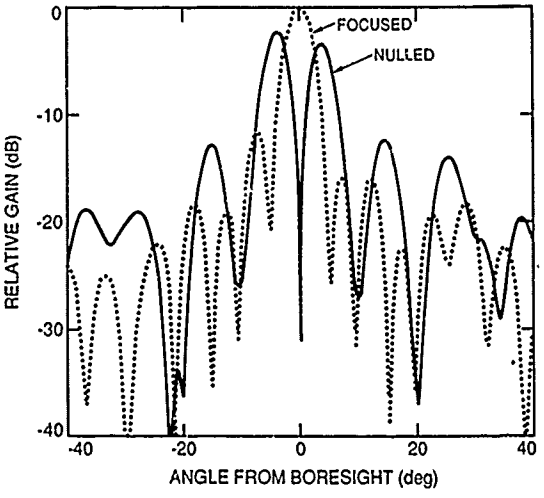


Figure 3. Beam Nulling of Elevation Pattern.

B. PRF Hopping

SIR-C uses a pulse repetition frequency (PRF) hopping technique, whereby the azimuth Doppler spectrum can be located unambiguously. As illustrated in Figure 4, if the azimuth spectrum is offset in frequency because of pointing errors, for example, the centroid of the spectrum falls into different portions of the azimuth processing bandwidth when sampled at various PRF rates. By knowing the ambiguous location of the centroid for three different PRFs, the location of the unambiguous location, and thereby the yaw angle, can be calculated. Present plans are to perform the PRF hopping at the beginning and end of each data take. There will be a one second dwell at each of the three PRFs.

C. Antenna Beamwidth Broadening

For a uniformly illuminated aperture, the aperture dimension and RF wavelength determine the antenna beamwidth. SIR-C uses a technique of tapering the power in elevation, which provides lower sidelobes than for the uniform illumination case. The elevation aperture weighting is shown in Figure 5 for half the aperture. Also, this configuration uses less

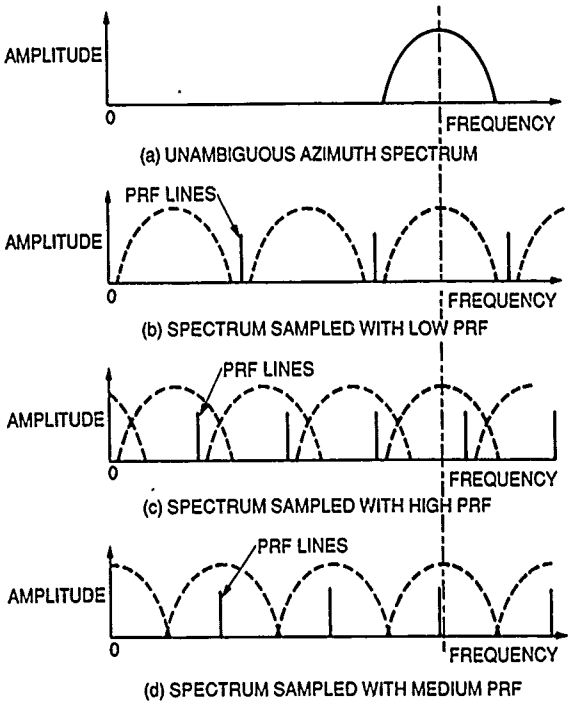


Figure 4. PRF Hopping Technique.

T/R modules and, thus, less power. Lower sidelobes are achieved in elevation. The resultant antenna pattern is lower in amplitude and has a wider beamwidth. Figure 6 shows the typical antenna pattern in elevation.

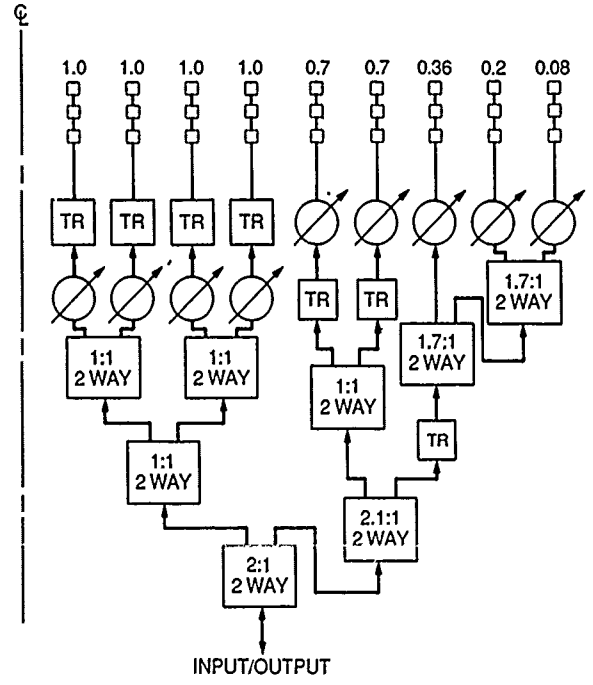


Figure 5. Elevation Aperture Weighting for Half-aperture.

Within the constraints of the Shuttle payload bay volume, the aperture sizes were optimized to have sufficient aperture gain, yet sufficiently low azimuth and range ambiguities over the PRFs selected. The sizes of the full apertures are 12.1 m x 2.95 m for L-band and 12.1 m x 0.75 m for C-band, which yield a fully focused elevation beamwidth of 4.9 degrees for each frequency. Variable beamwidths in elevation can be accommodated by selecting a preprogrammed phase function across the array. One of eight selectable beamwidths from 4.9 degrees to 18.0 degres is selected by uplink command. The amplitude decreases and the sidelobe levels increase slightly with increasing beamwidths. The beamwidth is optimized for swath illumination at the various look angles. Figure 6 shows the antenna elevation patterns for several selectable beamwidths.

D. Digital Chirp

To attain fine resolution in range, SIR-C encodes each transmitted pulse, such that each pulse of duration t can resolve targets as if a much shorter pulse of duration t/TBW, where TBW is the time-bandwidth product, had been transmitted. SIR-C attains this large time-bandwidth product by distributing the energy of each pulse over the frequency bandwidth, linearly with time. The SIR-C chirp signal is generated digitally, in that a tone is successively stepped across the bandwidth within the pulse duration, approximating an LFM (or chirp) signal. Figure 7 illustrates

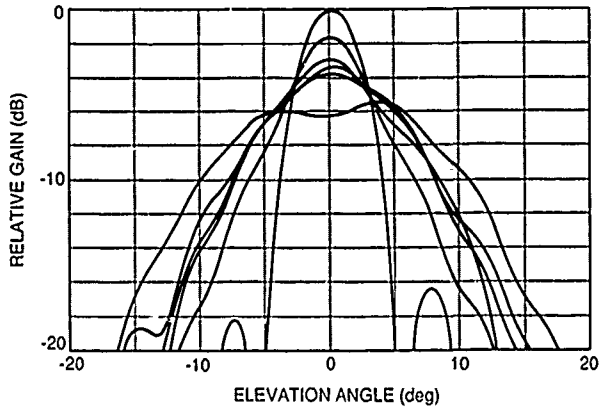


Figure 6. Beam Broadening in Elevation.

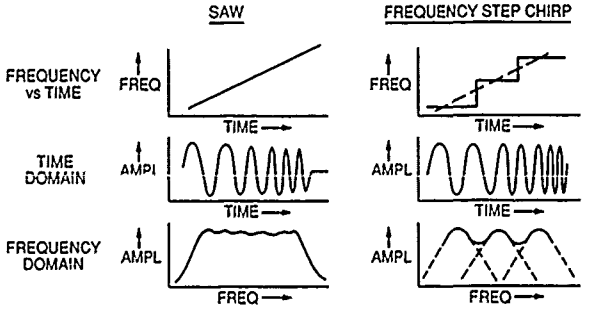


Figure 7. Comparison of SAW Chirp and Frequency Step Approximation of Chirp.

the characteristics of the surface acoustic wave (SAW) device used in SIR-B versus the frequency step chirp used in SIR-C. For ease of illustration, only three steps are shown for the digital chirp. The actual number of steps for SIR-C trades off the integrated side-lobe ratio with realistic switching rates.

The frequency step chirp technique provides flexibility in the selection of certain radar parameters: (1) sense of chirp slope, (2) pulsewidth, (3) calibration tone frequency, and (4) bandwidth. One can distinguish the polarization of the signals by alternately changing the sense of the slope of the transmitted chirp signal. The transmitted pulse width can be decreased by a factor of 2 or 4 to lower the average dc power usage. Fixed frequency calibration tones can be generated from the same digital chirp device, and the frequency of the calibration tone can be selected. The RF bandwidth can be increased from 10 MHz to 20 MHz, and perhaps experimentally even to 40 MHz, to improve the range resolution.

E. Block Floating Point Quantizing

The offset video output from each of the four receivers is digitized to 8 bits per sample with uniform quantization at a constant rate of 45 MHz. A SIR-C feature allows formatting the raw samples as 4 bits, 8 bits, or (8,4) block floating point. The block floating point quantizer (BFPQ) derives its name from "blocks" of data being uniformly quantized, and subsets of the available bits being selected by a predetermined algorithm, equivalent to moving the "floating point" marker in binary data. The data transmission rate to the on-board recorders is constant, and doubling the number of bits per sample approximately halves the swath width, therefore trading off dynamic range for swath width.

The (8, 4) BFPQ is a data compression technique allowing an output rate similar to a 4-bit uniform scheme but with the dynamic range of an 8-bit system. For each block of data, a series of the sign bit followed by the optimally selected 3 contiguous bits per sample, is followed by a common exponent for the block. The data are transmitted as a serial bit stream. Figure 8 shows the signal-to-noise ratio (SNR) versus the ratio of saturation-to-signal level for the 4-bit uniform quantizer (UQ), the 8-bit UQ, and the (8,4) BFPQ.

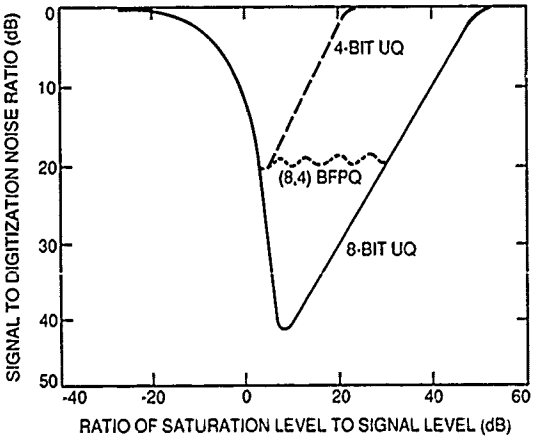


Figure 8. Performance of BFPQ Versus Uniform Quantizers.

F. Integer PRFs

SIR-C has incorporated integer PRFs, whereby there are an exact integer number of PRF pulses in one second. This facilitates timing for the radar since the PRF and timing circuitry provides basic control of the radar for PRF changes, receiver gain, exciter timing, digitized window position, and data channel switching, all of which change state synchronously on one second time ticks. Also, SIR-C chose the PRFs such that there are an integer number of 8-bit bytes in a range line. This feature facilitates ground processing. Since the PRFs are derived from a very stable local oscillator (STALO), the requirement of integer PRFs affected the choice of the STALO frequency. By analysis, it was determined that by using a STALO with frequency 89.994240 MHz, there were at least 16 integer PRFs in the range from 1240 Hz to 2160 Hz which were compatible with SIR-C's altitude and look-angle range.

G. Data Steering

The SIR-C exciter operates in an alternating pulse mode such that the oppositely polarized pulse is delayed by half an interpulse period. The radar effectively operates at twice the nominal PRF. In the quad-polarization mode, both like- and cross-polarization echoes are received in the same channel.

In normal swath modes, each receiver output is directed to a previously selected, separate digital data handling assembly (DDHA) for digitizing, buffering, and formatting. This steering occurs at offset video. Once the individual offset video signals have been digitized, formatted, and serialized, there is no flexibility in steering to the on-board recorders. Each channel of the recorder is hard-wired to an individual DDHA output. Thus, the flexibility in steering data to the recorders really occurs at offset video. This data steering allows for recovery in case one or more of the DDHAs should fail.

In extended swath modes, one receiver output is directed to two DDHAs, such that the resulting data represent an image swath twice that of a single DDHA. However, this puts half of the swath on a separate recorder channel. The SAR processor will have to combine the two channels in order to extend the image swath.

IV. Summary

Several innovative techniques used in the SIR-C instrument have been described. The instrument operating modes include squint alignment, extended aperture, scansar, and interferometry modes. The instrument has been designed to incorporate innovative engineering techniques such as beam nulling, PRF hopping, antenna beamwidth broadening, frequency-step chirp generating, block floating point

quantizing, integer PRFs, and data steering. The instrument flexibility is shown in its choice of values for various radar parameters, such as pulse-width, RF bandwidth, formatting, and beam width, the selection of which requires trading off image quality parameters, such as resolution, image SNR, and swath width.

Acknowledgments

The author would like to thank Charles Elachi, Ed Caro, and Neil Herman of the Jet Propulsion Laboratory; Gary Salisbury of Ball Communication Systems Division, Ball Corporation; Richard Monson of NASA Headquarters; and the many individuals whose contributions made possible the preliminary design of the SIR-C hardware. The work on the instrument as described in this paper was carried out by the Jet Propulsion Laboratory, California Institute of Technology, under contract with the National Aeronautics and Space Administration.

References

1. Jordan, R. L., "The SEASAT - A Synthetic Aperture Radar System," IEEE J. Oceanic Engineering, Vol. OE-5, No. 2, pp. 154-164, April 1980.

2. Cimino, J.B. and Elachi, C., eds., Shuttle Imaging Radar A (SIR-A) Experiment, JPL Publication 82-77, Jet Propulsion Laboratory, Pasadena, California, December 15, 1982.

3. Granger, J.L., "Shuttle Imaging Radar - A/B Sensors," Spaceborne Imaging Radar Symposium, JPL Publication 83-11, Jet Propulsion Laboratory, Pasadena, California, pp. 26-31, July 1, 1983.

4. IEEE Transaction on Geoscience and Remote Sensing, Vol. GE-24, No. 4, July 1986.

5. Shuttle Imaging Radar - C Science Plan, JPL Publication 86-29, Jet Propulsion Laboratory, Pasadena, California, September 1, 1986.

Advances in Spaceborne Active Microwave Sensing

Charles Elachi

Jet Propulsion Laboratory/Caltech
4800 Oak Grove Dr., Pasadena, CA , USA 91109
(818) 354-5673

Abstract

During the 90's, a number of new capabilities and techniques will be used to observe planetary (including earth) surfaces and atmospheres with active or microwave sensors. In 1992, SIR-C will provide the capability of full imaging radar polarimetry at two frequencies L and C-band. In combination with X-SAR, simultaneous imaging at three frequencies will be possible for the first time from space. The SIR-C-XSAR is the predecessor of the EOS SAR, which in addition to the multi-frequency and multi-polarization capability, will be able to provide very wide swath (500 to 600 km) imaging at moderate resolution of few hundred meters. The EOS SAR will also be capable of radar interferometric mapping of the surface topography by using data from repeat orbits.

In the area of altimetry, the next major advance is in the development of scanning altimeters which are capable of providing wide strip topography coverage instead of the line profile required with present systems. Interferometric techniques are also being considered for high resolution ocean topography mapping.

A particular challenge will be the development of a combined imaging/altimetry capability for a Titan radar sensor to be carried on the Saturn Cassini mission. This sensor must have broad flexibility due to our complete ignorance of the properties of Titan's surface. This has to be done in a very restricted weight and power environment.

In the area of scatterometry, considerations are being given to conical scanning scatterometers in contrast to the multiple fan beam scatterometers. This approach allows observation at constant incidence angle across the swath, each point being observed with two azimuth angles. Four azimuth angles are possible with a double beam conical scan.

Finally, active sensors will be used in the 90's to observe and map precipitation from space. Rain radars will allow observation of the rain extent horizontally and vertically and measurement of the rain rate profile.

BISTATIC RADAR IN SPACE
A NEW DIMENSION IN IMAGING RADAR

by

Hans Martin Braun
DORNIER GMBH, Postfach 1420, 7990 Friedrichshafen,

and

Philipp Hartl, Institut for Navigation, University of Stuttgart
F.R.G.

ABSTRACT

Based on a set of examples, this paper shows that bistatic space radar systems can have important capabilities for special imaging radar applications. It explains the different categories of bistatic space radars as there are single orbit systems, crossing orbit systems, tethered systems and systems with geostationary/geosynchronous transmitters (GEO-BISAR). It discusses the ideas of mixed systems, where aircrafts are operating together with spaceborne illuminators either in a cooperating or a parasitic way (BIPAR). It shows that a coordinated operation of monostatic and bistatic radars in co- orbits can effectively be used for measuring the positions of targets on Earth in three dimensions (topographic measurements).
In general, the intension of this paper is to provide a basic understanding of bistatic radar in space for civil applications, its advantages and its drawbacks. It shows that bistatic systems might be viable candidates for future space radars and that further investigations towards them are worthwhile.

Keywords: Bistatic radar, spaceborne radar, BIPAR, BISAR.

INTRODUCTION

Bistatic radar in space is an extremely broad area for investigations. They are systems in which spatial separation exists between the transmitting and the receiving part. Their fundamental principles have been known from the beginning of radar history (Skolnik, ref. 4). However, interest in them declined early, doughtlessly driven by the desire of users, particularly military users, to have radars operated from a single site. This resulted in intensive developments of monostatic radars to a very sophisticated state, until bistatic systems received new interest, when the technological status of advanced data and signal processing allowed bistatic radar to deploy their advantages.

BISTATIC SYSTEMS

Three classes of bistatic space radars can be distinguished:

- low Earth orbit (LEO) systems
- systems with geostationary (GEO) transmitters
- space-/airborne systems, with its most important candidate, the bistatic parasitic radar (BIPAR)

LEO Systems

LEO systems are defined as systems, where both, transmitter and receiver, are flying in low Earth orbits. There are the following major options: Single orbit systems, crossing orbit systems, and tether systems.

In SINGLE ORBIT SYSTEMS transmitter and receiver satellites are flying in the same orbit, one after the other (Fig. 1). Two separate antennas (one for transmit and one for receive) provide a higher transmit/receive isolation than a criculator or TX/RX-switch commonly used in monostatic systems. Whereas in case of monostatic radars the echo reception time and hence the width of the swath is limited to the pause interval between two subsequent pulse transmissions, the bistatic configuration allows to use the entire pulse repetition interval due to the fact that echo reception is not blocked during pulse emission. This leads to a larger swath width being independent from the transmission-pulse-length. Therefore, very large pulse compression ratios up to continuous transmission (cw) can be applied lowering the peak power requirement or improving the radiometric performance (gray level resolution). CW transmission can be advantageous for the overall efficiency of the high power amplifier especially in case of solid state transmitters. These advantages have typically to be payed by the implementation costs for 2 satellites and complex systems for transmitter/receiver synchronization.

If transmitter and receiver are carried on free flying satellites being separated in across track or nadir direction, the system is called CROSSING ORBITS SYSTEM. Fig. 2 shows this for orbits with different inclinations. Due to orbit dynamics the flight paths are crossing at the equator and the satellite distance varies from approximately zero at the crossing points to a maximum at the polar regions. Fig. 3 shows a system with orbit crossings at the poles (different ascending nodes). Its behaviour is similar to that with orbits crossings at the equator, but the maximum cross track separation appears at the equator.

Another type of a CROSSING ORBIT SYSTEM is defined by spatial separation in nadir direction (Fig. 4). In this case the transmitter and the receiver orbits are in the same orbital plane, slightly excentric with equal excentricity but different apogees. This results in different altitudes for transmitter and receiver. The orbit cycle times are kept equal. A similar effect can be obtained by a TETHER SATELLITE SYSTEM with the

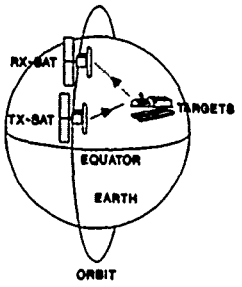


Fig. 1
SINGLE ORBIT SYSTEM

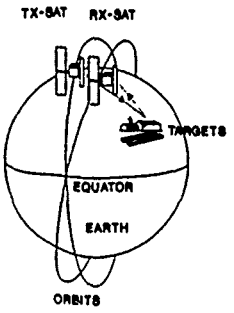


Fig. 2
CROSSING ORBITS SYSTEM WITH DIFFERENTLY INCLINED ORBITS

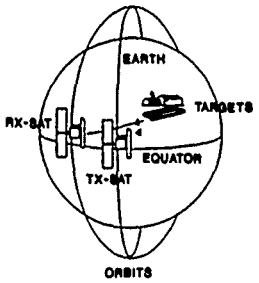


Fig. 3
CROSSING ORBITS SYSTEM WITH DIFFERENT ASCENDING NODES

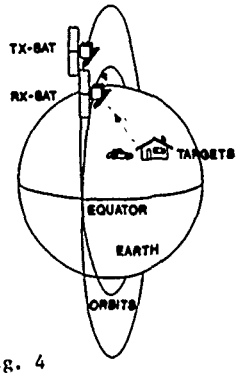


Fig. 4
CROSSING ORBITS SYSTEM WITH SEPARATION IN NADIR DIRECTION

advantage of a constant distance between the satellites (simplified radar geometry), where transmitter and receiver satellites are tethered by a cable of several kilometer length. Tether systems are described in reference 6 for communication systems. These results can be applied to radar systems as well. In this case the synchronization signals and the radar data can be transmitted via the tether cable.

Crossing orbits or tether systems are important candidates for RADAR STEREO IMAGERY. In this case the TX-SAT should additionally be equipped with a receive channel operating in monostatic mode or two receiver satellites should cover the same target area simultaneously from different positions in across track. The monostatic image or the second bistatic image provides the second look required for stereo imagery. Single orbit systems (spatial separation along track), however, can provide stereo imagery only with real aperture operation. Synthetic aperture methods applied to bistatic single orbit systems destroy each kind of stereo information within the radar data.

GEO/LEO Systems

The idea of operating a radar with a powerful transmitter in a GEO and clusters of receivers flying in LEO was born from the desire to provide a high repetition rate of radar coverage for selected target areas on ground. A multibeam transmitter in GEO is illuminating these ground areas continuously or in a scanning mode synchronized to the LEO receivers. Each receiver is operated as soon as it flies over these areas. It sends the received echoes either back to the GEO transmitter, which might act as a relay satellite for these data, or directly down to a ground station. The receivers are passing by the targets closely enough in order to improve the link budget and to provide a radial velocity between targets on ground and the radar allowing for synthetic aperture processing. Fig. 5 shows a typical geostationary transmitter system indicating that many receivers are flying in low Earth orbits.

BIPAR Systems

A BIPAR is an airborne radar receiver typically operating with transmitters of opportunity combining the advantages of spaceborne and airborne radars. The transmitter (e.g. a communications satellite) is located in a geostationary orbit and illuminates continuously the area of interest on ground. The receiver is carried on an airborne platform implying a high mission flexibility, high repetition capability and reducerd power requirements compared to systems with spaceborne receivers. It is called "parasitic" because it uses non-radar signals from communications, navigations or direct broadcasting satellites. The system geometry is shown in Fig. 6.

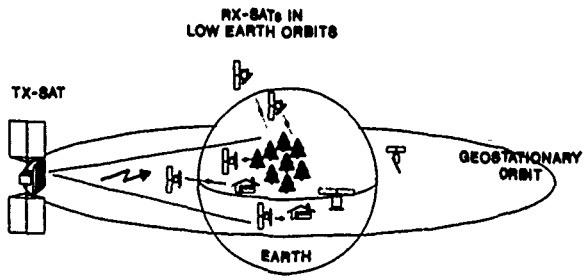


Fig. 5 GEOSTATIONARY TRANSMITTER SYSTEM

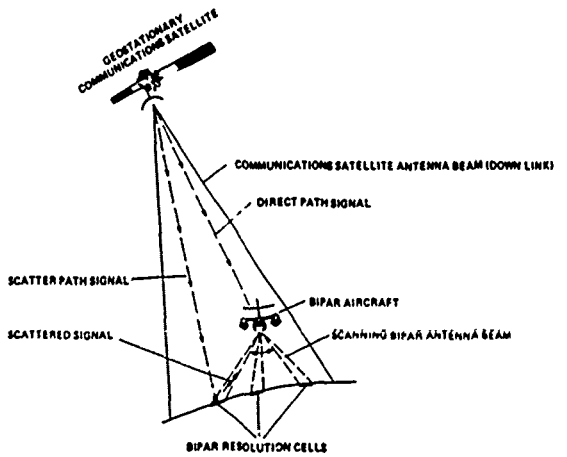


Fig. 6 BIPAR SYSTEM GEOMETRY

ELECTRICAL CONFIGURATION

The general electrical configurations for a TX-SAT and an RX-SAT are shown in Fig. 7. From the mother oscillator a beacon is sent via the direct-path link to the receiver for time and phase synchronization. Initiated by the TX-SAT control computer and based on the mother oscillator, the radar pulse is generated by the pulse generation subsystem, routed through the high power amplifier and the transmit antenna, and sent to ground. The RX-SAT control computer having received the position data for RX-SAT and TX-SAT from the ground control station via general telecommand channels (not shown) estimates the echo time of arrival in advance as input for the timing unit. This unit controls the coherent detection of the received radar echoes. The detected signals are multiplexed with auxiliary data required for data processing on ground. These data are sent to the ground receiving stations, either directly or via an in-orbit data relay station. The overall block diagram of a BIPAR receiver is shown in Fig. 8. It differs from fig. 7 by a more sophisticated signal processing.

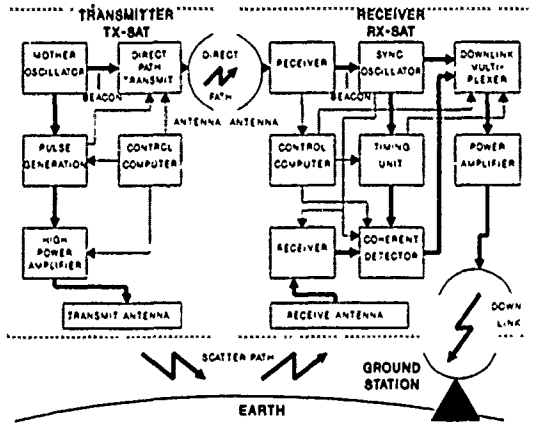


Fig. 7 BLOCK DIAGRAM OF A BISTATIC RADAR

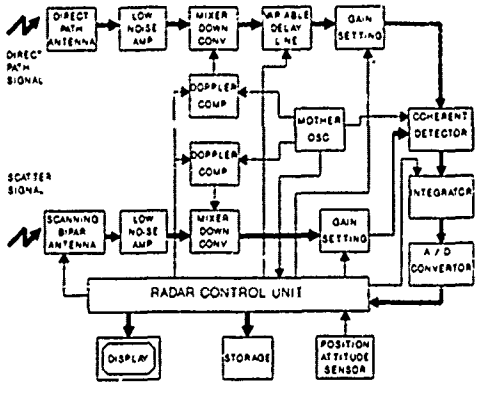


Fig. 8 BLOCK DIAGRAM OF A BIPAR RECEIVER

PERFORMANCE CONSIDERATIONS

Fundamentals concerning orbit geometry, isodops and isodels, and link budgets are described for completely space based systems by Kiyo Tomiasu, 1978 and for BIPAR systems by P. Hartl and H.M. Braun, 1988. The following approximations can be used for performance estimates.

Maximum width of swath w_{max}:

$$w_{max} \approx \frac{PRI \cdot c}{\cos \gamma_{TF} + \cos \gamma_{RF}} \qquad (1)$$

PRI is the pulse repetition frequency, c is the velocity of light, γ_{TF} is the grazing angle of the transmit- to-target path, and γ_{RF} is the grazing angle of the target-to-receive path at far end of the swath.

Range resolution on ground δ_R:

$$\delta_R = \frac{\tau_s \, c}{\cos \gamma_T + \cos \gamma_R} \qquad (2)$$

τ_s is the effective radar pulse-length after range compression, and γ_T and γ_R are the grazing angles of the transmit and the receive path. δ_R is a function of target position within the swath.

Doppler resolution δ_D:

$$\delta_D = \frac{1}{\lambda} [v_T \left(\frac{R}{R+A_T}\right) \Delta\beta_T \sin \beta_T + v_R \left(\frac{R}{R+A_R}\right) \Delta\beta_R \sin \beta_R$$

$$- v_s \left(\Delta\beta_R \sin \beta_{SR} + \Delta\beta_T \sin \beta_{ST}\right)] \qquad (3)$$

In order to estimate the azimuth resolution on ground δ_A, two targets (T1 and T2) can be defined, located on an iso-range (isodel) line with a distance of δ_A determining the transmit and receive viewing angles $\Delta\beta_T$ and $\Delta\beta_R$. A viewing angle is the angle between the paths to/from the targets 1 and 2 at the TX or RX satellite. β_T and β_R are the angles between the target paths and the satellite velocity vectors. β_{ST} and β_{SR} are the angles between the respective target path and the target velocity vector. R is the radius of the Earth, A_T and A_R are the altitudes of the transmit and the receive satellite, v_T and v_R are the velocities of the satellites, v_S is the target velocity, and λ is the radar wavelength.

The total SAR integration time T_i is given by

$$T_i = \frac{N_L}{\delta_D \, \eta_i} \qquad (4)$$

where N_L is the number of independent looks and η_i is the integration efficiency.

The signal to noise ratio $\left(\frac{S}{N}\right)$ is determined by

$$\left(\frac{S}{N}\right) = \frac{P_{Tm} \, T_i \, \eta_i \, \sigma_b}{(4\pi)^3 \, \eta_p (k \, T_R) \, F_s} \cdot \frac{\lambda^2}{D_{TS}^2 \, D_{SR}^2} \cdot \frac{G_{TS} \, G_{RS}}{L_{PT} \, L_{PR} \, L_S} \qquad (5)$$

where P_{Tm} is the average transmit power, σ_b is the bistatic radar cross section, G_{TS} and G_{RS} are the antenna gains transmit-to-target and receive-to-target, η_p is the time-bandwidth product, k is the Boltzman's constant, T_R is the receiver temperature, F_S is the noise factor, D_{TS} and D_{SR} represent the distances between transmit satellite - target - receive satellite, and L_{PT}, L_{PR} and L_S are margins for rf propagation losses on transmit/receive path and overall system losses.

SYSTEM EXAMPLE: GEO-BISAR

A transmitter in geosynchronous orbit illuminates the Earth's surface below the receiver satellites in low Earth orbits with a large multibeam phased array antenna. The RX-SATs synchronize their timing by use of the direct path signal from the TX-SAT. They receive the echo data as long as the TX-SAT phased array antenna beam is steered to their antenna footprints. They detect and format these data and send them to the ground receiving station via the TX-SAT data relay channels (Fig. 9). Typical system parameters are given in Tab.1. The radiometric performance is shown in Tab.2. The 8dB overall system losses cover propagation losses, hardware deficiencies and integration efficiency.

Spatial resolution	30 m x 30 m
TX-SAT orbit altitude (GEO)	36.000 km
RX-SATs orbit altitude (LEO)	600 km
Off-nadir angle of antenna beams	~ 20 deg
Velocity of RX-SATs	~ 6.5 km/s
Doppler resolution	1 Hz
Doppler bandwidth	250 Hz per look
Number of looks	3
Integration time per look	1 s
Total integration time	3 s
TX-SAT antenna dimension	30 m x 30 m
RX-SAT antenna dimension	10 m x 2 m
Radar frequency	1 GHz
RF bandwidth	18 MHz

Tab.1: GEO-BISAR System Parameters

	value	dezibel
+ Mean RF power	1 kW/beam	30 dBw
+ Integration time per look	1 s	0 dBs
+ Minimum RCS	~ 90 m^2	20 dBm^2
+ (Wavelength)^2	(0.3 m)^2	-10 dBm^2
+ Transmit antenna gain		50 dB
+ Receive antenna gain		34 dB
- $(4\pi)^3$		-33 dB
- ($k\ T_0$)		+204 dB/Ws
- System noise figure		-4 dB
- (Transmit path length)^2	(36.000 km)^2	-151 dB/m^2
- (Receive path length)^2	(640 km)^2	-116 dB/m^2
- Overall system losses		-8 dB
Signal-to-noise ratio		+16 dB
Radiometric resolution (3 looks)		2 dB

Tab.2: GEO-BISAR Link Table

CONCLUSION

Bistatic radar in space is an extremely broad area for investigations. A detailed discussion of all possible concepts and their analytics would fill up a book completely. Hence, this short paper tries to concentrate on major aspects in order to provide a basic overview. Bistatic space radar might be a viable sensor for future remote sensing instrumentations and further investigations are worthwhile.

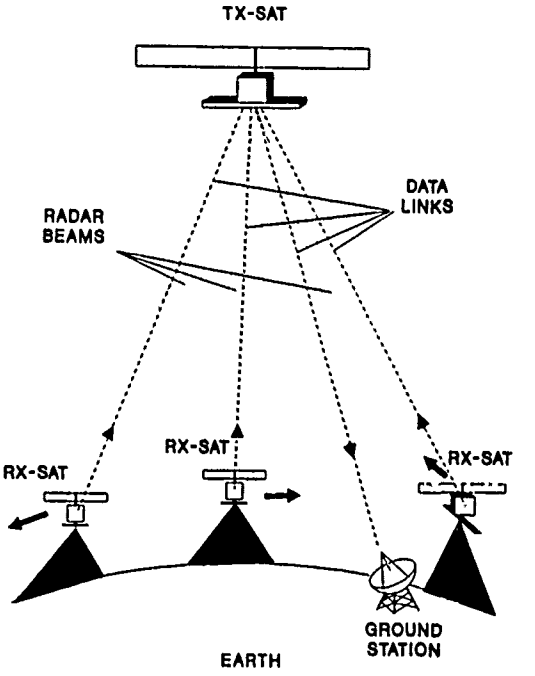


Fig. 9 A BISAR SYSTEM WITH ONE GEOSTATIONARY TRANSMITTER AND RECEIVERS IN LOW EARTH ORBITS (GEO-BISAR)

REFERENCES

1. George Heilmeier
"Radar: Bistatic Arrays in Space?"
Microwave System News, Jan. 1978, pp. 39 - 40.

2. K. Milne
"Principles and Concepts of Multistatic Surveillance Radars",
Radar-77, Conf. Publ. No. 55, IEE, pp. 46 - 52,
ISBN: 0852961812.

3. Kiyo Tomiyasu
"Bistatic Synthetic Aperture Radar using two Satellites"
1978 IEEE EASCON, Sept. 1978, Publ. 78CH-1352-4AES

4. M. Skolnik
"Introduction to Radar Systems",
McGraw Hill, 1962, pp. 585 - 594.

5. A.P. Hulbert et.al.
IEE Proceedings on Communications, Radar and Signal Processing
"Special Issue on Bistatic and Multistatic Radar",
Dec. 1986, Vol. 133, No. 7, UK ISSN 0143-7070
IEE Publications Sales Department, UK,

6. Tetsuo Yasaka, Takeski Hatsuda
"Geostationary Tether Satellite System and its Application to Communication Systems", IEEE Trans. on Aerospace and Electronic Systems, Vol. 24, No. 1, Jan. 1988, pp. 68 - 75.

7. P. Hartl, H.M. Braun
"A Bistatic Parasitical Radar (BIPAR)", ISPRS Kyoto 1988, Com. 1, pp. 45 - 54

Wide Swath Ocean Topography Mapping with Interferometric Altimeters

E. Rodriguez, F. Li

Jet Propulsion Laboratory, California Institute of Technology
4800 Oak Grove Drive, Pasadena, CA 91109

C.L. Parsons, and E. J. Walsh

NASA/Goddard Space Flight Center, Wallops Flight Facility
Wallops Island, VA 23337

Abstract: An interferometric radar altimeter is proposed to provide wide swath, high resolution ocean topography. Several system design issues of such an interferometric altimeter are presented. We show the tradeoffs between processing of the interferometric signal using the so-called amplitude approach and the so-called phase approach. We also discuss the systematic errors associated with uncertainties in the interferometer baseline and the attitude of interferometer orientation. An approach using the measurements at orbit cross over regions, together with the topography measurements from a traditional nadir looking altimeter which are not contaminated by the baseline and attitude noises, is described. Preliminary simulation results show that such an approach can generate an acceptable error level if the ocean surface does not change appreciably between the observations.

1. Introduction

The use of nadir looking radar altimeters for measuring ocean topography is a well established technique through the flight of such altimeters on SKYLAB, GEOS, SEASAT and GEOSAT. The altimeter system that is planned for the TOPEX/POSEIDON mission will perform high accuracy topography mapping over the global oceans over a period of years. The data generated will be used in various ocean circulations and modelling studies.

The usefulness of such nadir looking altimeters for the study of mesoscale ocean features, however, is limited by the swath that the altimeter generates. Even using the beam-limited footprint on the ocean, the typical swath achieved is less than 25 km. In many missions, the satellites are placed in repeat ground track orbits and with the limited swath, a significant portion of the space-time domain of ocean features cannot be adequately sampled. For example, there are ocean circulation eddies that have spatial extent that are comparable or smaller than the orbital track spacings. These features cannot be mapped accurately with a single nadir looking altimeter beam. A wide swath ocean altimeter system is required.

The simple-minded approach of using several off-nadir antenna beams to measure the topography over a wide swath has significant difficulties. In particular, the sharp rising edge in the return waveforms of nadir looking altimeters will no longer be present. The return waveform will spread over the equivalent range spacing of several tens of meters and to extract an ocean toppography measurement with centimetric level accuracy is nearly impossible. Another difficulty is that the attitude of the antenna beam must be accuractely known in order to allow a proper location of the topopgraphy measurement. A slight error in the antenna pointing angle will translate into large topography systematic errors.

Bush et al. (Bush et al., 1984) proposed an interferometric altimeter technique to effectively generate a much reduced antenna beamwidth, which reduces the range extent of the return waveform. Parsons and Walsh (Parsons and Walsh, 1989) discuss some of the issues associated with that approach. A similar technique has also been applied to synthetic aperture radar systems (Zebker and Goldstein, 1986). This interferometric SAR approach is different in that it utilizes the phase difference observed between the two interferometer antennas to infer surface topography. Recently, this interferometric SAR concept has been extended into a spaceborne configuration by Li and Goldstein (Li and Goldstein, 1989). In this paper, we present a comparison of the conceptual design between these two approaches. We list the systematic error sources that are present in two approaches. Furhermore, we show that for ocean topograpahy applications, the interferometric baseline and attitude have to be known to extremely high accuracies. We propose an approach to utilize the wide swath topography measurements over the orbit crossing regions to determine the interferometric baseline and attitude. Simulated results using a seven-beam system are shown. Finally, we briefly describe the additional tradeoffs required to confirm the viability of this concept.

2. Analysis of Systematic Error Sources

From Figure 1, one can see that the height above a reference surface can be obtained by means of the equation

$$z = H - r\cos\theta \tag{1}$$

where H is the interferometer height above the reference surface, r is the range from the interferometer to the selected height, θ is the incidence angle, and z is the height to be measured. Both of the measurement concepts discussed in this paper, which we call the phase and amplitude approaches, use this equation to estimate z. The difference between them is that, for the phase approach, r is known and θ is estimated, while, for the amplitude approach, θ is known and r is estimated.

To analyze the systematic error sources in the interferometric measurement, we model the return voltage, v_k $(k = 1, 2)$, at each of the receivers from a point x across-track from the interferometer by the expression

$$v_k(x) = f(x)q(x)G(x)\exp\left(-\frac{2i\pi r_k}{\lambda}\right) \tag{2}$$

$$\langle f(x)f^*(x')\rangle = \sigma_0(x)\delta(x-x') \tag{3}$$

where we have assumed that the phase shift due to the surface height is random and uncorrelated from point to point. In the above expression, G denotes the antenna gain, r_k is the distance from antenna k to x, and σ_0 is the normalized radar cross section. We have also included the term $q(x)$, which represents the "gate gains," ripples in the return power due to the filtering of the return signal to remove aliasing. In deriving the expression above, we assumed that, for the purpose of our analysis, the contribution from the pulse width, and the along-track contribution to the signal could be ignored.

Using this expression, we derive the following statistical properties of the return signals:

$$\langle v_1(x)v_1^*(x)\rangle = \langle v_2(x)v_2^*(x)\rangle = \sigma_0(x)G^2(x)q^2 \tag{4}$$

$$\langle v_1(x)v_2^*(x)\rangle = \sigma_0(x)G^2(x)q^2\exp(-i\phi) \tag{5}$$

where ϕ is the phase difference between the two signals and is given by

$$\phi = \frac{2\pi}{\lambda}(B_x\sin\theta - B_y\cos\theta) + \psi \tag{6}$$

$B_x = B\sin\xi$, $B_y = B\cos\xi$, ξ is the baseline attitude (nominally assumed to be 0), B is the interferometer baseline, λ is the electromagnetic wavelength, and ψ is a factor introduced to simulate phase contamination due to multipath effects, waveguide deformation, or receiver imbalance. The amplitude term in these equations can be modeled as $\sigma_0(x)G^2(x)q^2 = A_{\rm est}(1+n)$, where $A_{\rm est}$ represents the value of this term estimated from *a priori* knowledge (calibration and modeling of the ocean scatter) and n represents the noise due to the error in this knowledge.

In the phase approach, the product $v_1(x)v_2^*(x)$ is formed and the phase difference is estimated by means of the equation

$$\phi = \arctan\frac{\mathrm{Im}(v_1(x)v_2^*(x))}{\mathrm{Re}(v_1(x)v_2^*(x))} \tag{7}$$

The incidence angle for each range bin is then solved for using equation (6). A height estimate for each range bin is then obtained by replacing in equation (1). Height noise can be reduced by averaging different estimates over the spatial resolution cell.

Notice that the amplitude terms cancel in equation (7), so that the phase difference estimate is not contaminated by errors in antenna or filter gain calibration, or in modeling the ocean return. Notice, also, that the phase can only be obtained modulo 2π. The problem of "unwrapping" the phase to obtain the true phase has been addressed elsewhere (Goldstein et al., 1988). Let us mention, however, that making a 2π error in unwrapping the phase results in a height error estimation of tens of meters. Since the ocean geoid is known to much greater precision, we do not anticipate that phase unwrapping will present significant problems for ocean altimetry

In the amplitude approach, the voltage from both antennas is summed and the resulting signal is square law detected The resulting signal is given by

$$P(x) = 2A_{\rm est}(1+\cos\phi(x))N \tag{8}$$

where $N = 1+n$. The height estimate is then made by tracking the range to the peak of a preselected lobe in the interference pattern. By selecting a given lobe, say the mth lobe, the incidence angle is automatically selected by the equation $\sin\theta = m\lambda/B$. The detailed nature of the tracking algorithm determines the accuracy with which the range may be estimated. A fundamental limitation of all tracking algorithms, however, will be the actual range to the interference peak which, using equation (6), can be shown to equal

$$r_m = \frac{(H-z)\left(1+(B_y/B_x)^2\right)}{-\frac{B_y}{B_x}\frac{\lambda(\phi_m-\psi)}{2\pi B_x}+\sqrt{1-\left(\frac{\lambda(\phi_m-\psi)}{2\pi B_x}\right)^2+\left(\frac{B_y}{B_x}\right)^2}} \tag{9}$$

Systematic errors in determining B_x, B_y, and ψ will result in systematic errors in the height estimate. By differentiating equations (1) and (6) (for the phase approach), or equations (1) and (9) (for the amplitude approach), we obtain the systematic error associated with each parameter:

$$\delta z_1 = \frac{r\sin^2\theta}{B_x\cos\theta+B_y\sin\theta}\delta B_x \tag{10}$$

$$\delta z_2 = \frac{r\sin\theta\cos\theta}{B_x\cos\theta+B_y\sin\theta}\delta B_y \tag{11}$$

$$\delta z_3 = \frac{\lambda r\sin\theta}{2\pi(B_x\sin\theta+B_y\sin\theta)}\delta\psi \tag{12}$$

Notice that the sensitivities for both approaches are identical, even though the parameters estimated are different.

To see the practical implications of these equations for ocean satellite interferometry, consider a system with orbit height of 800 km, baseline of 12 m, wavelength of 0.8 cm (Ka-band), and incidence angle of 3.5°. If we assume that each of the error sources contribute at most 1 cm height error, the knowledge requirement for the various parameters is given by

$$\delta B < 4.0\text{x}10^{-5}\text{meters} \tag{13}$$

$$\delta\xi \approx \frac{\delta B_y}{B_x} < 4.2\text{x}10^{-2}\text{arcseconds} \tag{14}$$

$$\delta\psi < 0.11° \tag{15}$$

If one assumes that the phase contamination is due to leakage, the last requirement is equivalent to requiring power isolation to be better than -62 dB. As can be seen, these requirements are extremely stringent and probably beyond the capability of present day technology In the next section, we shall discuss a possible approach to overcoming these limitations by using satellite cross-over information.

We next examined the effects of errors in *a priori* knowledge in the antenna calibration, filter gains, and scattering model on the amplitude approach estimated height. We assumed that the parameter N in equation (8) could be modeled as a Gaussian variable, independent from sample to sample, whose mean is equal to 1. We implemented the tracking algorithm described in the paper of (Bush et al., 1986) and performed a Monte Carlo simulation of the tracker error induced by variations in N. The system parameters are the ones in the previous example, and the system bandwidth is 300 MHz The results are presented in the following table.

N (dB)	Height Noise (cm)
0.1	3.0
0.05	1.5
0.01	0.95

As can be seen, the accuracy requirement in the calibration of the antenna pattern and the filter gains is very stringent. On the other hand, these error sources are not present when the phase estimation approach is used.

3. Removal of Systematic Errors Using Cross-Over Tracks

The accuracy requirements on the knowledge of the interferometric attitude and baseline are so stringent that they are nearly impossible meet by present day technology In this section, we examine the feasibility of using the data obtained by the interferometric altimeter and a nadir looking altimeter over the regions where the orbital tracks cross to determine the baseline and attitude We assume that the surface height change is small in the time between the two satellite passes For an n beam interferometer in which the n beams are symmetrically placed about nadir, there will be n^2 cross-over points arranged in a diamond pattern when the orbit cross-over occurs.

The estimated height error for the k^{th} pass (k =1,2) will be modeled as

$$z_{ij}^{(k)} = z_{ij}^{(k)} + \epsilon_{(\mathrm{bx})ij}^{(k)} + \epsilon_{(\mathrm{by})ij}^{(k)} + \epsilon_{(\mathrm{orbit})ij}^{(k)} + \epsilon_{(\mathrm{range})ij}^{(k)} \tag{16}$$

where the indeces i and j refer to the beams of the first and second interferometer, respectively; $z_{ij}^{(k)}$ is the estimated height at this cross-over point, $z_{ij}^{(k)}$ is the true surface height, $\epsilon_{(\mathrm{bx})ij}^{(k)}$ is the baseline error in the cross-track direction and given by

$$\epsilon_{(\mathrm{bx})ij}^{(k)} = \frac{r_{ij}^{(k)}\sin^2\theta_{ij}^{(k)}}{\cos\theta_{ij}^{(k)}B_x^{(k)}+\sin\theta_{ij}^{(k)}B_y^{(k)}}\delta B_x^{(k)} \equiv a_{ij}^{(k)}\delta B_x^{(k)} \tag{17}$$

$\epsilon_{(\mathrm{by})ij}^{(k)}$ is the baseline error in the height direction

$$\epsilon_{(\mathrm{by})ij}^{(k)} = \frac{r_{ij}^{(k)}\sin\theta_{ij}^{(k)}\cos\theta_{ij}^{(k)}}{\cos\theta_{ij}^{(k)}B_x^{(k)}+\sin\theta_{ij}^{(k)}B_y^{(k)}}\delta B_y^{(k)} \equiv b_{ij}^{(k)}\delta B_y^{(k)} \tag{18}$$

$\epsilon_{(\mathrm{orbit})ij}$ is the error in the orbit estimation over the cross-over diamond which we have modeled as

$$\epsilon_{(\mathrm{orbit})ij} = c^{(k)} + \epsilon_o \tag{19}$$

where $c^{(k)}$ is a constant and ϵ_o is a Gaussian random variable of adjustable variance which mimics residual orbit error and random surface motions; finally, $\epsilon_{(\mathrm{range})ij}$ is the interferometer random range error which is assumed to be the result of 2cm noise for the nadir beam and 5cm noise for the off-nadir beams.

Assuming that the ocean surface does not change between the two tracks, one can take the estimated height difference at each cross-over point to obtain

$$\Delta z_{ij} \equiv \hat{z}_{ij}^{(1)} - \hat{z}_{ij}^{(2)} \quad (20)$$

$$= a_{ij}^{(1)}\delta B_x^{(1)} + b_{ij}^{(1)}\delta B_y^{(1)} - a_{ij}^{(2)}\delta B_x^{(2)} - b_{ij}^{(2)}\delta B_y^{(2)} + \Delta c + \epsilon_{ij} \quad (21)$$

where $\Delta c = c^{(1)} - c^{(2)}$ and ϵ_{ij} is a Gaussian random variable whose standard deviation σ_{ij} which can be can be calculated using the range and orbital noise standard deviations. The problem is now in the standard form for weighted least squares estimation of the five linear parameters $\delta B_x^{(1)}$, $\delta B_y^{(1)}$, $\delta B_x^{(2)}$, $\delta B_y^{(2)}$, and Δc. Errors in the parameters $\delta B_x^{(1)}$ and $\delta B_x^{(2)}$ are the most critical in the calculation of ocean vorticity since they introduce quadratic errors in the cross-track distance. The parameters $\delta B_y^{(1)}$ and $\delta B_y^{(2)}$, on the other hand, will, to first order, only introduce surface tilts.

The standard deviations in the estimated parameters can be easily calculated by using linear error propagation. The accuracy of the estimated parameters depends on the number of cross-over points which, in turn, depend on the number of beams. To study this dependence, we examined interferometers with 3, 5, and 7 beams and assumed that the beams were separated by 1.75°. The results for the standard deviations in B_x and B_y are presented in Figure 2 for two values of the orbit error. The implication of these figures for the height noise measurement at 50 km cross-track is presented in Figure 3.

To study the biases induced by surface motion, we simulated a simplified moving ocean eddy. The eddy was assumed to be Gaussian with a height of 50 cm and standard deviation of 60 km. It was assumed that the eddy moved 100 km per month and that it decayed linearly so that it completely dissipated also in one month (this was found to be the worst case among several eddies studied). The biases in the estimated baseline as a function of the cross-over repeat time is plotted in Figure 4 for 5 and 7 beams. In this figure, the "1st pass", and "2nd pass" refer to the estimated baseline error in the first and second passes, respectively. Notice that due to the shape of the simulated moving eddy, the baseline estimate for the five beam interferometer is in greater error for the second pass than for the first pass. As can be seen, for a 7 beam interferometer, height biases of less than three centimeters will be incurred for cross-overs separated by as long as 9 days. On the other hand, for a 5 beam altimeter, acceptable errors are only obtained when the cross overs are separated by about three days.

4. Summary

We have presented a set of system design issues associated with an interferometric altimeter for wide swath ocean topography mapping.The systematic error sources for the so-called phase and amplitude approaches are summarized. We also show the feasibility of using the data obtained in the orbit crossover regions to determine the attitude and baseline for the interferometer configuration. Additional issues that must be studied in detail include: receiver phase imbalances, relative phase knowledge between the antenna beams, spatial gradient of radar cross sections (Parsons and Walsh, 1989), etc. The removal of atmospheric and ionospheric delays from the height measurements must be evaluated too.

Acknowledgement

The research described in this paper was carried out by the Jet Propulsion Laboratory, California Institute of Technology, under contract with the National Aeronautics and Space Administration.

References

Bush, G.B., Dobson, E.B., Matyskiela, R., Kilgus, C.C., and Walsh, E.J., "An analysis of a satellite multibeam altimeter," Marine Geodesy, vol. 8, nos. 1-4, pp. 345-384, 1984.

Goldstein, R.M., Zebker, H.A., and Werner, C.L.,"Satellite radar interferometry: Two-dimensional phase unwrapping," Radio Science, vol. 23, n0. 4, pp. 713-720, 1988.

Li, F.K., and Goldstein, R.M.,"Studies of multi-baseline spaceborne interferometric synthetic aperture radars," to be published in J. Geophys. Res.

parsons, C.L., and Walsh, E.J.,"Off-nadir radar altimetry," IEEE Trans. Geosci. Remote Sensing, vol. 27, n0. 2, 1989.

Zebker, H.A., and Goldstein, R.M.,"Topographic mapping from interferometric synthetic aperture radar observations," J. Geophys. res., vol. 91, pp. 4993-4999, 1986.

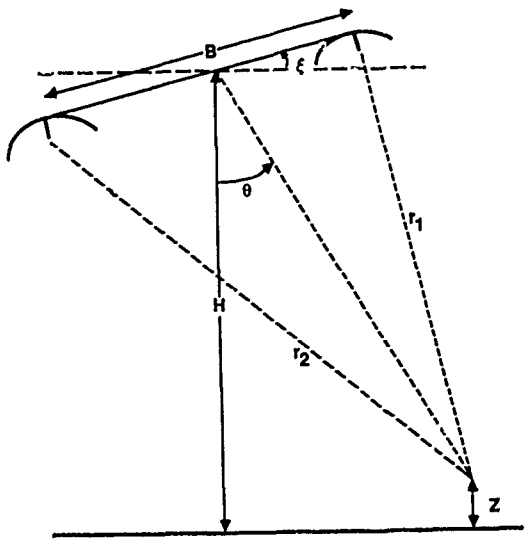


Figure 1

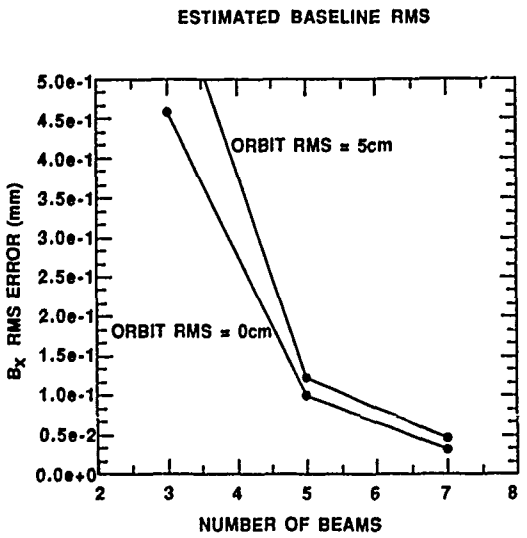


Figure 2(a)

BY ESTIMATION NOISE

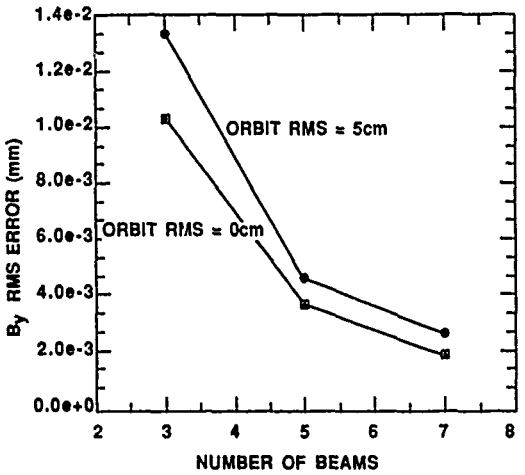


Figure 2(b)

HEIGHT ERRORS AT 50 KM

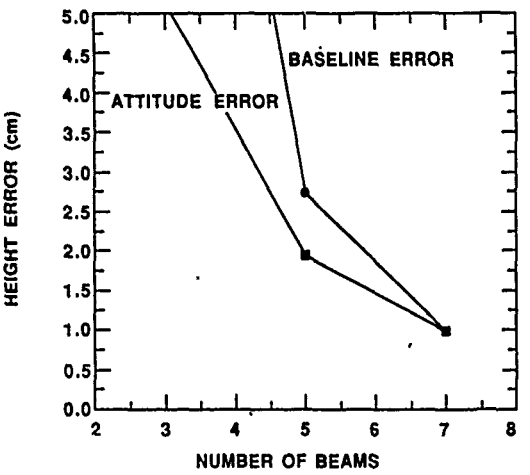


Figure 3

BASELINE ERRORS FROM EDDY MOTION
DECAYING EDDY

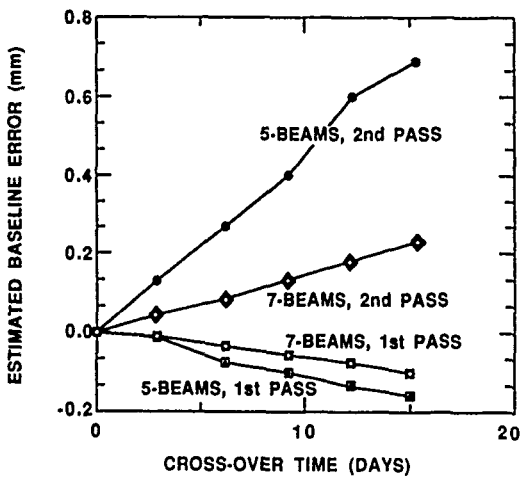


Figure 4

A PHASED ARRAY BREAD BOARD FOR FUTURE REMOTE SENSING APPLICATIONS

R.W. Zahn, E. Schmidt

Dornier GmbH
7990 Friedrichshafen
West Germany

ABSTRACT

The next generation of SAR antennas will be of the active phased array type. This paper describes the ongoing development of a phased array bread board for remote sensing. Starting from a detailed system design a functional representative bread board was developed. The design and the performance are discussed in the following.

Keywords: Phased Array Antenna, Radar, Synthetic Aperture Radar, Microwave Technology, Microstrip Antenna.

1. Introduction

Future Synthetic Aperture Radar systems (SAR) will introduce flexible mode operations to scan the earth's surface. This requires phased array technologies with adaptive beam steering and beam forming capabilities.

Dornier is presently leading a multi national project under ESA contract to define the technology of advanced SAR systems. This work includes the development of a phased array antenna bread board to demonstrate the capabilities for advanced SAR applications. It is composed of a radiating board, T/R modules, beam forming networks, a power supply and distribution, a control system and structural parts. Main contributors to the antenna breadboard hardware are Dornier, Ericsson Radar Electronics, and Selenia Spazio.

The nominal performance and requirement parameters for the bread board are derived from a comprehensive system design to ensure the adequacy to a space system. The bread board undergoes detailed testing on component and system level to prove the expected performance.

The present paper will show the overall layout of a space-borne phased array SAR antenna, the design of the bread board model, the subsystem performance, and a dicussion of the implications for the complete system.

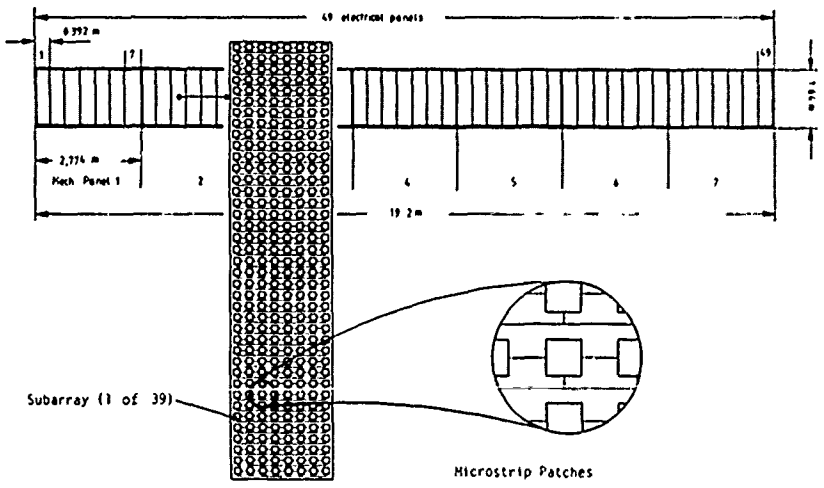


Fig. 2-1: Principal antenna lay-out

The present system is operating at 5.3 GHz with 50 µs pulsewidth and 50 MHz bandwidth. Due to the required flexibility of the SAR instrument the antenna has to cover a wide range of possible beamwidths and steering angles. The swath width shall be selectable between 50 km and 500 km using high and low resolution SAR modes. An additional mode to normal SARs is the spot beam mode which requires azimuth steering capabilities of the antenna. In all modes either H or V polarisation shall be transmitted and both H and V shall be measured on receive.

2. Space-borne Phased Array Antnena Layout

In order to perform azimuth and elevation beam steering the radiating aperture has to be divided in several independently phase controllable sections or elements. The beam steering requirements are ±1.5 deg. in azimuth and ± 10.7 deg. in elevation.

Due to the small azimuth steering range the antenna length can be divided in sections larger than one wavelength, thus, simplifying the antenna feed network but causing grating lobes to appear in the radiation pattern. With an azimuth grating lobe level of -13 dB the maximum width of one section is restricted to about 0.4 m, which gives 49 electrically independent panels. This allows to combine groups of 7 electircal panels to 7 mechanical panels.

The elevation scan range requires an element spacing less than 0.8 λ_0. From the SAR system design the smallest elevational beamwidth is 1.75 deg., equivalent to an antenna width of 1.64 m, or 39 subarrays. Fig. 2-1 shows the principal antenna layout.

The radiating board consists of microstrip fed patch arrays with dual polarization capabilities. 8 patches form a subarray which has its own amplifier and phase shifter. Thus, almost 2000 electronic modules are distributed over the entire array.

The elevational transmit pattern has to fulfil certain requirements (antenna mask) concerning beam shape and gain. Fig. 2-2 gives a synthesized pattern for the wide swath mode. On receive, the beam shall dynamically follow the direction of the incoming pulse echo. This procedure allows a higher antenna gain, and therefore, a better signal to noise ratic for the SAR instrument. The improvement of the SNR is necessary because the cross polar components shall be detected with an acceptable quality.

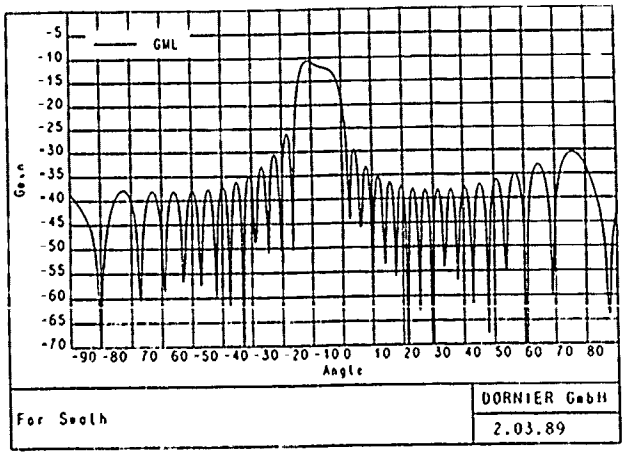


Fig. 2-2: Far swath synthesized antenna pattern

Apart from the T/R modules there are several other electrical subsystems: power supply, antenna control computer, beamforming networks, power and control networks.

The antenna mechanical design is largely determined by the electrical layout, e.g. the subdivision of the radiating aperture in electrical panels an the total aperture size which results from beamwidth constraints. Due to launcher constraint (ARIANE IV is envisaged) the antenna has to be folded to a stowed configuration, thus, the antenna is subdivided in several sections linked together via hinges.

3. Bread Board Design

The B/B functions are to a great extend derived from the space antenna functions. Essentially, the B/B shall be able to transmit fixed length RF pulses of a given frequency and with a given pulse repetition rate. The transmit beam shall be steerable to different directions by electronic means and shall also be variable in terms of beamwidth and beamshape. The polarization of the radiated field shall be switchable between horizontal and vertical.

On receive the antenna aperture is illuminated by a RF CW signal of the same frequency as transmitted. The receive beam shall be stepped over a distinct angular sector with a angular velocity corresponding to the space antenna receive scan mode, hence, the scanning shall be done in real time. Beside beam scanning, fixed receive beam positions shall be selectable within the angular scanning range. Both polarizations shall be received simultaneously. Therefore, the receive section comprises dual polarized receive channel pairs forming at the same time two receive beams of identical mainlobe direction.

The hardware required to carry out the functions listed before consists of several subsystems described herein after. These systems are the Radiating Board, the RF Front-End (T/R Module Boards) the Beam Forming Network (BFN), the Digital Control Subsystem (DCS) and the Power Supply Subsystem (PSS). A schematic drawing of the subsystems indicating also electrical interfaces is shown in figure 3-1.

Figure 3-2 shows the bread board hardware. It is composed of the subsystems described above. Due to cost reasons the mechanical lay out is different from a space antenna design. Especially each T/R module board will be a small T/R module unit in a later development stage. However electrically the bread board is designed to demonstrate the beam agility and other important parameters, like H and V operation, for a later space system.

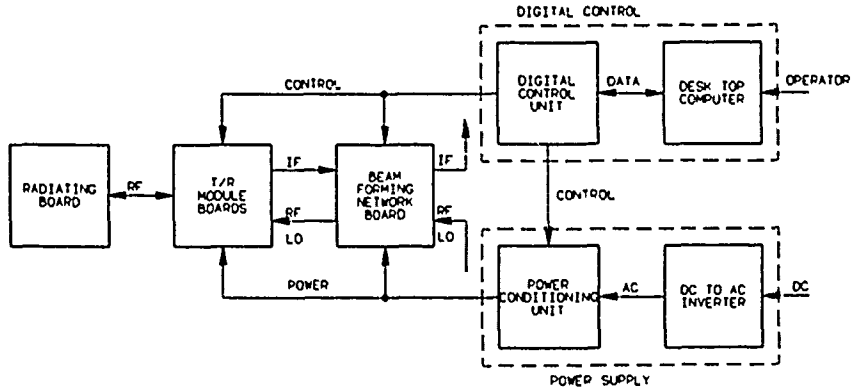


Fig. 3-1: Breadboard Subsystems

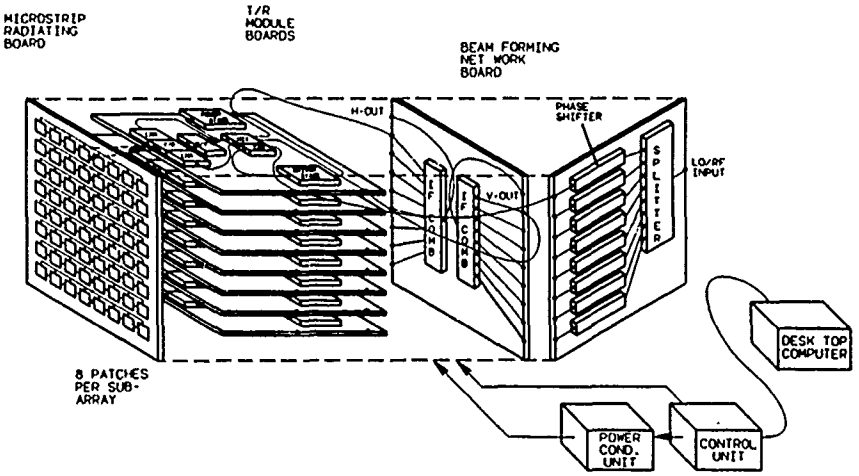


Fig. 3-2: Schematic of the bread board hardware

The radiating board subsystem performs the transmit and receive antenna functions. The radiating board is a planar array of dual polarized square microstrip patches printed on a low permittivity dielectric substrate. It consists of 8 horizontally arranged linear subarrays of 8 patches per subarray. The subarrays are fed by two sets (one for each polarization) of coaxial feed-throughs from the substrate rear side (ground plane) to the patch feeding microstrip lines on the front side. The radiating board is connected to the T/R modules by short coaxial cables with SMA connectors on both ends.

The RF Front End (T/R Module Boards) subsystem comprises the solide state driver and power amplifiers, the LNA's, the mixers, the horizontal/vertical polarization switches, limiters for LNA protection, and the transmit/receive duplexers. The power stages and part of the drivers stages will be operated with pulsed supply voltages/ currents. The total subsystem consists of 8 identical boards. The function of the boards can be seen from figure 3-3. The boards are connected to the radiating board and the BFN board by coaxial cables.

The basic functions of the BFN are signal distribution (RF/LO), signal combining (IF, H-pol. and V-pol.) and phase shifting (RF/LO). On both, the transmit and receive operation, the phase shifting is done by the same set of digital RF phase shifters, thus, minimizing their number required. On receive, the phase of the LO signals for the H/V mixer pairs is shifted before it is split and fed to the mixer LO ports. Therefore, the Rx H and V beams are steered simultaneously as required for the SAR system. However, because by this means no independent control of the H and V beam pointing is possible, the two channels of an Rx channel pair are subjected to tight phase matching requirements in order to get good pointing alignment.

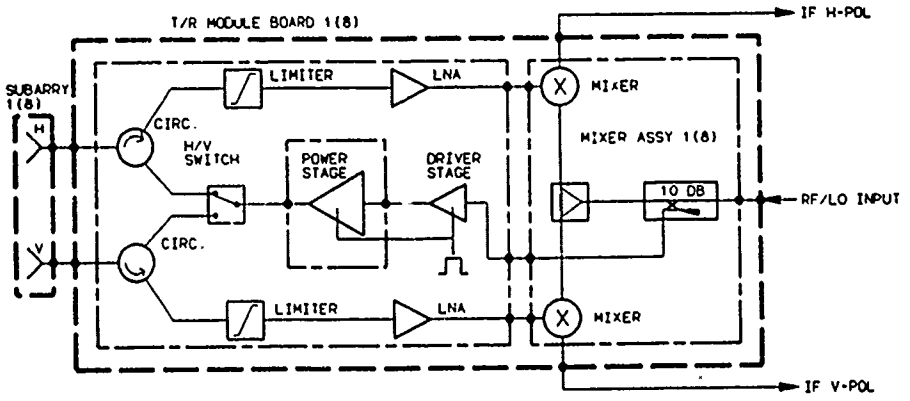


Fig. 3-3: T/R module board functional diagram

The Digital Control Subsystem (DCS) consists of two parts, the Digital Control Unit (DCU) and a Personal Computer (PC). The PC software presents to the operator a window-oriented operating menu from which he selects the tasks to be performed by the B/B, e.g. one of the Tx or Rx operating modes. The PC supervises the DCU which is the PC's interface to the B/B hardware. The DCU comprises look-up table memory for phase shifter data for 256 antenna beam directions. The pulse repetition frequency (PRF) and the control signals for driver and power stage gating and others are generated by the DCU. The high speed data transfer to the phase shifters, which requires fast read-out of the lock-up tables, is the major task of the DCU. The computation of the phase shifter data required for a certain beam steering angle under consideration of the phase correction data (known from calibration) will be performed by a dedicated high level language software package running on the PC.

The Power Supply Subsystem (PSS) consists of the AC fed Power Conditioning Unit (PCU) and the DC to AC inverter. This concept is very similar to the space antenna power distribution concept, where the power distribution to the numerous panels is done on high voltage AC level with the AC voltage generated from the space craft DC voltage by a DC to AC inverter.

The PCU provides DC power for all active components on the T/R module boards and on the BFN board.

Part of the RF power stage supplies, the power modulators, are located inside the power stage housings in order to have short distances for the high current pulses needed there. The control signal for power stage current switching is fed directly from the Digital Control Unit to the power modulators. The Digital Control Subsystem has its own build-in power supplies.

4. Bread Board Subsystem Performance

The following gives a summary of bread board performance data for the radiating board, the pulsed power amplifier and the receiver chain.

The radiating board comprises 8 linear subarrays of 8 dual feed square microstrip patches. Some performance data are given below.

Center frequency		5.300 GHz
Frequency bandwidth		50 MHz
Gain H-pol, V-pol		25 dB
Cross polarization at boresight max.	H-pol	-26 dB
	V-pol	-28 dB
VSWR max. across bandwidth	H-pol	1.5
	V-pol	1.8
Peak gain variation across bandwidth	H-pol	<0.5 dB
	V-pol	<0.2 dB
Losses, ohmic, mismatch		<1 dB

The amplifier chain in the T/R module transmit path consists of 5 GaAs-FET stages whereof the last 4 stages are operated with pulsed drain currents. The current switches and pulse energy storage capacitors are integrated in the power stage housing. Some of the main characteristics of the amplifier chain are listed below.

Pulse rep. frequ.	1024 Hz
RF pulse length	50 µs
Output power	40 dBm
Gain	40 dB
In pulse gain variation	< 0.9 dB
In pulse phase variation	< 5 deg
Unit to unit gain tracking	± 0.5 dB
Efficiency (last 2 stages)	>25 %

Each of the two parallel receive channels is constituted of a diode limiter and a two stage LNA. Some key data are:

Gain	19 dB
Noise figure	4.7 dB
Unit to unit gain tracking	±0.5 dB
Gain diff. H/V	<0.9 dB
Phase diff. H/V	<9 deg

5. Conclusion

The present activities to define active antennas for space-borne applications show promising results. The bread board activities show that most of the envisaged performance data can be met. However, the total mass and power consumption figures require further detailed investigations. A reduction of these figures is necessary to make space operations feasible for large active phased array SAR antennas.

There are several ideas to reduce mass and power which have to be traded against each other in terms of performance, cost and feasibility for space applications.

PERFORMANCE OF A PROPOSED SYNTHETIC APERTURE RADAR WITH VARIABLE ANTENNA HEIGHT

K. Tomiyasu

General Electric Company
P.O. Box 8048
Philadelphia, PA 19101 U.S.A.

Abstract

A variable antenna area for a spaceborne synthetic aperture radar (SAR) offers the possibility of broadside imaging with high area rates over a wide range of grazing angles. The antenna area must exceed a minimum value to avoid ambiguous responses in range and Doppler. The rectangular antenna length is fixed by the azimuth resolution. The antenna height is initially determined to fulfill the ambiguity constraint and then possibly increased to meet the SAR power constraint. A simplified active-element phased array with discrete heights is suggested and analyzed for SAR performance. An imaging rate up to 651 km^2/sec has been calculated for a 5.3 GHz SAR in an 800-km altitude orbit and 10° grazing angle.

Key words - synthetic aperture radar, satelliteborne SAR, imaging

Introduction

The antenna area of a spaceborne synthetic aperture radar (SAR) must exceed a minimum value to avoid responses ambiguous in range and Doppler for the particular geometry and radar wavelength. (Tomiyasu, 1978, Elachi, 1982) For the farthest range and smallest grazing angle at the scene, a large area and hence narrow beamwidth is required to fulfill the ambiguity constraint. The beam footprint size traversing at the orbital ground speed will yield the SAR imaging rate. If this initial beam is slewed towards nadir with a higher grazing angle and shorter range, the beam footprint size would be significantly smaller that at the farthest range. For this high grazing angle geometry, a significantly smaller antenna could be used to meet the ambiguity constraint. A smaller antenna area would result in a larger beam footprint and restore much of the loss in imaging rate. With a different approach a wide swath can be imaged using a narrow antenna beam scanned in elevation angle. (Tomiyasu, 1981, Moore, 1981)

Typically with a spaceborne platform there is a limit on the amount of power available to operate a SAR. With an ambiguity constrained antenna if the required SAR power exceeds that available, the power can be reduced by increasing the antenna area. Further, the terrain scene scattering coefficient is generally weak at small grazing angles and becomes significantly stronger at large grazing angles. Thus, for a number of reasons a spaceborne SAR with variable antenna height offers advantages that warrant a brief analysis of its performance.

Ambiguity Limited Antenna Area

To avoid ambiguous responses in range and Doppler the antenna area must exceed the ambiguity value given approximately by: (Tomiyasu,1978)

$$A_{amb} = \frac{8 v_{sc} \lambda R}{c \tan g}$$

where v_{sc} = spacecraft velocity
λ = wavelength
R = radar slant range
c = light velocity
g = grazing angle at scene

The minimum antenna area, A_{amb}, as a function of g is plotted in Figure 1 for a 5.3 GHz SAR on a satellite in a 800-km altitude circular orbit. For this example v_{sc} is 7450 m/s and λ is 0.0566 m. A very large spread in A_{amb} is noted.

Terrain Scattering Coefficient

The scattering coefficient σ^o of numerous scenes for different wavelengths and polarization have been measured and reported. Although there are large variations in the coefficient, the following relationship has been arbitrarily selected for this analysis:

$$\sigma^o = 0.03 \tan g$$

This relationship is also shown in Figure 1.

Azimuth Resolution

With a broadside imaging geometry the azimuth resolution for a single azimuth "look" image is approximately one-half of the in-track or azimuth dimension L_A of a rectangular aperture antenna. For multiple N-look imaging, the azimuth resolution δ_{az} is given by:

$$\delta_{az} = N L_A/2$$

In this analysis the azimuth resolution is assumed constant for all grazing angles.

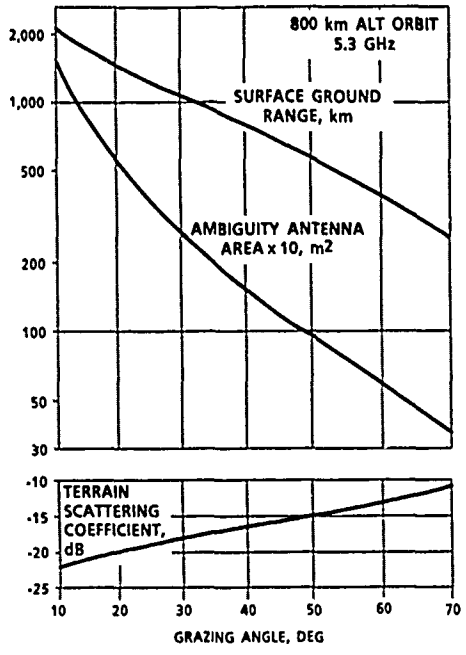


Figure 1

Antenna Height

For the variable area antenna it is assumed that only the antenna dimension in the elevation direction L_E will be varied so that

$$A_{ant} = L_A L_E.$$

Scene Ground Range

The distance from satellite nadir to the imaging scene along the earth's surface as a function of grazing angle is plotted in Figure 1. The satellite height above earth is 800 km.

SAR Transmitter Power

A very important constraint on spaceborne SAR is the amount of power required by the transmitter. As stated earlier the required power level can be reduced by increasing the antenna area which reduces the beam footprint size. The transmitter average power P_{ave} is given by (Tomiyasu, 1981):

$$P_{ave} = SNR \frac{\pi c^2 k T_S L_S}{8 \eta^2 v_{sc}} \frac{R \sin^2 g}{\sqrt{N}\, \delta_{sr} \cos g\, \sigma^o} \left(\frac{A_{amb}}{A_{ant}}\right)^2$$

where SNR = signal-to-noise ratio at beam center
k = Boltzmann's constant
T_S = system noise temperature
L_S = two-way system loss, greater than unity
η = antenna efficieny
δ_{sr} = slant range resolution

Constant Power Imaging Rate

In this analysis a three-look azimuth resolution of 30 m was chosen so that L_A = 20 m. For a satellite in circular orbit at 800 km altitude the satellite velocity is 7.45 km/s and the beam footprint velocity is 6.62 km/s. For simplicity earth rotation effects are neglected. The range swath width W is determined from the elevation geometry and is approximately

$$W = R\, \theta_E / \sin g$$

where θ_E = elevation beam width, in radians.

The imaging rate IR is given by

$$IR = W \times 6.62,\ km^2/sec.$$

In calculating the required transmitter average power the following assumptions were made:

f	5.3 GHz
SNR	10 dB at beam center, 4 dB at beam edge
T_S	450 K system noise temperature
L_S	6 dB two-way system loss
N	3 looks
δ_{sr}	7 m slant range resolution

In the computation, the ambiguity antenna areas were first determined as a function of grazing angle. The antenna area (antenna height) was then increased to reduce the power level to 50 watts of average power. A different antenna height was obtained for each grazing angle. With these antenna heights the range swath widths were calculated for grazing angles from 10° to 70°. The imaging area rates vary from 651 km^2/s at 10° to 1348 km^2/s at 70° with a minimum of 562 km^2/s at 20°. This is plotted in Figure 2.

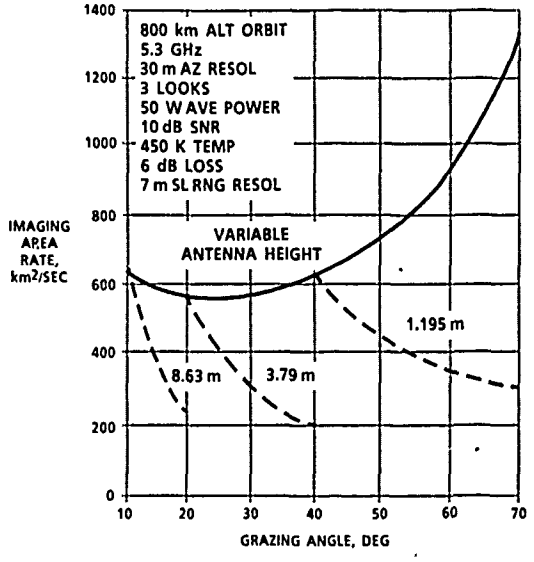


Figure 2

Variable Height Phased Array Parameters

For this type of SAR it is assumed that a phased array with active elements will be employed. If the elements on the aperture are located on a 0.7λ square grid which would permit a scan angle of 25° from array surface normal, the average power per element assuming uniform aperture illumination will range from 0.45 to 14.26 milliwatts. These power levels appear to be reasonable values for active solid-state elements. These levels as well as the number of elements required in the array for each grazing angle are plotted in Figure 3.

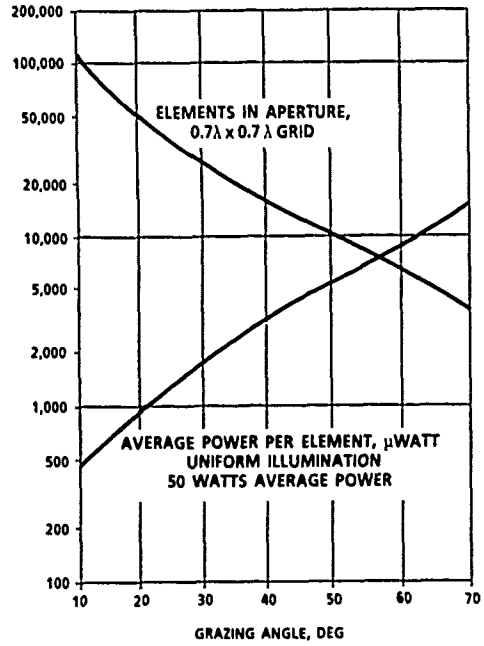


Figure 3

Simplified Array Design

In order to simplify the design of the active array, three specific antenna heights were chosen to image over the range of grazing angles. For the smallest grazing angles of 10 to 20 degrees an antenna height of 8.63 m is chosen which provides an imaging rate from 651 to 247 km^2/s. This is shown by the dashed lines in Figure 2. If the average transmitter power is fixed at 50 watts, the signal-to-noise ratio at beam center varies from 10 to 17.1 dB as the grazing angle is changed from 10° to 20°. At the 20° grazing angle an optional antenna height of 3.79 m can be used and the imaging rate can be restored to 562 km^2/s with a SNR of 10 dB at beam center. This antenna height can be used up to 40° grazing angle where the height can be decreased to 1.195 m to restore the imaging rate. The latter height can be used to the limit of 70° grazing angle.

Discrete Height Phased Array Parameters

The density of active elements in the phased array is governed by the appearance of grating lobes when the beam is steered from array surface normal. If the elements are placed on a square grid of 0.7 wavelengths the beam can be steered up to 25° from normal. The number of elements for the 8.63, 3.79, and 1.195 m height antennas are 109,954, 48,288 and 15,225 elements respectively. The transmitter average power per element assuming uniform aperture illumination are 0.451, 1.03, and 3.28 milliwatts respectively. These element power levels appear to be reasonable for solid-state devices at 5.3 GHz.

Conclusion

For imaging with a broadside-viewing spaceborne synthetic aperture radar over a wide range of grazing angles from 10° to 70°, the ambiguity limited minimum antenna area varies by a factor in excess of 30. The goal of high rates of imaging area requires the antenna size to be near the ambiguity limit. However, to meet a SAR power constraint it may be necessary to increase the antenna area thereby sacrificing imaging area rate. The antenna length in the azimuth direction is fixed by the azimuth resolution. To simplify the antenna design three antenna heights are suggested to cover the 10° to 70° grazing angle range. With this approach, a 50-watt 5.3 GHz SAR borne on a satellite in an 800-km altitude orbit can generate 30-m azimuth resolution terrain images at rates of 247 to 651 km^2/sec.

Bibliography

1. Tomiyasu, K., "Tutorial review of synthetic aperture radar (SAR) with application to imaging the ocean surface," Proc. IEEE, vol. 66, pp. 563-583, May 1978.

2. Elachi, C., T. Bicknell, R.L. Jordan, and C.Wu, "Spaceborne synthetic aperture imaging radars: application, techniques and technology", Proc. IEEE, vol. 70, pp. 1174-1209, October 1982.

3. Tomiyasu, K., "Conceptual performance of a satellite borne, wide swath synthetic aperture radar", IEEE Trans. Geosci. Remote Sensing, vol. GE-19, pp. 108-116, April 1981.

4. Moore, R.K., J.P. Claassen, and Y.H. Lin, "Scanning spaceborne synthetic aperture radar with integrated radiometer", IEEE Trans. Aerosp. Electron. Syst., vol. AES-17, pp. 410-421, May 1981.

DESIGN OF A DUAL POLARIZED MICROSTRIP PATCH ARRAY FOR AN ADVANCED SAR ANTENNA

C.A. MCDONACH, D. NGUYEN, G.S. GUPTA, A.P. LUSCOMBE

SPAR AEROSPACE LIMITED
21025 TRANS CANADA HIGHWAY
STE ANNE DE BELLEVUE
QUEBEC H9X 3R2

ABSTRACT

Future Synthetic Aperture Radar (SAR) systems in space are required to be capable of imaging in several modes calling for antennas which have sophisticated beamforming and rapid beamswitching capabilities. The instrument concept for such an Advanced SAR, implemented in microstrip, is discussed. The electrical design of a dual polarized microstrip patch array, suitable for this application, is presented. The array is designed for a uniform illumination giving a narrow pencil beam in azimuth and a broad elevation coverage. The measured results of a sixteen element prototype array are presented and compared to theoretical predictions. (Keywords: microstrip, synthetic aperture).

1.0 INTRODUCTION

The antenna design described in this paper is intended for use in an advanced Synthetic Aperture Radar (SAR) system to be flown on a European Space Agency (ESA) satellite in the latter half of the next decade. This system is required to provide dual polarization images of the earth's surface at a range of different incidence angles and to be capable of limited spotlight operation [1]. The SAR antenna is approximately 20m in length operating at 5.3 GHz providing vertical and horizontal polarizations for both transmit and receive modes of operation. The current work covers the technology definition phase for the advanced SAR antenna and is aimed at providing 16 breadboard linear arrays (each with uniform illumination in the azimuth direction) which may be stacked to form a planar array representative of a future SAR panel.

2.0 INSTRUMENT CONCEPT AND SYSTEM DESIGN

As mentioned above, the proposed antenna design is intended for use in ESA's next generation of Synthetic Aperture Radars. The SAR on ESA's first Remote-Sensing Satellite, ERS-1, which is due for launch in 1990, will operate with an antenna that produces a single horizontally-polarized beam pattern pointed at a fixed angle in the plane broadside to the satellite. This beam enables the system to image a swath approximately 100 km in width at an incidence angle of around 23° to the right of the flight path. The Advanced SAR system is intended to be significantly more flexible in its operations, and to have enhanced capabilities in terms of coverage, resolution and polarization. Each of these capabilities is translated into specific requirements which have had to be addressed in the antenna design.

The Advanced SAR is required to be capable of imaging in several different modes. At one extreme is imaging of a very wide swath (nominally 500 km) at a moderate resolution. At the other, fine resolution (10m or better) imaging of a swath of limited width. Additionally, the system must be able to image regions at any of a range of different incidence angles to the side of the satellite. These various requirements can all be satisfied if the antenna is designed to have elevation beamforming and rapid beamswitching capabilities. The various swaths require elevation beamwidths which range in width from 16° down to 3°.

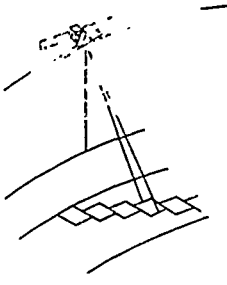


FIGURE 1A ADVANCED SAR ELEVATION BEAM AGILITY

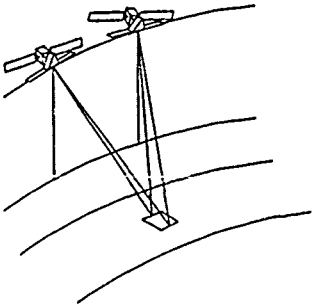


FIGURE 1B ADVANCED SAR 'SPOTLIGHT' MODE

Figure 1a shows a schematic representation of imaging of a series of different regions using this elevation beam agility. Provided that the swath width required for a given imaging period is not much greater than about 100 km, a single beam of the necessary width and direction can be formed. When the requirement is for coverage of a wider swath, the ScanSAR technique, involving rapid switching between beams covering adjacent swaths, can be used.

In conventional SAR operations, the antenna beam is pointed continuously broadside to the flight path. The total period of signal that can be processed coherently to image any given point on the ground is limited to the period that the point is in the beam. The optimum resolution that can be achieved in the along-track (azimuth) direction is then approximately half the antenna length. A 0.15° azimuth beamwidth, for example, will allow a resolution of about 10m to be obtained. In order to obtain even finer resolution in this dimension of the image, the Advanced SAR is required to be capable of limited 'Spotlight' operation. This mode of operation involves a backward rotation of the beam which allows the area of interest to be kept illuminated for a longer period (see Figure 1b). This extended period of signal can then be processed coherently to provide the required finer resolution. In terms of the antenna, this translates into a requirement for a beamforming capability in the second dimension, in azimuth.

The system is required to produce images simultaneously in co- and cross-polarizations. To achieve this, it must be able to transmit with either horizontally or vertically polarized waves, and to receive simultaneously in both polarizations. These convert directly into equivalent requirements for the antenna. Although the antenna design described in the following sections is specifically for the one system, the technology and approach are equally applicable to other advanced SAR systems. One additional capability that is of particular interest is quad-polarization imaging (i.e. with all four combinations of linear polarization on transmit and receive). This form of imaging requires changes in the system's operating parameters, but could be provided with the same basic antenna technology and design.

3.0 ANTENNA DESIGN

Overall SAR Antenna:

As shown in Figure 2 the antenna is 20 m in length and is made up of ten identical panels each 2m x 2m in size. Each of these panels consists of a RF radiating surface of microstrip patches connected to a number of transmit/receive (T/R) modules at the rear. A mechanical support structure houses the T/R modules and also serves to fix the position of the radiating surface relative to all other panels. In order to reduce design complexity the feed lines are etched on the same surface as the microstrip patches, thus eliminating the need for an additional substrate layer.

The spacing between elements is chosen as approximately 0.7 wavelengths resulting in a 50 x 500 element array for the complete antenna and a 50 x 50 element subarray for each panel. The proposed design concept maintains a uniform amplitude distribution over the complete antenna.

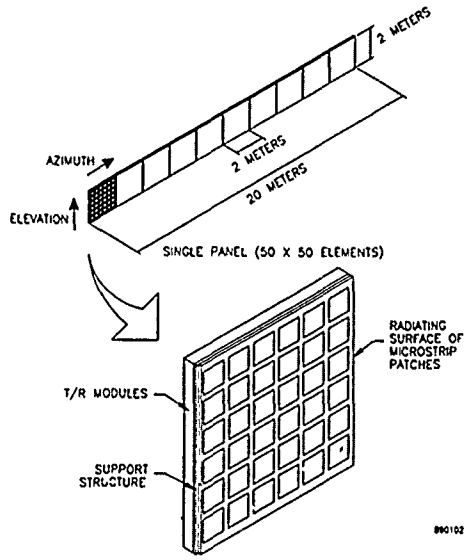


FIGURE 2 ADVANCED SAR ANTENNA CONCEPT

Consequently the required beamshaping in elevation is achieved by phase only synthesis controlled by the 50 T/R module elements in that plane. Similarly in the azimuth direction beamscanning of the narrow pencil beam may be achieved by controlling the phase of each panel resulting in ten effective azimuth elements. This arrangement assumes one T/R module per row. However finer control of beamscanning may be achieved by increasing the number of T/R modules in each row. The main antenna RF requirements are summarized in Table 1.

TABLE 1 SUMMARY OF ELECTRICAL PERFORMANCE REQUIREMENTS

Operating Frequency	5.3 GHz
Bandwidth	52 MHz
Polarization	Dual Linear
Aperture Illumination	Uniform Amplitude and Phase
Element Spacing	0.7 Wavelengths
VSWR	<1.3:1
Loss	<2 dB
Cross-Polarization Level	<20 dB
Gain /Frequency Response	<0.5 dB

Breadboard Antenna:

In demonstrating the proposed SAR concept the current work was aimed at providing a limited breadboard antenna, consisting of a 16 x 16 planar array, which forms a representative portion of a full 2m panel satisfying the requirements outlined in Table 1. The breadboard array is made up of 16 identical linear microstrip arrays, each with 16 elements in azimuth, and is implemented in RT DUROID 5870 with a substrate thickness of .062 inches.

Microstrip Patch:

The microstrip patch offers a low profile lightweight option in the choice of radiator for many array applications. As shown in Figure 3, the patches consist of a planar conducting geometry etched onto one side of a dielectric substrate and backed by a ground plane on the other. They are usually around a halfwavelength in size and radiate by virtue of the fringing fields at the edges. In this application where the array is required to provide identical performance for two orthogonal polarizations the square patch is chosen as the radiating element for the SAR antenna. The patch is fed from the edge at the point of highest impedance (~350 ohms) and may be represented by an equivalent RLC circuit valid near resonance [2].

The resonant frequency of the patch varies with the patch dimensions and may be calculated within an accuracy of a few percent [3][4][5]. In practice a number of patch samples with slightly different dimensions around a computed value were fabricated.

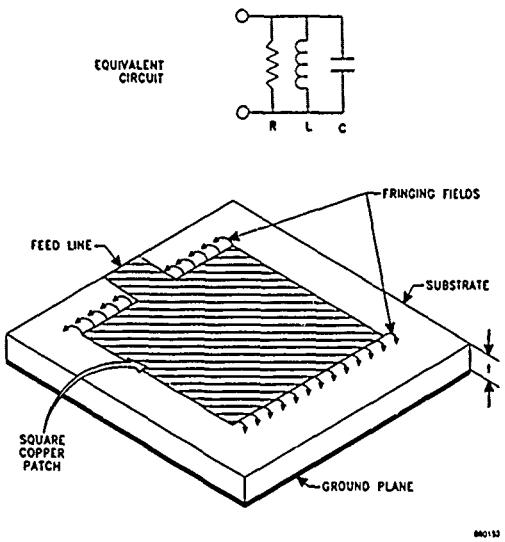


FIGURE 3 MICROSTRIP PATCH RADIATOR

Feed Network:

In designing the single layer feed network the main considerations were bandwidth, feed line radiation and layout efficiency. The bandwidth is determined by design parameters such as substrate thickness, and dielectric constant. However extraneous radiation from the feed lines may cause some distortion of the radiation pattern and should be minimized by judicious choice of the above parameters at the outset. To avoid feed line radiation and generation of surface waves high impedance lines were used and the substrate thickness was kept below $0.03\lambda_o$. Similarly layout in the vertical plane was a major constraint. Due to scanning and grating lobe considerations the element spacing was specified as 0.7 wavelengths (40 mm) resulting in a 50 element elevation aperture. Consequently both VP and HP networks for each horizontal subarray were confined to this 40 mm limit. This is a stringent requirement when one considers that the patch size itself accounts for approximately 50% of the available space ($\sim.5\lambda_o/\epsilon_r$).

The layout chosen to satisfy the above requirements was that of the series feed shown schematically in Figure 4. This design is a linear structure consisting only of a main feed line from which power is tapped off and delivered to the patches periodically. The series feed is a 16 element resonant array fed from the centre to give improved VSWR bandwidth. To avoid coupling between feed lines a minimum edge to edge separation of three ground plane spacings was maintained throughout. As can be seen from the diagram the elements are fed from points which are λ_g apart on the main feed line so that all patches are fed in phase. Similarly the input impedance at any patch is Z/16 so that all elements have equal amplitude.

In the HP network there is a 180° offset at the input in order to compensate the space phase reversal which exists at the patches on the left hand side of the array compared to those of the right hand side.

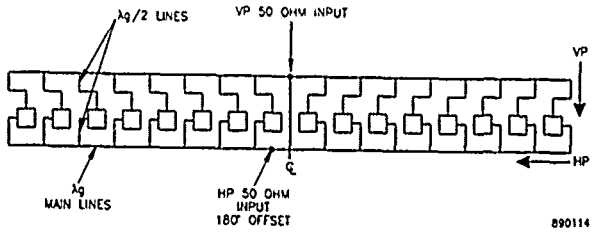


FIGURE 4 SERIES FEED LAYOUT

The series feed was modelled using the SUPERCOMPACT software package which is commercially available and is used as an industry standard in printed circuit design. A circuit model was constructed using a node description from which the s-parameters were obtained. This model was analyzed using a number of equivalent circuits for microstrip discontinuities such as bends, steps, and tee junctions. Since the array design is a resonant structure good electrical performance can only be achieved by accurate characterization of the above discontinuities. Reactive contributions in such discontinuities may limit the bandwidth and cause a shift in the resonant frequency if not corrected. Consequently, a number of test samples were made to verify the accuracy of the equivalent circuits. It was found that the tee-junction and step models were very accurate. However, in the case of the 90° bend it was necessary to use a measured value for insertion phase in the design of the feed. Having constructed the computer model the performance of the entire array was readily assessed. Figure 5 and Figure 6 show the predicted VSWR and mid band far field pattern predictions for the 16 element series feed array.

A prototype 16 element dual polarized linear array was fabricated and found to meet the requirements outlined in Table 1. Typical measured radiation patterns are shown in Figure 7 and Figure 8. As can be seen from the plot of Figure 7 the azimuth pattern is a narrow pencil beam with a 4.3 deg half power beamwidth as expected for a 16 element array.

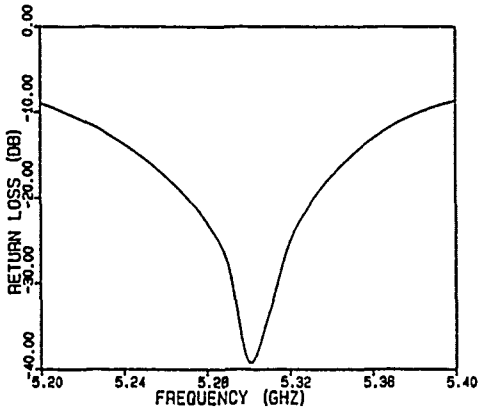


FIGURE 5 PREDICTED VSWR PERFORMANCE

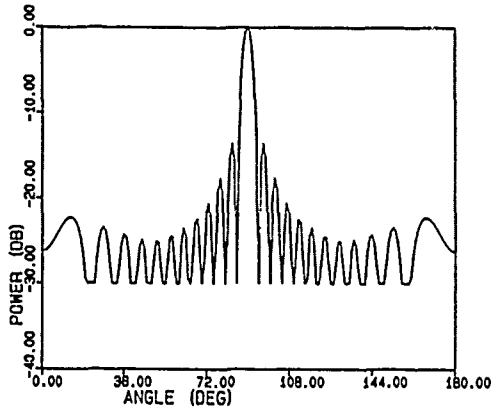


FIGURE 6 PREDICTED FAR FIELD RADIATION PATTERN

The pattern follows a sin x/x profile which agrees well with theory and the nulls are sharp indicating that the required equi-phase distribution across the array has been achieved. The elevation pattern of Figure 8, on the other hand, provides broad beam coverage and is almost identical to the pattern of a single microstrip patch as expected.

Conclusion

Future SAR systems will be significantly more sophisticated and flexible in their operations calling for electrically scanned antennas. An active array architecture implemented in microstrip offers a low profile light weight solution for advanced SAR systems.

In particular a dual polarized array with a square microstrip patch as radiating element has been designed to meet these needs. In designing the array the reactive components of microstrip discontinuities must be compensated to ensure good performance.

A prototype sixteen element array has been fabricated satisfying the anticipated requirements of an Advanced SAR system. Thus microstrip has been demonstrated to be a suitable technology for future SAR arrays.

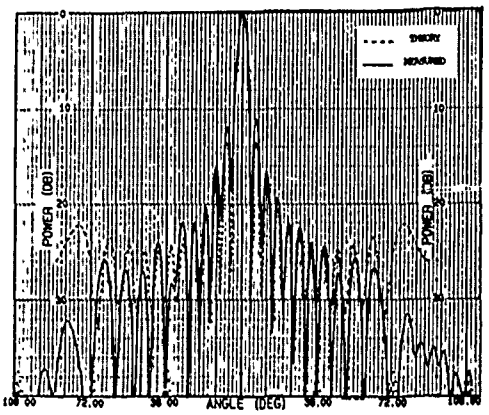


FIGURE 7 MEASURED AZIMUTH PATTERN

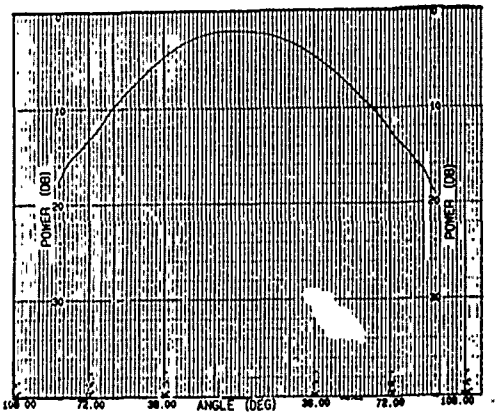


FIGURE 8 MEASURED ELEVATION PATTERN

REFERENCES

[1] ITT for Technology Definition of an Advanced SAR ESA Document AO/1-2010/87/NL/JG

[2] Richards WF, bo YT, Simon P, and Harrison DD 'Theory and Application for Microstrip Antennas' Proc. Workshop printed circuit Antennas Tech. New Mexico State University, Las Cruces, Oct. 1979

[3] Carver KR 'Practical Analytical Techniques for the Microstrip Antennas' Proc Workshop on Printed Circuit Antennas' New Mexico State University, Oct. 1979

[4] Coffey EL 'A new analysis technique for Calculating Self and Mutual Impedance of Microstrip Antennas' proc workshop on Printed Circuit Antennas New Mexico State University, Jan. 1979

[5] Carver KR and Coffey EL 'Theoretical Investigation of Microstrip Antennas' Tech. Rep PT 00929 Physical Science Lab, New Mexico State University, Las Cruces, Jan 1979

A generalized method for evaluation of the mutual coupling effect on the radiation and scanning characteristics of microstrip array antenna.

Abdul Amir M.Ali and Babau R.Vishvakarma.

Department of Electronics Engineering, Institute of Technology,
Banaras Hindu University, Varanasi, India.

Abstract

A generalized analytical method for the evaluation of mutual coupling in microstrip array antenna is proposed in order to study the effect of mutual coupling on the radiation and scanning characteristics of the system. Based on cavity model the microstrip antennas are analysed using the dyadic Green's function to obtain the far field radiation. It is found that because of mutual coupling effects, the input impedance of a radiating antenna in an array is different from its impedance value when it is isolated, and this difference varies with the angle of scan of the array. Also it is found that the mutual coupling between the two patch elements decreases with increasing spacing between them.

I. Introduction

Microstrip antennas are ideally suited for arrays. Also the various characteristics such as high gain, beam scanning or steering capability are possible only when discrete radiators are combined to form arrays (Derneryd, 1976). In microstrip array antenna since the elements are close to each other, there is a mutual coupling between them. In practice the mutual coupling between the radiators alters the array performance significantly(Jadlicka, Poe and Carver, 1981). It becomes more important in case of electronically scanned array antennas.

In the present endeavour therefore, an attempt has been made to investigate the mutual coupling between rectangular elements of microstrip array antenna. Effect of mutual coupling on the radiation and scanning characteristics are also studied.

The mutual coupling between a pair of antenna(Fig.1) is studied in terms of the mutual impedances. The elements of the array are excited by microstrip or coaxial line (Bahl and Bhartia, 1982). Based on cavity model the microstrip antennas are analysed using the dyadic Green's function to obtain the far field radiation. An equation is developed for evaluating the mutual impedance.

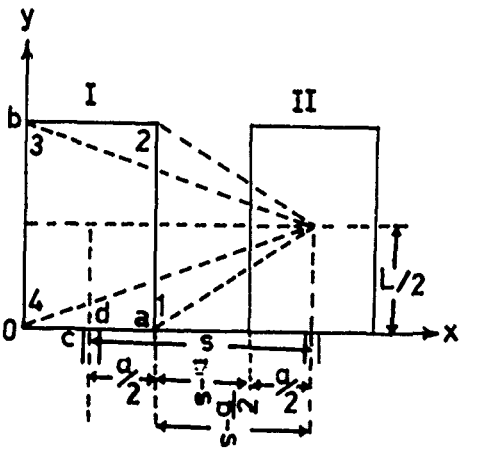


Figure 1 Configuration for coupling calculation between elements of array antenna.

From the E and H-Plane pattern it is observed that the coupling increases the radiated power level and improves the directivity and beam width.

II. Theoretical Considerations

The mutual impedance between a pair of antennas is defined as the ratio of the voltage induced at the terminals at one antenna, Say a (by the current in the other antenna), to the input current in the second antenna, Say b (Kraus, 1984)

$$Z_{12} = \frac{V_{12}}{I_2} = \frac{V_{ocb}}{I_b} \qquad \ldots\ldots(1)$$

where

$V_{12}=V_{ocb}$ is the open-circuit voltage at the terminals of antenna a which is induced by the field generated by antenna b.

We consider that the far field at point $\bar{R}$

is due to a harmonically oscillating current in free space that is given by (Tai,1971)

$$\bar{E}(\bar{R}) = Jw\left[1+\frac{1}{k^2}\nabla\nabla\right]\bar{A}(\bar{R}) \qquad .....(2)$$

where

$$\bar{A}(\bar{R}) = \mu_o\int_v Go(\bar{R}/\bar{R}')\,\bar{J}(\bar{R}')dv \qquad .....(3)$$

with

$$Go(\bar{R}/\bar{R}') = \frac{e^{jk/\bar{R}-\bar{R}'/}}{4\pi/\bar{R}-\bar{R}'/} \qquad .....(4)$$

The function $Go(\bar{R}/\bar{R}')$ is called the free space Green's function, where $\bar{R}'$ denotes the position vector of a source point and $\bar{R}$ denotes the position vector of a field point or that of an observer.

The field given in equation (2) depends upon the distribution of the source current which is considered to be an infinitesimal source located at $\bar{R}'$ with a current moment equal to $1/JW\mu_o$ in the X-direction which is given by

$$\bar{J}(\bar{R}) = \frac{1}{Jw\mu_o}\,\delta(\bar{R}-\bar{R}')\,\hat{x} \qquad .....(5)$$

Now V_{ocb} will be obtained in terms of transmitting parameters. The use will be made of the Reciprocity theorems.

If the transmitting antenna b is a point dipole then the reciprocity theorem gives

$$V_{oc} = \int \bar{E}_b(\bar{R}).\bar{J}_a(\bar{R})\,dv = \int \bar{E}_a(\bar{R}).\bar{J}_b(\bar{R})\,dv = \bar{E}_a(\bar{R})\,\hat{P} \qquad .....(6)$$

Where $\hat{P}$ is the unit vector in the direction of the transmitting point dipole b.

Equation (6) is very difficult to evaluate because it depends on the total field $\bar{E}_b(\bar{R})$ which results when the incident field interacts with metal and dielectric structure of the antenna. To determine this field, it is necessary to solve a complex scattering problem.

An alternative form is obtained by considering that when $\bar{J}_a(\bar{R})$ is connected to the antenna terminals, currents flow on the antenna structure. We can regard these later currents as induced currents and refer to them as $\bar{J}_i(\bar{R})$. The source current $\bar{J}_a(\bar{R})$ and the antenna currents $\bar{J}_i(\bar{R})$ form some current distribution in space. Therefore equation(6) can be rewritten as

$$V_{oc} = \int \bar{E}_b(\bar{R}).(\bar{J}_a(\bar{R})+\bar{J}_i(\bar{R}))\,dv \qquad .....(7)$$

If the current at the terminals is I_{in} instead of unity, then the open circuit voltage is

$$V_{oc} = \frac{1}{I_{in}}\int \bar{E}_{inc}(\bar{R}).\bar{J}_t(\bar{R})\,dv \qquad .....(8)$$

Where $E_{inc}(\bar{R})$ is the field incident (in the absence of the antenna structure), $\bar{J}_t(\bar{R})$ is the current distribution of the antenna when it is acting as a transmitter, and I_{in} is the input current associated with the transmitting current distribution.

Hence, the mutual impedance between a pair of antenna is given by the formula.

$$-Z_{12} = \frac{1}{I_{in1}\,I_{in2}}\int \bar{E}_{inc,2}(\bar{R}).\bar{J}_{t1}(\bar{R})dv \qquad ...(9)$$

where

$$\bar{E}_{inc,2}(\bar{R}) = \bar{E}_{inc,1}(\bar{R})(Z_{12}/Z_{22}) \qquad .....(10)$$

$\bar{E}_{inc,2}(\bar{R})$ is the field generated by antenna 2 at the position of antenna 1(but in the absence of antenna 1), $\bar{J}_{t1}(\bar{R})$ is the transmitting current distribution on the first antenna(with the second antenna structure present but not operating), and $I_{in,1}$ and $I_{in,2}$ are the input currents to the antennas.

Using the cavity model(Derneryd,1979)with radiating edges extended slightly to account for the fringing field, the field in the cavity of a rectangular microstrip antenna must be of the form

$$E_z = \sum A_m \cos\frac{m\pi}{a}X \cos B_m(Y-b) \qquad .....(11)$$

where

$$B_m = \sqrt{K^2-(\tfrac{m\pi}{a})^2},\quad K = W\sqrt{\mu\epsilon_d}, \qquad .....(12)$$

$$A_m = \frac{j4w\mu}{m\pi}\,\frac{\sin\frac{m\pi}{2a}(d-c)\cos\frac{m\pi}{2a}(d+c)}{B_m \sin B_m b} \qquad .....(13)$$

and

$$m = 1,2 .... \qquad .....(14)$$

When only one element of an array is excited its field induces currents in other unexcited elements hence they radiate fields. Therefore,with the help of equation(2),(9)and (11) the radiation field of first antenna when second is present but not excited(Fig.2) can be written as :

$$\bar{E}_{12} = \bar{E}_1\left\{(1+(Z_{12}/Z_{22})\cos\Psi)^2+(Z_{12}/Z_{22}\sin\Psi)^2\right\}^{1/2} / \tan^{-1}\frac{Z_{12}\sin\Psi}{Z_{22}+Z_{12}\cos\Psi} \qquad .....(15)$$

where

$$\Psi = \frac{2\pi d}{\lambda}\cos\Psi + d$$

Similarly, field when only second antenna is excited with first is present, is given by

$$\bar{E}_{21} = \bar{E}_2\left\{((Z_{12}/Z_{11})+\cos\Psi)^2+\sin^2\Psi\right\}^{1/2} / \tan^{-1}\frac{Z_{11}\sin\Psi}{Z_{12}+Z_{11}\cos\Psi} \qquad .....(16)$$

if both the antennas are excited(Fig.3) then the resultant field will be

$$\bar{E}_R = \left\{(E_{12}+E_{21}\cos\Psi)^2+(E_{21}\sin\Psi)^2\right\}^{1/2} / \tan^{-1}\frac{E_{21}\sin\Psi}{E_{12}+E_{21}\cos\Psi} \qquad .....(17)$$

where Z_{11}, Z_{22} denote the self-impedances.

To simplify the integral of equation(9), following assumptions have been made

a) Current flowing on the outer surface of the antenna is negligible.
b) Variation of field in ϕ direction on edge of microstrip antenna is negligible.This

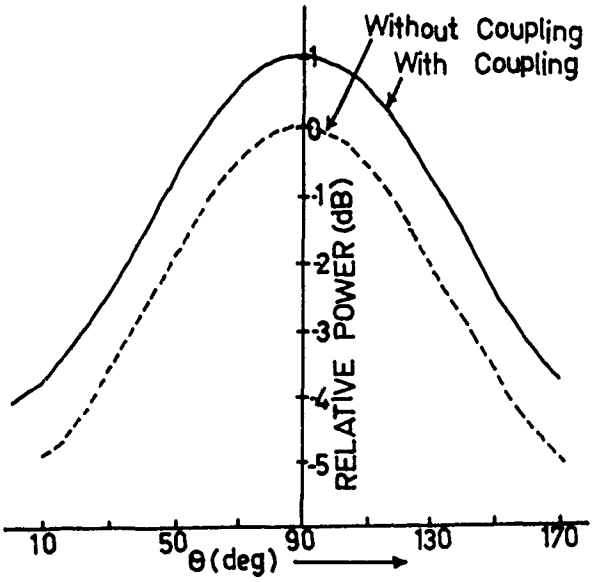


Figure 2 E-plane radiation pattern when only first element excited.

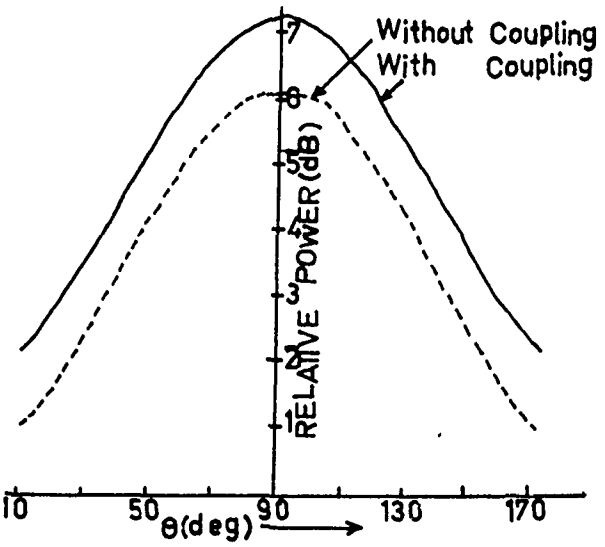


Figure 3 E-plane radiation pattern when both elements are excited.

is true because width of element is usually very small.

Then equation (9) reduces to

$$-Z_{12} = \int_{1,2,3,4} h\, \bar{E}_{inc,2}(\bar{R}) \cdot \bar{J}_{t1}(\bar{R})\, dl \quad \ldots(18)$$

Which is a line integral to be done on the whole boundary of the surface with h as the substrate thickness. Integration performed separately for all four boundary walls i.e.

$$-Z_{12} = \int_1^2 h\, \bar{E}_{inc,2}(\bar{R}) \cdot \bar{J}_{t1}(\bar{R})\, dl + \int_2^3 \longrightarrow + \int_3^4 \longrightarrow + \int_4^1 \bar{E}_{inc,2}(\bar{R}) \cdot \bar{J}_{t1}(\bar{R})\, dl \quad \ldots(19)$$

In this expression function to be integrated is function of r, θ, X and Z. However two variables are sufficient to define the function. Hence X and Z are found in terms of r and Q.

Because of the complexity of the function, numerical integration is used and Trapezoidal rule for numerical integration is applied.

The final expression for Z_{12} for m = 0 is given by

$$-Z_{12} = Y' \int_{\theta=\tan^{-1}\frac{-L/2}{S-W/2}}^{\theta=\tan^{-1}\frac{L/2}{S-W/2}} \left(\frac{\exp(-Jk_o(S-(W/2))\mathrm{Sec}\theta)}{(S-(W/2)).\mathrm{Sec}\theta} (H_s) \right) \left\{ W\sqrt{\mu\epsilon}d \left(\frac{L}{2} - (S+\frac{W}{2})\tan\theta \right) - L \right\} (S-\frac{W}{2})\mathrm{Sec}\theta \, d\theta +$$

$$\int_{\theta=\tan^{-1}\frac{-L/2}{S+W/2}}^{\theta=\tan^{-1}\frac{-L/2}{S-W/2}} \left(\frac{\exp(-Jk_o(S+\frac{W}{2} - \frac{L}{2}\mathrm{Cot}(-\theta))}{L/2\mathrm{Cosec}(-\theta)} \cdot (H_s) \right) \left\{ W\sqrt{\mu\epsilon}d \ (L-L) \right\} \frac{L}{2}\mathrm{Cosec}\,(-\theta)\,)\, d\theta +$$

$$\int_{\theta=\tan^{-1}\frac{-L/2}{S+W/2}}^{\theta=\tan^{-1}\frac{L/2}{S+W/2}} \left(\frac{\exp(-Jk_o(S+\frac{W}{2})\mathrm{Sec}\theta}{(S+W/2)\mathrm{Sec}\theta} \cdot (H_s) \right) \left\{ W\sqrt{\mu\epsilon}d \ (\frac{L}{2} - (S-\frac{W}{2})\tan\theta - L) \right\} (S+\frac{W}{2})\mathrm{Sec}\theta) \, d\theta +$$

$$\int_{\theta=\tan^{-1}\frac{-L/2}{S-W/2}}^{\theta=\tan^{-1}\frac{L/2}{S+W/2}} \left(\frac{\exp(-Jk_o(\frac{L}{2}\mathrm{Cosec}\theta))}{L/2\ \mathrm{Cosec}\ \theta} \cdot (H_s) \right) \left\{ W\sqrt{\mu\epsilon}d \ \mathrm{Cos}(-L) \right\} \frac{L}{2}\mathrm{Cosec}\ \theta)\, d\theta \quad \ldots(20)$$

where

$$Y' = 2h(J2V_oWk_o)\left(\frac{1}{4\pi}\right) \cdot J(d-c)\sqrt{\mu} / (W - \mathrm{Sin}\, k_oL\sqrt{\epsilon d}) \quad \ldots(21)$$

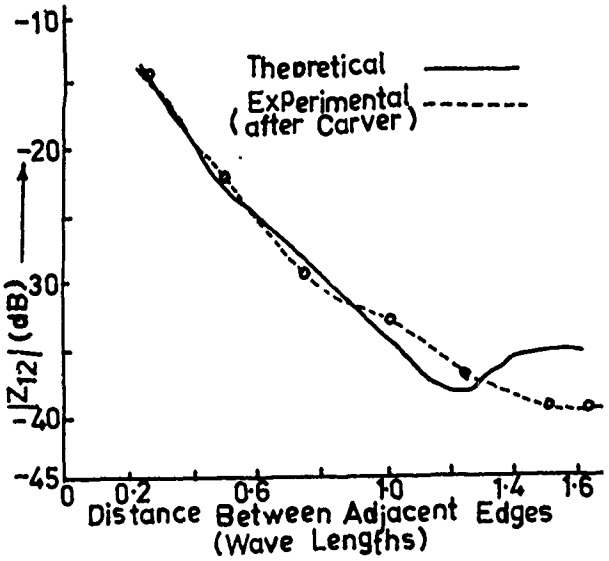


Figure 4 $/Z_{12}/$ values for rectangular patch of size 5 cm.(radiating edge)x6 cm. at frequency 1.6 GHz.

and $$H_s = \frac{\sin(\frac{Koh}{2}\sin\theta)}{\frac{Koh}{2}\sin\theta} \cdot \frac{\sin(\frac{Kow}{2}\cos\theta}{\frac{Kow}{2}\cos\theta} \cdot \sin\theta\cos\theta$$

$$(K_o \frac{L}{2}) \cos \qquad .....(22)$$

III. Discussion of Results

Using equation (15),(16) and (17) both E-plane and H-plane radiation patterns were calculated when both antennas were excited, antenna 1 is excited with the presence of antenna II, and antenna II is excited with antenna I is present. The perusal of the E-plane patterns (Fig.2-3), indicate that the presence of unexcited element in the proximity of the excited element causes considerable coupling between them which increases the relative radiated power by 1 dB as compared to the case when there is no coupling. The resultant radiation beam is found to be tilted toward the parasitic element. Therefore, coupling increases the radiated power and improves the beam width. Similar results were also obtained for H-plane patterns. For the case when both antenna elements are excited, coupling is observed to improve the radiation level by 2.5dB. The coupling thus, increases the radiated power level and improves the directivity and beam width. Also it is found that the mutual coupling does have the effects on the input impedance of the element and the radiation pattern of an element in the array. Because of mutual coupling effects, the input impedance of a radiating antenna in the array is different from its impedance value when it is isolated, and this difference varies with the angle of scan of the array. However the edge effects were ignored at this stage.

Equation (20) is developed for evaluating the mutual impedance. The equation developed for Z_{12} is quite complex and involves four integrals. The calculations for the mutual coupling between the antenna elements of typical dimensions(5 cm x 6 cm) on alumina substrate reveal that the mutual coupling between the two patch elements decreases with increasing spacing between them. This is in accordance with the fact that increasing spacing between the elements reduces the inter-element interaction resulting in weak coupling. The calculated values (Fig.4) of Z_{12} are similar to the experimental data as reported by Carver. There is a very small deviations in Z_{12} for higher spacing between the elements. This, therefore, varifies the validity of the developed method for mutual coupling in case of rectangular patches.

REFERENCES

1. Derneryd, A.G., "Microstrip Array Antenna", Proc. 6th European Microwave Conference, PP 339-343, 1976.

2. Jedlicka, R.P., Poe, M.T., and Carver, K.R., "Measured Mutual Coupling between microstrip antennas", IEEE Trans. Antennas Propagat, vol. Ap-29, PP 147-149, Jan. 1981.

3. Bahl, I.J., and Bhartia, P., "Microstrip Antennas", ARTECH House, Inc., PP 23-24, 1982.

4. Derneryd, A.G., "Analysis of the microstrip disc antenna element". IEEE Trans. Antennas Propagat., vol. Ap-27, PP 660-664, Sept. 1979.

5. Tai, C.T., "Dyadic Green's Functions in Electromagnetic Theory", Scranton, PA : Educational, 1971.

6. Kraus, J.D., "Electromagnetics," McGRAW-HILL, Inc., PP 679-681, 1984.

DESIGN OF A SPACEBORNE SAR AT HIGHER INCIDENCE ANGLE USING ALTERNATE TRANSMITTED CHIRP SLOPE REVERSAL TECHNIQUE

A.M.Jha , Archana Majmudar & N.S.Pillai

Microwave Sensors System Division
Microwave Sensors Development Group
Space Applications Centre
AHMEDABAD INDIA - 380053

Abstract

The system design of a spaceborne SAR at higher look angle calls for a trade-off among antenna complexity, wider swath and the required range ambiguity ratio (RAR).The requirement of minimum acceptable RAR even for a nominal swath may necessitate larger antenna width, complex weighting and highly asymmetrical antenna elevation pattern. The SCANSAR concept for wider swath leads to a trade-off against the azimuth resolution and system complexity. The scheme proposed, known as alternate transmitted chirp slope reversal (ATCSR) method, is capable enough to suppress the range ambiguity and reduce the antenna complexity consequently, without increasing the system complexity. A case study of a conceptual C-band SAR at ~ 50° incidence angle for a swath of ~ 150 kms., conducted utilising ATCSR method for uniform antenna elevation pattern, is presented in the paper.

Key words: Spaceborne SAR, Range ambiguity suppression, Higher incidence angle, Wider swath, Chirp slope reversal technique .

1. Introduction

The potentiality of the spaceborne SAR for monitoring the earth and its resources has created intensive interest among the user community. The various earth observation applications require wider swath for repeated coverage to monitor various time varying surface phenomena. Further, applications like crop monitoring and assessment, and various geological applications demand higher incidence angle operation along with the wider swath capability. The SAR launched on-board SEASAT-1 during 1978, and the SARs proposed on-board ERS-1 and JERS-1, are all designed to operate at lower incidence angle for a nominal swath. The shuttle imaging radar series (SIR-A,SIR-B and SIR-C) and the Canadian satellite RADARSAT have the capability to operate at higher angles but the swath mapped is very limited. Both the wider swath and higher incidence angle operation requirements, are severely constrained by range ambiguity limitations [1].A number of schemes have been suggested to suppress the range ambiguity and improve upon the obtainable swath . One of these is to design the antenna elevation pattern with the a-priori knowledge of the ambiguous return power which in turn is a function of the SAR elevation geometry and target reflectivity [2]. This technique results generally in a highly asymmetric pattern and requires a complex weighting. A pulse coding scheme to suppress the ambiguity, and subsequently improve upon the swath, has been suggested as another possibility [3]. A vertically scanned beam synthetic aperture radar system (SCANSAR) has been further suggested as an alternative and is presently being proposed for RADARSAT campaign[1].Here, the potential length of the synthetic aperture is shared between beam positions making the azimuth resolution poorer by a proportionate amount. Hence, the wider swath is achieved as a trade-off against the azimuth resolution and increased system complexity .

In this paper , an alternative to the above schemes is presented and is termed as alternate transmitted chirp slope reversal method (ATCSR). By this technique, significant reduction in ambiguity is achieved at the expense of minor modifications in the conventional SAR approach. This scheme can be effectively utilized to obtain wider swaths even at higher incidence angle with the least complex uniform antenna pattern .

2. System Design Constraints

The concept and design philosophy of spaceborne SAR is well documented and hence, is excluded in the discussion. As has been mentioned ,the present paper will discuss the impact, design complexity and subsequent solution through the ATCSR method, for a spaceborne SAR having the capability to map a swath of ~150 kms. around an incidence angle of 50°. The baseline system parameters required for the conceptual spaceborne SAR system are listed in table-1.

As antenna's mass and volume are very critical for any spaceborne campaign, it has been attempted to keep the area of the antenna around 10 m^2 for the analysis of the proposed C-band SAR.

As the average power requirement at higher

TABLE-1

Spacecraft altitude	800 kms
Operating frequency	5.3 Ghz (C-band)
Incidence angle	~50°
Range resolution	36 m.
Azimuth resolution	36 m.
Radiometric resolution	3 dB
Swath	~150 kms
Antenna area	~10 m^2
Back scattering coeff.	-22 dB
Range ambiguity ratio	< -20 dB
Azimuth ambiguity ratio	< -20 dB

incidence angle increases due to increase in slant range and decrease in the value of noise equivalent backscattering coefficient, the average power requirement of the system is going to increase significantly.This, for the proposed antenna area, is going to increase the average power requirement of the tube.For the present system, the availability of a C-band tube with an average power of 600 watts, has been presumed.

The dependence of radiometric resolution (γ), image SNR and no. of independent looks (N) can be expressed as

$$N \geq [(1+1/snr)/(10^{0.1\gamma} -1)]^2$$

Assuming a worst SNR of zero dB as acceptable for image, the above inequality constraints the maximum antenna length to 18m. for an azimuth resolution of 36 m. This corresponds to a minimum requirement of four independent looks in azimuth.

The wider swath requirement is significantly going to reduce the operating PRF. However, the data sampling window increases beyond proportion. Nonetheless, the smaller system bandwidth corresponding to the required range resolution, makes the data rate within manageable limits and will not pose any problem .

The PRF has to be selected so that it sufficiently oversamples the maximum doppler bandwidth and hence, satisfies the worst AAR requirement .

The main problem in the realization of larger swath at higher.angles is the large value of range ambiguity. Various methods with their limitations and inherent complexities, as mentioned, in section 1, have been suggested to overcome this. The present paper envisages the ATCSR method for the range ambiguity suppression and is discussed below.

3. Range Ambiguity Suppression

The ambiguity caused by the time overlap of the returns due to different transmit pulses, is termed as range ambiguity and has been discussed much in detail [4,5,6,7]. The range ambiguity ratio (RAR) is defined as the ratio of the sum of the ambiguous echo returns to the unambiguous signal and is dependent on the swath, PRF and most critically on the antenna elevation pattern. Referring to fig.1, the unambiguous signal is represented by the return from the 0^{th} pulse, whereas the overall return from the remaining pulses constitute the range ambiguity. A significant contribution to the ambiguity comes from the immediate adjacent pulses (marked +1 & -1 in fig.1). Let the odd ambiguity denote the overall contribution due to odd numbered pulses and even ambiguity denote the contribution due to the even numbered

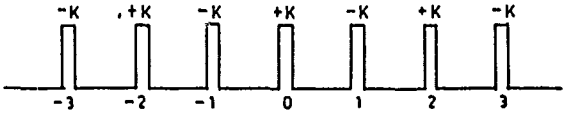


FIG. 1
SEQUENCE OF LFM PULSE TRAIN

pulses. Fig. 2 shows the variation of percentage contribution due to odd ambiguity for an incidence angle of 50.8°. A PRF of 1075 hz. and an antenna width of 0.63m with uniform illumination pattern have been considered for the analysis.

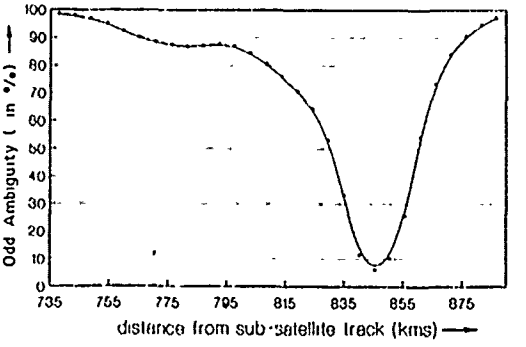


FIG. 2
VARIATION OF ODD AMBIGUITY ALONG THE SWATH

The analysis shows that towards the either ends of the swath, the odd ambiguity contribution is more than 90 % .Generally, RAR deteriorates significantly towards the end of the swath. Hence, if these odd ambiguities can somehow be suppressed, RAR can be significantly improved.

3.1 Alternate Transmitted Chirp Slope Reversal (ATCSR) Method

In case of conventional SAR , when LFM pulses are transmitted and the received echo is range compressed, both the signal of concern and the associated ambiguities are similarly compressed. The ambiguity,hence, is not affected by the range compression. If the characteristic of some or all of the pulses causing the ambiguities are changed from that of the transmitted pulse, the ambiguous power will be mismatched during the range compression. As stated earlier, the odd ambiguity contributes significantly to the overall ambiguity of the system. If the chirp slope of all the odd transmitted pulses are made equal and opposite to the even transmitted pulses, including the 0^{th} pulse, the odd ambiguity will be suppressed during the range compression and hence, will result in significant improvement in RAR. '+K' and '-K',inscribed above the pulses shown in fig 1 ,denote the slope of the chirp modulation.

This scheme has been proposed to suppress the range ambiguity and has been utilised to finalise the antenna configuration alongwith the evaluation of other relevant system parameters, and are discussed in next section.

4. System Design Analysis

Based on discussions carried out in the previous section, a software was developed to evaluate the various parameters relevant to the RAR analysis for the conventional and suggested ambiguity suppression technique.

In section 2., it was observed that the maximum antenna length was constrained by the azimuth and radiometric resolution requirements. Though the theoretically derived maximum antenna length is 18m., in practice, it needs to be further reduced to get four number of independent subapertures. The RAR optimization study for maximum swath around an incidence angle of 50.8°, leads to an antenna length of 16 m and width of 0.63 m

Fig 3 shows the variation of RAR with ground distance from the sub-satellite

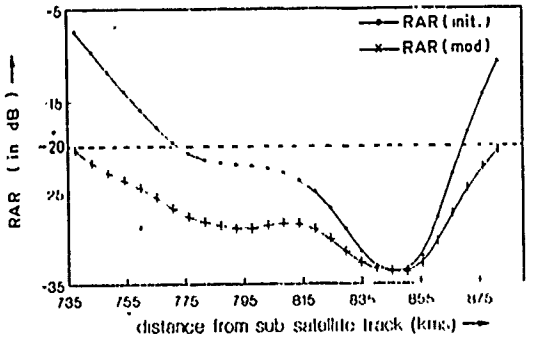


FIG 3

VARIATION OF RAR ALONG THE SWATH

track. In order to demonstrate the corresponding improvement due to the modified approach over the conventional one, both the curves, RAR(init.) for conventional and RAR(mod.) for modified approach are plotted. The variation of image SNR with ground distance is plotted in Fig 4. It is observed

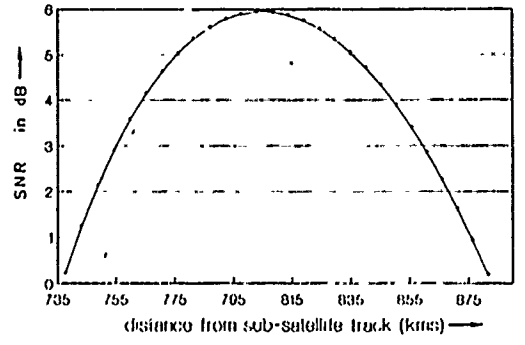


FIG. 4

VARIATION OF SNR ALONG THE SWATH

that a swath of 145 kms., from a ground distance of 737 kms. to 882 kms., satisfies the SNR criteria. However, RAR wise the conventional analysis gives the worst RAR of -7.4 dB. The odd ambiguity variation for this configuration was shown in fig 2 (presented earlier). This figure shows that wherever, especially at the ends of the swath, the contribution due to the odd ambiguity significantly rises, the RAR becomes poor. The modified approach suppresses the odd ambiguity and makes it possible to obtain a swath of 145 kms. As higher PRF deteriorates RAR , a PRF of 1075 hz has been considered for the present analysis. The final swath geometry is shown in fig 5. The swath corresponds to a

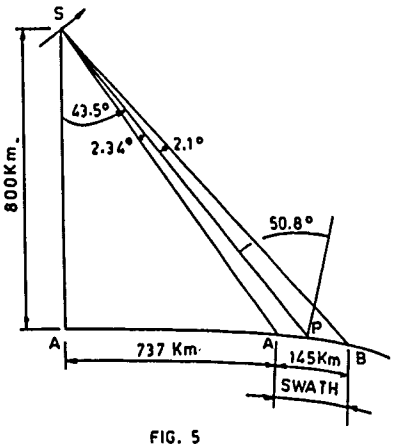


FIG. 5

SAR ELEVATION GEOMETRY

data window of ~767 us. The worst range resolution requirement of 36 m. gives rise to a compressed pulse width of 158 ns and a signal bandwidth of ~6.4 MHz. If a 20 % margin for reducing the aliasing is considered, a sampling frequency of 7.6 MHz. is obtained. For a sample represented by 6I+6Q bits (similar to ERS-1 SAR), the data rate in case of OBRC (on board range compression) mode comes out to be 83 Mbits/sec (considering a 10 % margin). The results of these analysis are summarized in table-2.

TABLE -2

Altitude	800 kms.
Operating frequency	5.3 GHz.
Incidence angle	50.8°
Range Resolution	36 m.
Azimuth Resolution	36 m.
Swath	145 kms.
Antenna length	16 m.
Antenna width	0.63 m.
Antenna peak gain	43.76 dB
Backscattering coeff.	-22 dB
Range Ambiguity Ratio	-20 dB (worst)
Azimuth Ambiguity Ratio	-21 dB (worst)
Pulse Repetition Freq.	1075 hz.
Average Power	600 watts
Compressed Pulsewidth	158 ns
Signal bandwidth	6.4 Mhz.
Sampling freq.	7.6 Mhz.
Data Rate	~83 Mbits/sec.
Quantizing bits (OBRC mode)	6I + 6Q

5. Design Implications

The realization of the ATCSR method will require some minor modifications to be introduced at the transmitter and signal processor level. As the transmitted signal bandwidth remains the same, the receiver chain is totally unaffected. The transmitter section needs to be modified to facilitate the modulation of the transmitted pulses with up

and down chirps alternately . Minor modifications are necessary at the signal processor level to compress the return pulses with the required chirp slope.

6. Conclusion

A conceptual spaceborne SAR design having a capability of mapping a wide swath (~145 kms) around 50° incidence angle has been proposed. Satisfactory RAR level is obtained using ATCSR method which obviates the need for a complex antenna elevation pattern or to resort to other range ambiguity supression schemes with their inherent complexities and limitations. The average power and the data rate requirement are within the present state of the art.

References

[1] R.K.Moore, J.P.Classen, Y.H.Lin,'Scanning spaceborne synthetic aperture radar with integrated radiometer',IEEE Trans. on Aerospace and Electronic Systems, Vol. AES-17,No. 3,May 1981.

[2] Barbarossa S. and Levrini G.,'An antenna pattern synthesis technique for spaceborne SAR performance optimization',Proceedings of IGARSS '87 Symposium, Ann Arbor,18-21 May 1983.

[3] Westing House Electric Corporation, 'Spaceborne synthetic aperture radar pilot study',Final report, NASA Contract NA55-21951, Baltimore, Md., April,1974 .

[4] A.M.Jha, Archana Majmudar, N. S. Pillai, 'Range ambiguity suppression technique for the spaceborne SAR ',39th International Astronautical Congress of the IAF,Bangalore,India,Oct. 1988.

[5] Bayma R.W. and McInnes P.A.,'Aperture size and ambiguity constraints for synthetic aperture radar',IEEE, 1975, International Radar Conference.

[6] Mehlis J.G., 'Synthetic aperture radar range-azimuth ambiguity design and constraints', IEEE International Radar Conference, 1980.

[7] Li F.K, Johnson W.T.K.,'Ambiguities in spaceborne synthetic aperture radar systems', IEEE Transactions on Aerospace and Electronic Systems, Vol. AES-19,No. 3,May 1983.

RASS APPLIED TO WIND PROFILER RADARS

Peter T. May[1], Richard G. Strauch[2], and Kenneth P. Moran[2]

Abtract

Radio Acoustic Sounding (RASS) has been applied with wind profiler radars. Temperature profiles have been obtained up to several kilometers in altitude with RMS differences between the RASS measurements and conventional radiosonde observations of 1°C have been obtained. The technique shows great promise for a number of meteorological applications.

1. Introduction

Radio Acoustic Sounding Systems (RASS) are a combination of acoustic excitation and radar detection that measures profiles of the speed of sound. Our experiments combine high-power acoustic sources with wind profiling radars, which are sensitive enough to measure backscatter from the clear air in the free troposphere in a continuous manner. RASS has been used to obtain temperature profiles, primarily in the boundary layer, since the early 1970's (e.g. Marshall et al., 1972), but concerns over limited height coverage and the ability to operate only in low wind conditions have prevented wide acceptance of the technique. However, recent experiments with wind profiling radars have shown significantly greater height coverage and less dependence on wind than was found in early RASS experiments (Matuura et al., 1986; May et al., 1988).

RASS makes use of the resonant backscatter that occurs when the acoustic wavelength is half the radar wavelength (Bragg matched) and the focusing of the backscattered radiation onto a diffraction-limited spot by the spherical acoustic wavefronts. This latter condition is not required with wind profiler radars because sophisticated signal processing allows the detection of very weak signals (with signal-to-noise ratios as low as −35 dB into the receiver (May and Strauch, 1989)) and turbulence acts to defocus the spot. Thus Doppler velocity estimates can be made at heights where the radar does not observe the focussed spot. This allows the detection of the speed of sound c_a, which is related to the virtual temperature T_v, to a good approximation, by $T_v = (c_a/20.047)^2$ K. Vertical winds add to the measured c_a and thus can produce errors under strong vertical wind conditions such as downslope wind storms (Klemp and Lilly, 1975) or convection.

Wind profilers are pulsed Doppler radars designed to observe the backscatter from the clear air on a continuous basis in the troposphere. The WPL profilers (Table 1) used in this study use frequencies

Table 1. Characteristics of RASS/wind profiler systems operating at three radar frequencies

Site			
Characteristic	Platteville	Erie	Denver
Radar			
Frequency (MHz)	49.8	404.37	915
Wavelength (m)	6.0	0.7414	0.3
Antenna size (m²)	10,000	25	100
Beamwidth (1 way) (degrees)	3	7.8	2
Range resolution (m)	300	150	150
Mean power (W)	200	30	100
Sampling time (1 profile) (min)	1	1	1
Acoustic			
Frequency (Hz)	~ 110	~ 900	~ 2000
Beamwidth (1 way)	~ 60°	~ 17°	~ 8°
Acoustic power (W)	50	5	50
RASS Altitude Coverage			
Min. altitude (km)	2.0	0.4	0.2
Max. altitude (km)	4–11	1.5–2.5	1–2

[1] Cooperative Institute for Environmental Sciences (CIRES), Universitiy of Colorado/NOAA, Boulder, Colo., U.S.A.

[2] Wave Propagation Laboratory, ERL, NOAA, Boulder, Colo., U.S.A.

of 49.8, 404.37, and 915 MHz (Strauch et al., 1984). Note that the 404.37 MHz system used here is a low power radar and that the new profilers for the NOAA network will be about 26 dB more sensitive. RASS is incorporated into the current wind profilers by introducing a frequency offset into the receivers corresponding to about 300 m s^{-1}, while all the remaining signal processing for clear air observations remains unchanged. The Nyquist frequency corresponds to a velocity of about 50 m s^{-1}, so that an interval of 300 $\pm$ 50 m s^{-1} is sampled. A frequency modulated CW acoustic source is used such that somewhere within each pulse volume the acoustic wave is Bragg matched during the data acquisition. The acoustic power is about 50 W.

2. Factors affecting range coverage

There are three main factors affecting the height coverage. They are (1) wind displacing the acoustic waves, (2) acoustic absorption in the air, and (3) turbulent distortion of the acoustic wavefronts (May et al., 1988).

The effect of the wind has been a major limitation in the past, but the greatly enhanced sensitivity of wind profiler radars has allowed measurements even in wind storms with wind speeds of about 30 m s^{-1}. We use acoustic sources with beamwidths wide enough so that the acoustic pressure perturbations are in the radar beam even under high wind conditions. The 49.8 MHz radar is the most affected by winds. This radar has observed profiles as high as 11 km above ground under light wind conditions, but 4–6 km is typical under moderate to strong winds.

Acoustic attenuation is a very strong function of frequency. The attenuation measured in dB/100m increases as the frequency squared. It is unimportant at frequencies around 100 Hz (corresponding to the 49.8 MHz profiler), but is the dominant factor at 2000 Hz (for the 915 MHz profiler). The height coverage for the 915 MHz profiler depends on the temperature and humidity, which affect the acoustic absorption (e.g., Harris, 1966).

The effect of turbulence is ambivalent. Turbulent distortion of the acoustic wavefronts acts to decrease the radar signal power faster than range squared decreasing height coverage, particularly at higher frequencies. On the other hand turbulence acts to smear the spot over a large area, which dramatically lessens he effect of wind on the measurements.

In summary it is the latter two effects which most strongly limit the 404.37 and 915 MHz profilers to maximum altitudes of about 1–2 km while winds have the most dramatic effect on operations with the 49.8 MHz profiler.

3. Accuracy of RASS

Very little quantitative discussion on the accuracy of RASS measurements has been shown in the literature. Numerous temperature profiles measured with RASS and radiosondes have been shown, but the discussion has been limited to describing the agreement as excellent (e.g., Figure 1). In order to discuss our results more quantitatively a data set incorporating over 50 radiosonde ascents during two campaigns in the summer and winter of 1988 have been compared to 15 min averages of the RASS observations centered at the launch time. No attempt has been made to simultaneously measure the vertical component of the wind, so the averaging is an attempt to lessen the impact of vertical motions on the data set. Modifications to the radar to allow simultaneous measurement of the vertical motions and RASS are planned.

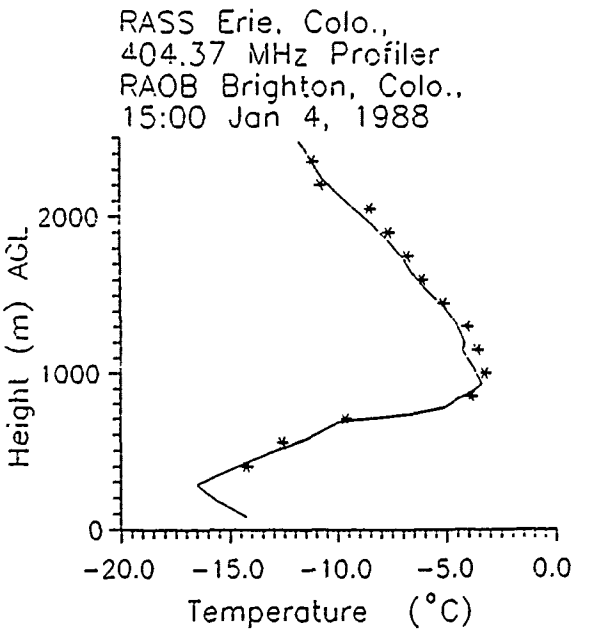


Fig. 1. A comparison between RASS measurements using the 404.37 MHz profiler at Erie Colo, (stars), and a radiosonde launched at Brighton (solid line), about 24 km away. The time is given in U.T. (after May et al., 1988)

Figure 2 shows a scatter plot comparing the data sets and the agreement is striking. The RMS difference is 1.0°C, which is comparable to that obtained with radiosonde-radiosonde comparisons. There is a section of the data which shows greater

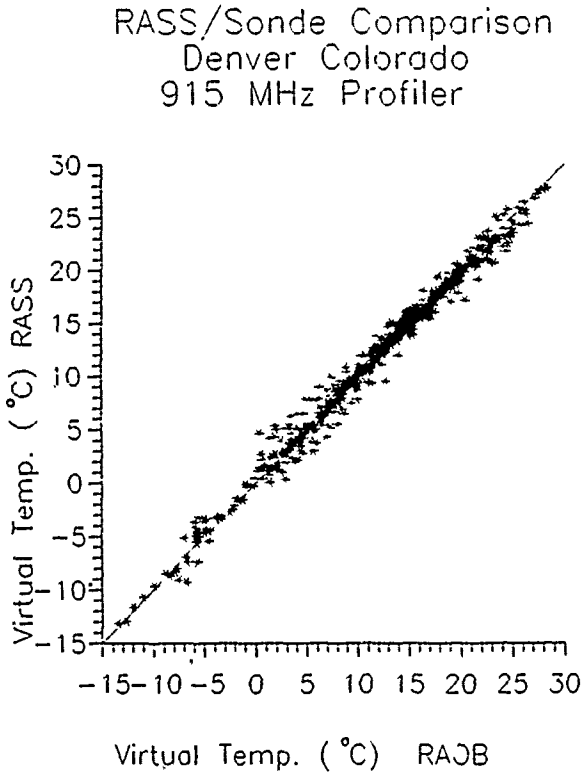


Fig. 2. Virtual temperature measured by radiosondes and by RASS using the Denver 915 MHz wind Profiler. The RMS difference is 1.0°C.

variation at temperatures around 3–5°C. Examination of the wind fields measured duringthese periods reveals vertical motions of several m s^{-1} associated with downslope wind storms. If data during events such as this are removed; the RMS difference between the temperature observations is reduced to about 0.6°C. Further evidence of the effect of vertical wind motions is that in the summer the RMS difference between the data sets decreased with height in the convective boundary layer while in the winter they increased, probably because of lee waves and wind storms.

Clearly correction of the errors caused by vertical motions is the next step in the RASS/wind profiler technical development. Such corrections are possible with the requirement of more computer power for data analysis. Observations under very light wind conditions have shown RMS variations of the RASS observations of only about 0.2°C at a given height over a period of 36 min. This illustrates the even greater promise of the technique when a vertical velocity correction is available.

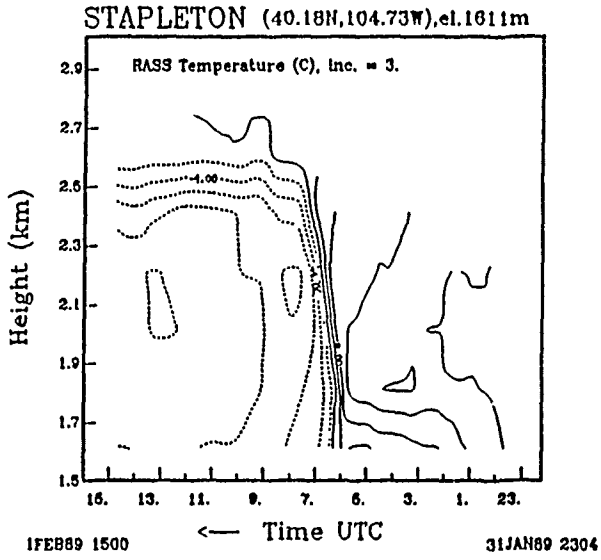


Fig. 3. Temperature field measured using the Denver profiler/RASS during the passage of an intense shallow cold layer on February 1, 1989.

4. Weather observations

This example shows data from the passage of an intense cold front over the Denver profiler and illustrates the promise and capabilities of the profiler/RASS combination.

On January 31, 1989, an arctic airmass was pushing down from the north along the eastern side of the Rockies. In the early morning of February 1 the initial cold surge associated with this cold front passed over Denver. The Stapleton radar was running a sequence of 6 min of RASS, 6 min of clear air observations, and 3 min for analysis during the entire weeklong period of the arctic event. Figure 3 shows the temperature field measured by RASS during the initial surge from the north. Prior to the passage of the front a ground based inversion developed because of radiational cooling. The temperature drop at the ground was about 10°C at the surface, and almost 20°C in an hour a few hundred meters above the ground. Notable features include the rapid temperature drop, the shallow depth of the cold air (~ 1 km), and the strong temperature inversion which resulted. The wind field during this period showed the shallow layer of northerlies associated with the surge and some evidence of a head, which was not seen in the temperature fields. The zonal component of the wind (Fig. 4) shows the

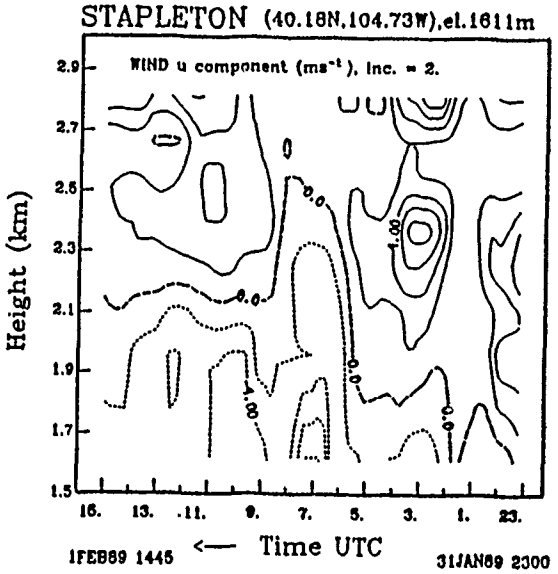


Fig. 4. The zonal component of the wind field measured using the Denver wind profiler during the frontal passage.

wind shear associated with the boundary. The speed of the cold surge (about 15 m s^{-1}) was fairly consistent with a density current with depth 900 m and a temperature contrast of 12°C, which is close to the drop over the period of the sharpest gradient. Subsequent surges deepened the cold air to only about 1.5 km with minimum temperatures around −30°C. The temperature inversion associated with the cold air was about 20°C over a single radar resolution cell (150 m). The actual gradient may have been even greater.

This is one case among many which have already been observed. Other applications include severe pollution events, aircraft icing, and possible operational applications, for example, in forecasts of convection.

5. Conclusions

The RASS technique applied with wind profilers clearly has great potential for temperature profiling as both a research and operational tool. The accuracy is of the order of a degree and the continuous time and height coverage gives an instrument analogous to a tower several kilometers high. A new network of wind profilers currently being built by NOAA should have a typical height coverage up to about 3–4 km above ground. The major problems to be solved are vertical velocity corrections made in real time and noise pollution, which is quite severe with the high powered acoustic sources which are currently used.

Acknowledgments: Contributions from B. Stankov and P. Neiman are gratefully acknowledged.

References

Harris, C.M., Absorption of air versus humidity and temperature, J. Acoust. Soc. Amer., 40, 148–160, 1966.

Klemp, J.B., and D.K. Lilly, The dynamics of wave induced downslope winds, J. Atmos. Sci., 32, 320–335, 1975.

Marshall J.M., A.M. Peterson, and A.A. Barnes Jr., Combined radar-acoustic sounding system, Applied Optics, 11, 108–112, 1972.

Matuura, N., Y. Masuda, H. Inuki, S. Kato, S. Fukao, T. Sato and T. Tsuda, Radio acoustic measurement of temperature profile in the troposphere and stratosphere, Nature, 333, 426–428, 1986.

May, P.T., R.G. Strauch and K.P. Moran, The altitude coverage of temperature measurements using RASS with wind profiler radars, Geophys. Res. Lett., 15, 1381–1384, 1988.

May, P.T. and R.G. Strauch, An examination of some algorithms for spectral moment estimation, Preprint Vol. 24th Radar Conf. Radar Meteorology, Tallahassee, Fl., March 1989.

Strauch, R.G., D.A. Merritt, K.P. Moran, K.B. Earnshaw and D. van de Kamp, The Colorado wind profiler network. J. Atmos. Oceanic Tech., 1, 37–49, 1984.

APPLICATION OF SEASAT SCATTEROMETER AND PASSIVE MICROWAVE RADIOMETER DATA TO REGIONAL SHORT-RANGE WEATHER FORECASTING

S. Peteherych

Atmospheric Environment Service
Downsview, Ontario, Canada

ABSTRACT

Three of Canada's six regional weather centres prepare weather forecasts for ocean areas. They all suffer from a lack of ship data for wind forecasting and inadequate data for precipitation forecasting. All regions have access to cloud imagery. Cloud-level storm centres can be identified but they can be displaced 100-500 km from the ocean-surface storm centres. In addition qualitative estimates of water content can be obtained from the brightness of clouds.

The Seasat scatterometer provided an abundance of high-resolution wind speed and direction data. It could locate a storm centre within a resolution slightly better than its footprint resolution. The passive microwave radiometer, also on Seasat, provided quantitative observations of column abundances of rain rate, liquid water and water vapour.

Both instruments can be expected to provide significant improvements for wind and precipitation forecasting.

Keywords: Scatterometer, Passive Microwave Radiometer, Weather Forecasting.

1. BACKGROUND

The Seasat Scatterometer is an active radar that has the ability to measure ocean wind speed and wind direction. The measurement technique depends on the fact that winds generate centimetre-length waves, which increase in number with increasing wind speed. The scatterometer ia essentially an all-weather device that can yield high-resolution data (50 km). Wind direction ambiguities occur because the backscatter in both the upwind and downwind directions is virtually identical. Similarly, there is no difference in backscatter between the two cross-wind directions. Direction ambiguities can be removed by meteorological analysts with an estimated accuracy of 90-95%. Future scatterometers will have the capability to reduce these ambiguities significantly. The incorrect directions are called aliases and the process of removing aliases is called "de-aliasing". Wind speed estimates are accurate to ± 10% or 2 ms^{-1}, whichever is larger, and wind directions after de-aliasing are accurate to ≤20°.

Six cases were studied to evaluate the utility of Seasat scatterometer data for regional short-range ocean wind forecasting and were published by Peteherych et al. (1988).

The SEASAT scanning multichannel microwave radiometer (SMMR) was a passive microwave radiometer that measured the radiant energy emitted from the earth and atmosphere at 5 microwave frequencies. This sensor was capable of measuring ocean surface temperature, ocean wind speed, column abundances of rain rate, non-raining water and water vapour. Verification studies have been reported by Alishouse (1983), Gurvich et al. (1970) and Katsaros et al. (1981).

2. INTRODUCTION

In Canada there are three regional weather centres that have the responsibility for forecasting weather over the oceans: the Maritme Weather Centre in Bedford, Nova Scotia,the Arctic Weather Centre in Edmonton, Alberta and the Pacific Weather Centre in Vancouver, British Columbia. They all share a common problem - a deficiency of ship data over their corresponding ocean areas. They all receive satellite visible and infrared imagery. Storm locations in cloud imagery are often different from their storm location at the ocean surface. Typically the differences are 100-200 km and occasionally as much as 500 km. Cloud imagery can provide a qualitative measure of the surface wind speed, whereas the scatterometer locates the storm location and intensity with precision.

The whiteness of a cloud image is a qualitative measure of the water content. A passive microwave radiometer identifies the location and amount of rain rate, non-raining water and water vapour.

3. RESULTS AND DISCUSSION

There are about 20 ship reports within 800-1000 km of an intense storm at 51°N, 150°W (Fig.1). This is an exceptionally large number for a Pacific coast cyclone.

These reports are sufficient to prepare an acceptable MSL pressure analysis for this storm. Ocean ships report wind speed in knots whereas Seasat obtains them in ms-1. The ratio is virtually 2 to 1. Wind direction is in the direction of the arrow

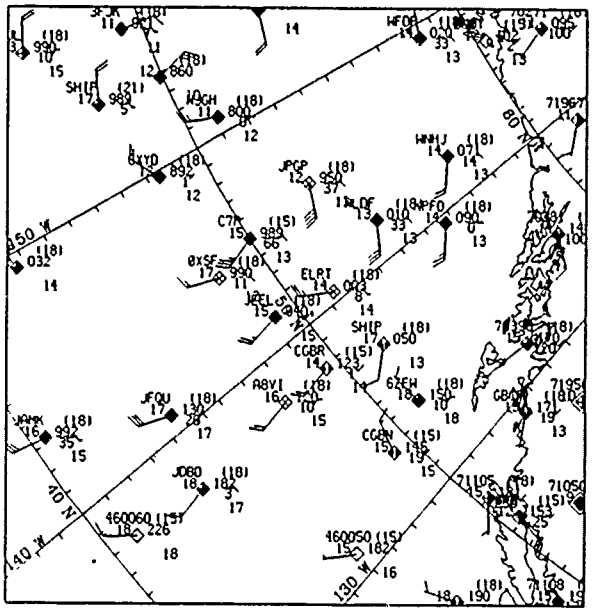


Fig. 1. OCEAN SHIP WINDS SEPT 11 1978

accompanying each wind speed number. There are 4 times as many Seasat observations, (Fig.2) for a smaller area. The ship data show one observation at 40 kts. due east of the storm. Seasat observations report 4 winds at 20 ms^{-1} north and west of the storm centre.

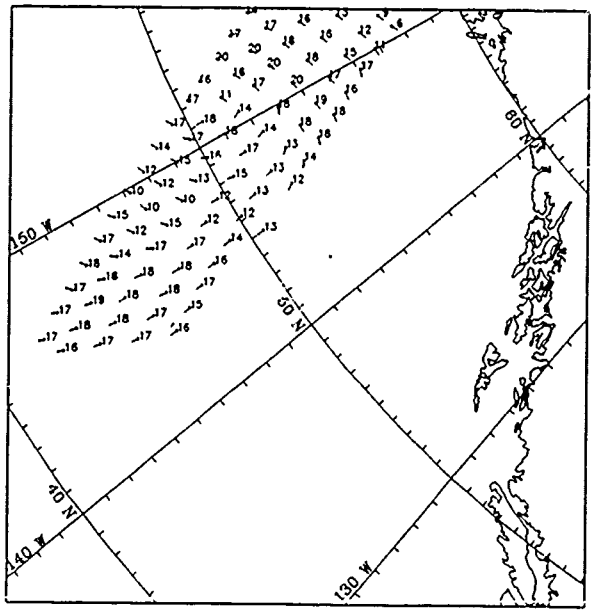


FIG. 2. SEASAT SCATTEROMETER WINDS SEPT 11 1978

The Canadian Meteological Centre (CMC) MSL charts for this period are available, but not included in this paper due to lack of space, show a much lower resolution analysis than shown in (Fig.3) which was performed manually by marine analysts.

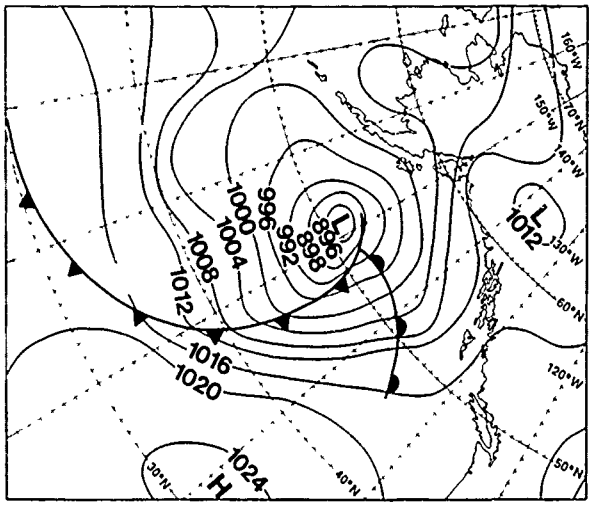


FIG. 3 MSL PRESSURE CONTOURS SEPT. 11, 1978

Rain (Fig.4) is located on the cold side of the warm and cold front as expected, and liquid water (Fig.5) is a maximum in the same areas where the rain rate is a maximum. Maximum water vapour (Fig.6) is also located in the same location as the maximum rain rate and liquid water. In addition the maximum values of rain rate, liquid water and water vapour occur near or at the frontal boundaries.

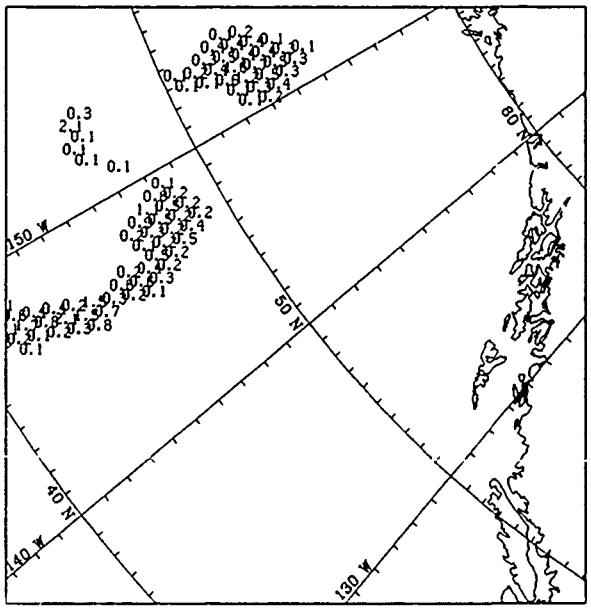


FIG. 4. SEASAT SMMR RAIN RATE (MM/HR) SEPT 11 1978

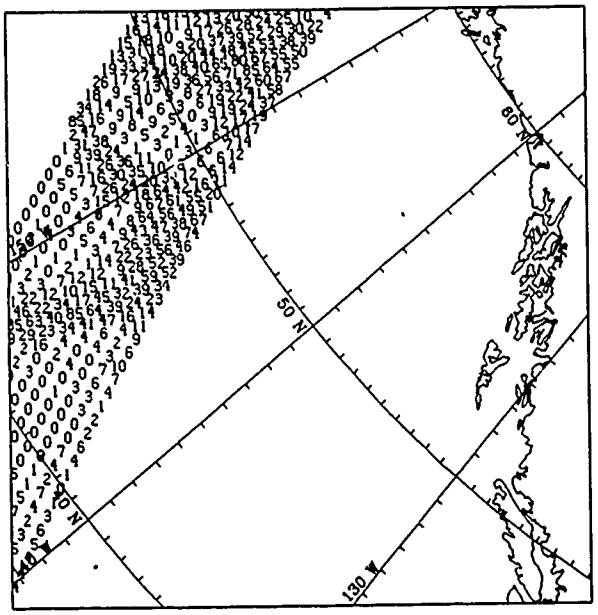


Fig. 5. SMMR LIQUID WATER (mg/cm²) SEPT 11 1978

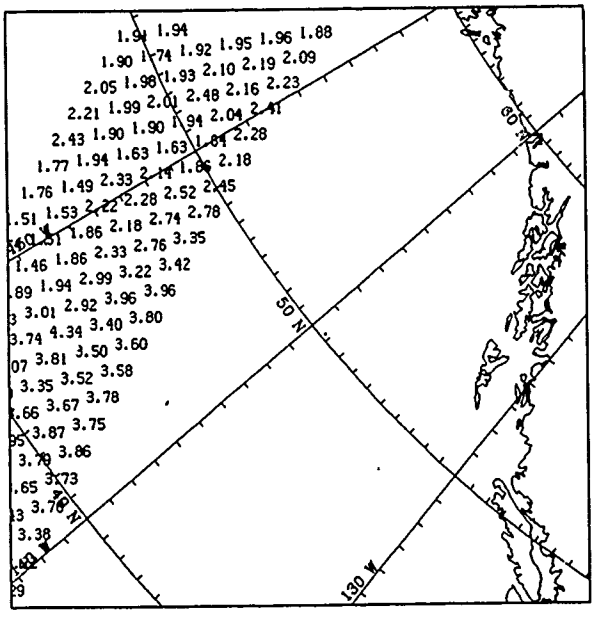


Fig. 6. SMMR WATER VAPOUR (g/cm²) SEPT 11 1978

4. CONCLUSION

Quality scatterometer data are expected to improve the forecasting of wind speed and direction for all coastal regions. A passive microwave radiometer will provide improved precipitation forecasting for these areas.

5. ACKNOWLEDGEMENTS

The author wishes to acknowledge the effort of A.F. Davies for providing comments, E. Truhlar for editing this paper, Mahesh Raghunathan, my student assistant, for preparing the figures and P. Pearson for typing the manuscript.

6. REFERENCES

1. ALISHOUSE J.C., "Total Precipitable Water and Rainfall Determination from the SEASAT Scanning Multichannel Microwave Radiometer", J. Geophys. Res. Vol. 88 No. C3, Pages 1929-1935, 1983.

2. GURVICH A.S., and V.V. Demin,"Determination of the Total Moisture Content in the Atmosphere from Measurements on the Cosmos 243 Satellite", Atmos. Oceanic Phys., 6, 453-457, 1970.

3. KATSAROS, K.B.; P.K. Taylor, J.C. Alishouse and P.G. Lipes. "Quality of SEASAT SMMR (Scanning Multichannel Microwave Radiometer) Atmospheric Water Determination". In: Oceanography From Space, J.F.R. Gower (ED.), pp 691-706, Plenum Press, New York, 1981.

4. PETEHERYCH S.; W,S. Appleby, P.M. Woiceshyn, J.C. Spagnol and L.Chu. "Applications of SEASAT Scatterometer Wind Measurements for Operational Short-Range Weather Forecasting", Weather and Forecasting Vol. 3 89-103, 1988.

DETERMINATION OF NEAR-SURFACE WIND SPEEDS FROM SUNGLINT DATA IN BAND 2 OF THE AVHRR AND COMPARISON WITH METEOROLOGICAL DATA FOR THE WESTERN MEDITERRANEAN

A.P. Cracknell, S. Khattak and R.A. Vaughan

Department of Applied Physics and Electronic
& Manufacturing Engineering,
University of Dundee,
DUNDEE DD1 4HN, Scotland, U.K.

ABSTRACT

We have used the well-known theory of Cox and Munk to determine near-surface windspeeds from AVHRR visible/near-infrared band data for 23 May, 21 June, 28 June and 5 July 1982 for areas of sunglint in the western Mediterranean. The results have been compared with meteorological data obtained from the UK Meteorological Office for these dates in this area. The discrepancies are generally not unreasonably large, moreover it is by no means clear whether the discrepancies arise from errors in the meteorological data or from errors in the retrieval from the satellite-data. The relevance of this work to the validation of wind-speeds that are expected to be derived from ERS-1 microwave data is considered briefly. An example of the temperature rise associated with the skin effect in an area of calm water in one scene has also been calculated.

Keywords: Sunglint, wind speeds, AVHRR, skin effect.

1. INTRODUCTION

The theory of Cox and Munk (1954) for determining near-surface windspeeds from sunglint data is well established and has been quite widely used. We have obtained some 1982 AVHRR data for areas of sunglint in the western Mediterranean for the purpose of extracting near-surface wind speeds and comparing the results with data from wind speed measurements in the area supplied to us by the U.K. Meteorological Office.

First of all we carried out a visual inspection of prints supplied to us by the Dundee University Satellite Data Receiving Station for the various spectral bands. In one of the scenes we noticed areas of low windspeed, which appear dark in band-1 (visible) and band-2 (near infrared) AVHRR data, according to the explanation given by McClain and Strong (1959). We found similar dark patches in the band-4 (thermal infrared) AVHRR data for the same scene and these correspond to areas where the surface temperature is higher than that of the surrounding water. The cause of the local heating of the surface layer is that the surface of the water is warmed by solar heating and in a region of calm water, associated with low wind speed, there is little mixing and the heated water remains at the surface. The spatial correlation between the dark patches in bands 1 and 2 [figure 1(a) and (b)] and in band 4 [figure 1(d)] is very striking.

The purpose of the work described in this paper was to quantify these effects which were observed visually by carrying out detailed calculations of the near-surface wind speed and to make some estimates of the temperature rise caused by local heating in areas of low wind speed.

2. SUNGLINT AND THE THEORY OF COX AND MUNK

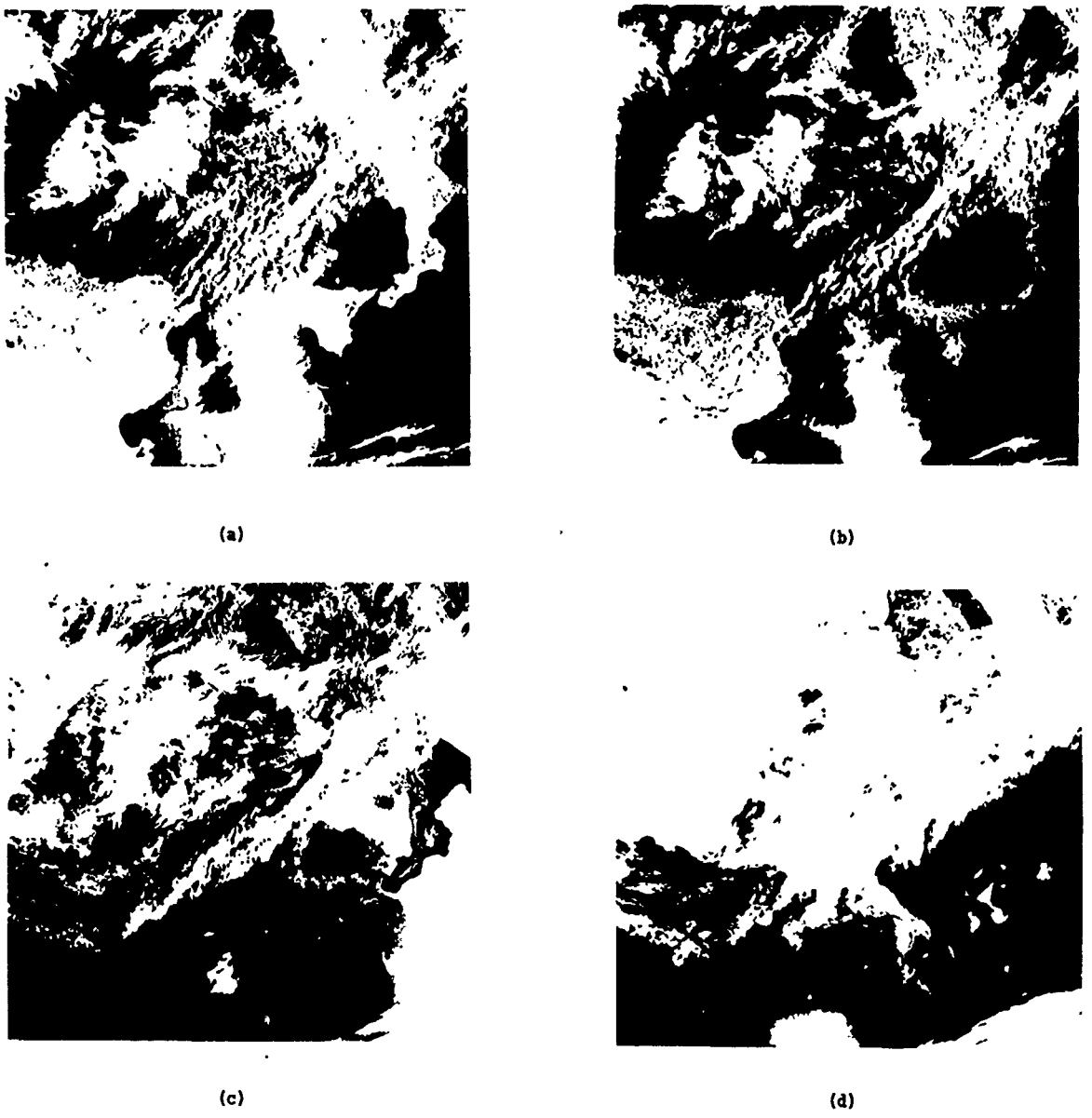
Sunglint, the direct specular reflection of sunlight at the surface of the sea, is something that people using remotely-sensed data usually regard as a nuisance and take steps to avoid. That is, in most situations one is aiming to study the surface of the sea by diffusely-reflected radiation. Thus, for example, on the NIMBUS-7 satellite there was special provision made in the Coastal Zone Colour Scanner (CZCS) to tilt the mirror so as deliberately to avoid gathering data from an area of sunglint. Also, very frequently, in flying an aircraft, one deliberately organises the flight in such a way as to avoid obtaining data that includes sunglint. In the present work we take the approach of deliberately using sunglint data to determine near-surface wind speeds using the theory of Cox and Munk (1954). It is not necessary to repeat the details of that theory here, it has recently been summarised elsewhere (Cracknell 1989). We shall simply summarise the main points very briefly.

The theory of Cox and Munk (1954) is based on an assumption of a Gaussian distribution to describe the various orientations of the tangent (or normal) to the water surface when waves are present. The probability distribution function assumed is

$$P(\theta_N, V) = \frac{1}{\sigma^2} \exp\left(\frac{\tan^2 \theta_N}{\sigma^2} \right) \qquad (1)$$

where θ_N is the zenith angle of the normal at the point at which reflection occurs, V is the near-surface wind speed and σ is the standard derivation of the distribution. A second assumption is that one can use an empirical relation

$$\sigma^2 = 0.00512\ V + 0.003 \qquad (2)$$



(a) (b)

(c) (d)

Figure 1. NOAA-7 AVHRR data for 1310 GMT on 23 May 1982 (a) band 1, (b) band 2, (c) band 3 and (d) band 4 (copyright Dundee University).

between the standard deviation, σ, and the near-surface wind speed; the value of V in equation (2) is in ms^{-1} and the values of the coefficient of V and of the constant term on the right-hand side of the equation were determined empirically. Subsequent to the original work of Cox and Munk (1954) other versions for the expression relating σ and V have been explored.

It is then necessary to relate the satellite-received radiance, $L_s(\lambda)$, at wavelength λ to the probability distribution of the slopes, i.e. of θ_N. We can write

$$L_s(\lambda) = L_w(\lambda) + L_{sk}(\lambda) + L_g(\lambda) \quad (3)$$

where

$L_w(\lambda)$ = water-leaving (diffuse radiance)

$L_{sk}(\lambda)$ = diffuse sky radiance

and

$L_g(\lambda)$ = sunglint radiance.

Then we have

$$\frac{L_g(\lambda)\ \cos^4\theta_N\ \cos\theta}{L\ (\lambda)\ r(\omega)} = P(\theta_N, V) \qquad (4)$$

where

θ = zenith angle of reflected ray

$r(\omega)$ = Fresnel reflectivity

and

$\omega/2\pi$ = frequency of light (= c/λ (where c is the velocity of light)).

For a part of the scene that is far removed from the area of sunglint the term $L_g(\lambda)$ will be absent. If we assume that there are no large concentrations of suspended sediment or of chlorophyll near the surface in the areas being studied, then we can assume that $L_w(\lambda)$ has the same value throughout the scene so that the value of $\{L_w(\lambda) + L_{sk}(\lambda)\}$ for the sunglint area can be taken as given by the satellite-received radiance from a region outside the sunglint area. This also assumes there are no horizontal spatial variations in the atmospheric conditions throughout the whole area.

Thus the procedure that we have implemented for calculating near-surface wind speeds is as follows. For a given pixel in the sunglint area we obtain $L_s(\lambda)$ from the AVHRR band-1 data for that pixel and we obtain $L_g(\lambda)$ for that pixel by the method described at the end of the previous paragraph. We perform the necessary trigonometry to obtain θ, θ_N and $r(\omega)$ and thus calculate $P(\theta_N, V)$ using equation (4). From this we use equation (1) to calculate σ^2 and then use this value of σ^2 in equation (2) to calculate V, the near-surface wind speed for the surface area corresponding to the given pixel in the image.

3. CALCULATIONS OF WIND SPEEDS

We have used AVHRR data from 4 dates in the summer of 1982 in which sunglint is present. The dates are 23 May 1982, 21 June 1982, 28 June 1982 and 5 July 1982. Images for the first of these dates have already been shown, in four bands, in fig. 1. The procedure described at the end of section 2 has been applied at about 20 points on a rectangular grid within the sunglint area of each scene and the sunglint-derived near-surface wind speeds, which we denote by V_s, have been obtained.

We obtained data on magnetic tape from the U.K. Meteorological Office giving observed values of wind speeds at some points in the Mediterranean on the dates for which we have used the AVHRR data. The number of points at which observations were made was quite small and we have had to interpolate to estimate the wind speed, which we denote by V_m, at the points on the rectangular grid used for our calculations. The means and standard deviations of the interpolated values from surface observations, V_m, and the sunglint-derived values, V_s, are given in table 1.

TABLE 1. Wind Speeds derived from meteorological data and sunglint data

DATE	23 MAY 1982		21 JUNE 1982		28 JUNE 1982		5 JULY 1982	
	V_m	V_s	V_m	V_s	V_m	V_s	V_m	V_s
$\bar{V}$ (ms^{-1})	2.6	5.75	3.97	5.13	5.1	4.73	5.39	5.36
σ (ms^{-1})	-	0.42	0.10	0.08	-	0.62	0.36	0.36

For two of these scenes, namely 23 May and 28 June, the meteorological observations are so sparse that the value of V_m is the same for each point in the sunglint area. Thus no standard deviation can be determined and the meteorological data provide very little useful information to help in assessing the reliability of the sunglint-derived wind speeds for these two dates. A further problem is that, for all 4 scenes, there was a time difference of the order of an hour or more between the meteorological observations and the reception of the AVHRR data. A comparison between the sets of values of V_m and V_s tells us more about the quality of the meteorological observations than about the quality of the sunglint-derived wind speeds. From table 1 we see that the meteorological observation of 2.6 ms^{-1} for 23 May 1982 is almost certainly far too low whereas the meteorological observation of 5.1 ms^{-1} for 28 June 1982 is probably a reasonable mean value for the area being studied. However, it is not really the main objective of our study to evaluate the quality of the meteorological observations.

For the remaining two scenes it is instructive to consider the mean and standard deviation of V_m, the meteorological wind speed, and V_s the sunglint-derived wind speed, see table 1. From the table we see that for 5 July 1982 the difference between the means (0.03 ms^{-1}) is much smaller than the standard deviation of either wind speed and there is no evidence of bias between the two data sets. But for the data from 21 June 1982 there is clear evidence of bias between V_m and V_s. For 21 June the difference between the means is 1.1(6) ms^{-1} while the standard derivations of V_m and V_s are about 0.1 ms^{-1}. For this date then the means of V_m and V_s differ by about 10 standard derivations and this is clear evidence of bias.

Having established clear evidence of bias in one case out of the four scenes studied, one should try to establish the cause of the bias. Contributory factors include *(i)* errors in the meteorological observations themselves *(ii)* a time difference between the time of the meteorological observations. What is interesting is the question of whether there is any systematic error, or bias, in the use of the procedure based on the Cox and Munk theory in our analysis of the AVHRR data. This is unresolved by our present work and it would require the analysis of far more scenes before this question could be resolved.

4. THE SKIN EFFECT

In this section we consider briefly an example of part of one of these scenes in which there is evidence of an increase in the surface temperature arising from solar heating and the absence of mixing by wave action. The scene is that of 23 May 1982.

In figure 1(a), which shows the band-1 AVHRR data for 23 May 1982, there are two calm areas showing as dark areas in the sunglint pattern; one is south-west of Sicily and the other is north-east of Sicily, adjacent to the edge of the clouds. There areas are clearly visible in band 2 (fig. 1(b)), while they appear dark, in band 3 (fig. 1(c)) where they appear light and in band 4 (fig. 1(d)) where they appear dark. The correlation of low wind speed and surface heating has been studied before by Sanders et al (1981) who used a combination of Meteosat data and AVHRR data.

We have calculated the brightness temperature, T_s, for a typical sunglint area and for the dark patches with the AVHRR data for 23 May 1982. We have not made any attempt to perform atmospheric corrections to convert the brightness temperatures into sea-surface temperatures. This is because we are only interested in the differences between the temperatures of areas with different near-surface wind speeds. This does of course mean that we are assuming a horizontally-stratified atmosphere over the whole area being studied. We have calculated T_s for 22 points in each of the following *(a)* outside the sunglint area, *(b)* in the bright part of the sunglint area (wind speed circa 3-4 ms^{-1}) and *(c)* in the dark area within the sunglint (wind speed negligible). The mean of the 22 values for each case was:

(a)	outside sunglint area	287.5(8) K
(b)	sunglint	289.4(3) K
(c)	dark patch	291.2(1) K

The difference between *(a)* and *(b)* is not untypical of the temperature differences between different parts of the Mediterranean. *(b)* and *(c)* were close together geographically and the temperature difference of nearly 1.8 deg K can quite reasonably be attributed to the skin-effect surface heating of the calm water.

There has been in the literature some discussion of the difficulties of measuring the extent of thermal heating of a thin surface layer of the sea in very calm conditions and the difficulty of measuring the magnitude of this effect in situ (see for instance Robinson et al (1984)). We submit that a considerable amount of information about this phenomenon can be obtained without recourse to in situ measurements, which are notoriously difficult to perform without disturbing the surface layer, by deliberately choosing to study this phenomenon in an area of sunglint.

5. CONCLUSION

We have shown that random errors in calculating near-surface wind speeds from AVHRR data in an area of sunglint are quite small. We have found some evidence of bias between meteorological observations of wind speeds and the sunglint-derived wind speeds but we are inclined to attribute most, or even all, of that bias to the meteorological observations. We are aware that suggestions have been made, e.g. in connection with the validation of algorithms for non-imaging active microwave instruments on SEASAT, that bias exists in the use of the formulae of Cox and Munk (particularly the relation between σ and V). However, to establish that by the method we have used would require many more sets of data to be analysed. Nevertheless, it is possible that sunglint-derived wind speeds might be used as an alternative to in-situ data for ERS-1 validation purposes.

We have investigated the surface layer heating and its relation with wind speed in areas of sunglint. We would suggest that this provides a very fruitful alternative to attempting to perform in-situ measurements of surface layer heating.

REFERENCES

1. Cox, C. and Munk, W., "Measurement of the roughness of the sea surface from photographs of the sun glitter", J. Opt. Soc. Amer., 44, pp 838-850 (1954).

2. Cracknell, A.P., "Sunglint and the study of near-surface windspeeds over the oceans". In "Microwave remote sensing for oceanographic and marine weather-forecast models" ed R.A. Vaughan (D. Reidel, Dordrecht), 1989, in press.

3. Robinson, I.S., Wells, N.C. and Charnock, H., "Review article. The sea surface thermal boundary layer and its relevance to the measurement of sea surface temperature by airborne and spaceborne radiometers", Int. J. Remote Sensing, 5, pp 19-45, 1984.

4. Saunders, R.W., Ward, N.R., England, C.F. and Hunt, G.E. "Sea surface temperature measurements around the UK derived from Meteosat and TIROS-N data". In "Matching Remote Sensing Technologies and their Applications, Proceedings of an International Conference held in London in December 1981" (Remote Sensing Society, Nottingham) pp 191-198, 1981.

APPLICATION OF SPECIAL SENSOR MICROWAVE/IMAGER DATA FOR ANALYSIS OF CYCLONIC STORMS IN MIDLATITUDES OVER THE SEA

Kristina B. Katsaros
Grant W. Petty
Iftekhar Bhatti
Douglas Miller

Department of Atmospheric Sciences
University of Washington
Seattle, WA 98195 USA

Abstract:

Microwave radiometers on satellites, operating at frequencies in the vicinity of 19, 22 and 37 GHz, can distinguish atmospheric water vapor, cloud liquid water and rain against the radiatively uniform and cold background of the ocean. In addition, preliminary algorithms also appear to promote at least qualitative fields of surface wind roughening. Since data on atmospheric weather systems are sparse over large expanses of the earth's ocean, microwave sensors could provide information to fill a serious data gap. Employing signals obtained by the Special Sensor Microwave /Imager (SSM/I) on the F8 satellite in the Defense Meteorological Satellite Program, we have recently studied the relationship between these atmospheric variables and the surface location of atmospheric fronts in midlatitude cyclones. In this note we present one representative case which illustrates the observed relationships, and we suggest ways in which the information thus obtained from the SSM/I may prove useful in the operational analaysis of oceanic storms

1. Introduction

Over the past decade radiometric measurements at microwave frequencies greater than 18 GHz have been available from satellite sensors at resolutions of 50 km or better. Two Scanning Multichannel Microwave Radiometers (SMMR's) were launched in 1978 on the Seasat and Nimbus 7 satellites (Gloersen and Barrath, 1977), and in 1987 an instrument with similar characteristics, the Special Sensor Microwave/Imager (SSM/I), was launched on the F8 satellite in the U.S. Defense Meteorological Satellite Program (DMSP).

By including channels in the frequency range 18-37 GHz (see Table 1) these instruments allow estimation of integrated atmospheric water vapor, integrated cloud liquid water and at least the *presence* of rain in the column. The integrated water vapor parameter is very robust and has the same r.m.s. accuracy as one obtained by integration of radiosonde ascents (Alishouse, 1983; McMurdie, 1989). Difficulties in interpreting the signal in terms of rain rate occur because most rain cells do not fill the approximately 30 km by 30 km footprint of the 37 GHz channels used for this purpose, and rain rates greater than about 3mm/hr would saturate this signal, if the rain covered the footprint.

In previous studies at the University of Washington, we have related patterns in integrated water vapor from the SMMR's to the surface location of fronts in midlatitude cyclones, and have examined the ability of simple algorithms to identify rain areas and estimate rain intensity in such weather systems (McMurdie and Katsaros 1985; Katsaros and Lewis 1986; McMurdie et al., 1987).

Because we found these patterns to be very characteristic, we also developed flagging routines to objectively locate frontal zones over the sea using SMMR data (Katsaros et al., 1989a). Recently, we have made preliminary tests of the same methods applied to data from the SSM/I. This instrument provides four times the data coverage of the Nimbus 7 SMMR, because it has twice the swath width and has a continuous duty cycle, while the SMMR was turned on only on alternate days due to power restrictions on the Nimbus 7 satellite. The SSM/I is therefore a more valuable data source for studying evolution of cyclones.

2. A cyclone 'seen' with the Special Sensor Microwave/Imager

In this short note we present one example of several types of information that SSM/I can provide about cyclonic storm systems over the ocean. More details are available in a technical report by the authors (Katsaros et al 1989b).

Table 1. Characteristics of the Two Polar Orbiting Radiometers, SMMR on Seasat and on Nimbus 7 and the SSM/I on the F8 DMSP Satellite

SMMR		SSM/I	
FREQ. GHz	APPROX. RESOLUTION (km)	FREQ. GHz	APPROX. RESOLUTION (km)
6.6	150	--	--
10.7	100	--	--
18	65	19.35	55
21	60	22.235	50
37	35	37	35
--	--	85.5	15

Figure 1 shows five fields determined from SSM/I signals for a storm that approached northern Europe on October 18, 1987. The frontal analysis superimposed on these SSM/I products is taken directly from the National Meteorological Center (NMC) surface map for 0:600 UTC on the same day. The contours of integrated water vapor are presented in Figure 1a. We see that the surface position of the cold front occurs at the leading edge of the strong gradient between the warm moist air ahead of the front and the cold dry air behind it. In our previous work (Katsaros et al 1989a), we found that by flagging gradients in integrated water vapor greater than a certain constant threshold value, we could identify the surface location of the fronts objectively in all regions of the world ocean, irrespective of the season. Applying a similar threshold (slightly different in magnitude, due to difference in sensor resolution we surmise) to the gradients in Figure 1a, we produce Figure 1b. The tail end of the cold front is well identified by this routine. Similarly, Figure 1c shows rain areas identified by flagging 37 GHz brightness temperatures greater than a certain threshold value. Again the frontal locations are correlated to the location of the flags, except in the occluded portion at 58N, 15W, where we expect that the rain flags may possibly have the correct answer.

Figure 1d depicts the location of depressed 85 GHz brightness temperatures resulting from scattering by large ice particles in the clouds. These large ice particles are typically associated with precipitation. An intensification in the scattering signal at about 49N, 7W could be indicative of enhanced convergence in this area. Both the water vapor pattern and the prominent region of rain behind the NMC frontal location seem to corroborate this interpretation.

Figure 1e shows surface wind speed calculated with the pre-launch algorithm documented by Hollinger et al. (1987). (Areas where heavy cloud liquid water and rain obscure the surface have been omitted.) The maximum winds of 15 to 20 m/s west of Ireland are supported by the NMC surface map, which shows a station with wind speed of 17 m/s at 54N, 15W. The NMC pressure analysis show weak gradients in the area, but stronger ones both to the north and to the south. If the algorithm is correct, the SSM/I could possibly have assisted the analyst in sharpening the pressure gradient at the location of these high wind speeds. No other information is available in the area.

3. Summary and Discussion

We see in Figure 1 that the fields of water vapor, rain, and scattering by large ice particles complement each other in the sense that they appear to give different information about the structure of the storm system. The information in each of the various fields is nearly orthogonal to the information in the other fields. We have not illustrated the quantity cloud liquid water in this short note, but it is another potentially available parameter, which however cannot be readily isolated from the rain water.

Our conclusion from looking at several hundred samples of cyclonic storms with microwave data from SMMR and SSM/I are that a) they provide a new valuable research tool for studying mesoscale structure of the various forms of atmospheric water (as vapor, cloud droplets, rain or large ice particles) in midlatitude cyclones, and b) the water vapor gradient and rain flags together provide an objective method for locating fronts over the ocean, which may have important practical value.

Difficulties occur due to the coarse resolution of the sensors and our incomplete knowledge of the physics of microwave emission from the weather systems, which makes algorithms currently in use subject to continuous revision for some time to come; the integrated amount of water vapor is a notable exception.

4. Acknowledgements

This work was performed with support from the Naval Environmental Prediction Research Facility under Contract N00014-86-K-0453 and from NASA Grant NAG 5-943. We are grateful to Frank J. Wentz and Erica Francis for providing us the SSM/I tapes and to Janet Meadows, who typed the manuscript with great care. We appreciate the continuous encouragement by Drs. Andreas Goroch and Paul Tag at the Naval Environmental Prediction Research Facility throughout this work.

5. References

Alishouse, J.C., 1983: Total precipitable water and rainfall determinations from the Seasat Scanning Multichannel Microwave Radiometer (SMMR). *J. Geophys. Res.*, **88**, 1929-1935.

Gloersen, P. and F.T. Barath, 1977: A Scanning Multichannel Microwave Radiometer for Nimbus-G and Seasat A. *IEEE Journal of Oceanic Engineering*, **OE-2**, 172-178.

Hollinger, J., R. Lo, G. Poe, R. Savage and J. Pierce, 1987: *Special Sensor Microwave/Imager User's Guide.* Naval Research Laboratory, Washington, D.C. 177 pp.

Katsaros, K.B. and R.M. Lewis, 1986: Mesoscale and synoptic scale features of North Pacific weather systems observed with the Scanning Multichannel Microwave Radiometer on Nimbus 7. *J. Geophys. Res.*, **91**, 2321-2330.

Katsaros, K.B., I.Bhatti, L.A. McMurdie and G.W. Petty, 1989a: Identification of atmospheric fronts over the ocean with microwave measurements of water vapor and rain. *Weather Forecasting*, (submitted).

Katsaros, K.B., G.W. Petty, I. Bhatti and D. Miller, 1989b: Application of Special Sensor Microwave/Imager data for analysis of cyclonic storms in midlatitudes over the sea. *Second Annual Report on Contract N00014-86-K-0453*, Nov. 1988. Dept. of Atmospheric Sciences, University of Washington, Seattle, WA 98195.

McMurdie, L.A., 1989: *Interpretation of Integrated Water Vapor Patterns in Oceanic Midlatitude Cyclones Derived from the Scanning Multichannel Microwave Radiometer.* Ph.D. Thesis, Dept. of Atmospheric Sciences, University of Washington, Seattle, WA 98195.

McMurdie, L.A. and K.B. Katsaros, 1985: Atmospheric water distribution in a midlatitude cyclone observed by the Seasat Scanning Multichannel Microwave Radiometer. *Mon. Wea. Rev.*, **113**, 584-598.

McMurdie, L.A., G. Levy and K.B. Katsaros, 1987: On the relationship between scatterometer-derived convergences and atmospheric moisture. *Mon. Wea. Rev.*, **115**, 1281-1294.

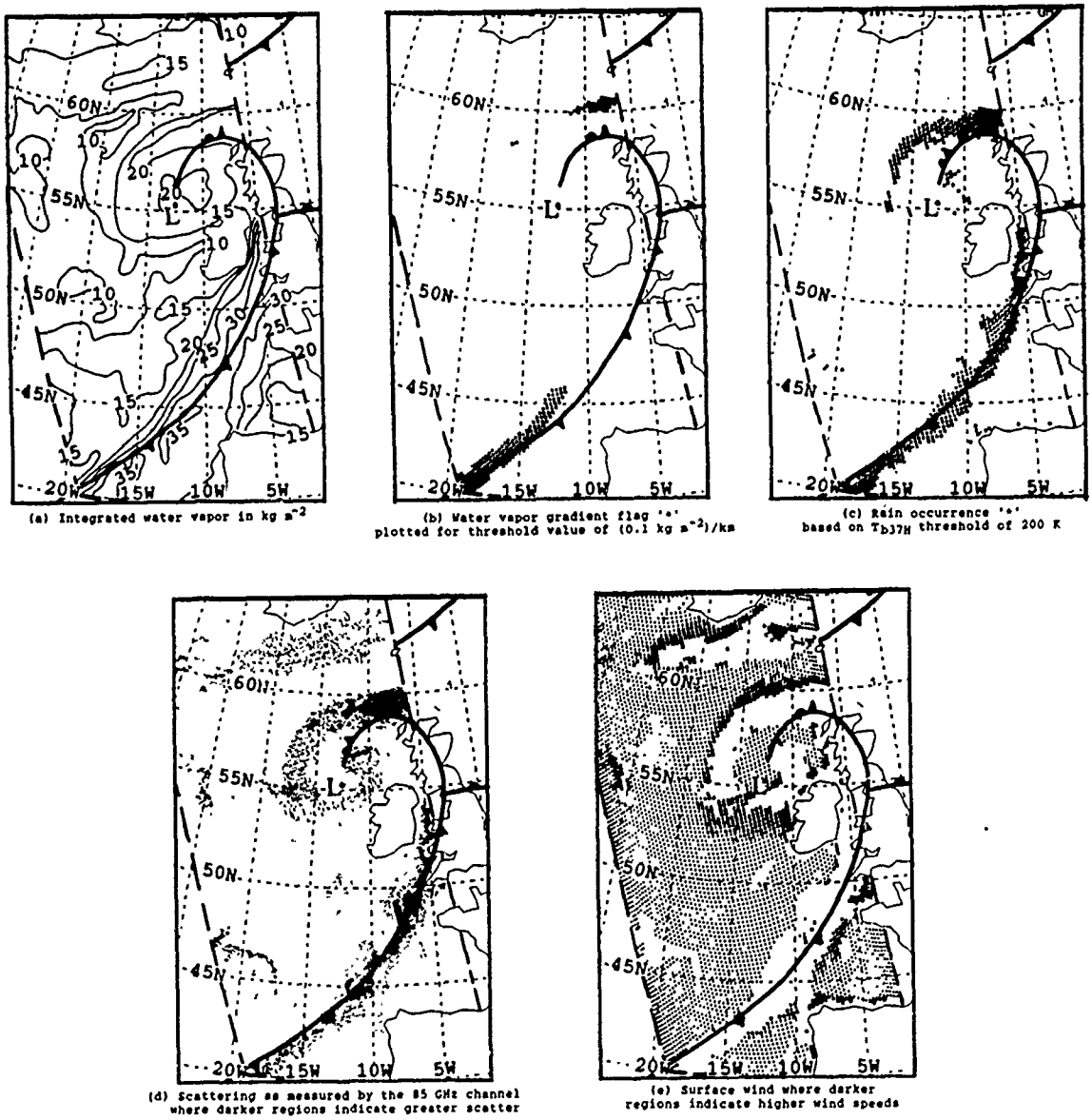


Figure 1: Occluded cyclone observed by DMSP F8 SSM/I, October 18, 1987 at 06:24 UTC, Orbit #1697. The superimposed frontal positions have been obtained from the NMC analysis at 06:00 UTC. In Figures (d) and (e) darker symbols indicate higher values.

RADIAN
CORPORATION

ESTIMATION OF ATMOSPHERIC THERMAL LAPSE RATES USING CONVENTIONAL ACOUSTIC RADAR MEASUREMENTS

John E. Letkeman, K. Russell Peterman
Radian Corporation
Austin, Texas 78720-1088

FAX (512) 454-7129
PHN (512) 454-4797

Conventional acoustic radar (SODAR) technology is commonly used to measure vertical and horizontal wind speed, direction, atmospheric turbulence, and an estimation of atmospheric stability. For the measurement of atmospheric temperature, radio waves are required, often in conjunction with acoustic radar. SODAR measurements are therefore not applicable to air quality dispersion models without substantial interpretations by a meteorologist. In particular, the thermal lapse rates of stable atmospheres are not directly measurable by SODAR. Plume rise is very sensitive to thermal lapse rates of stable layers and in many cases, the plume rise in relation to a stable layer may mean the difference between fumigation and lofting.

By simultaneous SODAR and radiosonde pilot balloons measurements, we are correlating various SODAR measured parameters to the thermal lapse rates measured by balloon. These correlations provide relationships to thermal lapse rates which have a physical basis. Indications of lapse rates are measured turbulence in the forms of vertical and horizontal wind variance, wind speed shear, wind direction shear, echo intensities, time of day, and solar radiation intensities. The above parameters are correlated to lapse rate and evaluated as to the effectiveness of the correlations. Data from an ECHOSONDE® SODAR in the Seattle, Washington area and other locations are used in developing the relationships. The correlations are a basis for a procedure for the estimations of thermal lapse rates, inversion base and top elevations, and stability estimates.

The method of arriving at a reliable relationship between SODAR measurements and thermal structure variables is an empirical one. As an empirical relationship, the method tends to be site-specific and not necessarily transferrable. With further study and data, however, significant transferrable relationships will probably become available. Speculation is made about methods of transferring this method from one site to another, and the feasibility of this approach.

A COMPARISON OF THE INDIVIDUAL AND COMBINED PERFORMANCE OF GROUND-BASED RADIO-ACOUSTIC AND RADIOMETRIC TEMPERATURE SOUNDING SYSTEMS

Judith A. Schroeder

NOAA/ERL/WPL
325 Broadway
Boulder, CO 80303
303-497-6304

ABSTRACT

An experiment was conducted to compare the accuracy of atmospheric temperature soundings obtained simultaneously with a ground-based radio-acoustic sounding system (RASS), a ground-based microwave radiometric system, and a combined RASS/radiometric system colocated at Stapleton International Airport in Denver, Colorado during July and August, 1988. Accuracy was defined as agreement with simultaneous in situ rawinsonde temperature measurements.

Rms errors for the 22 sets of soundings compared ranged from 0.5 to 2.0°C, depending on altitude. Temperatures retrieved from the radiometer measurements were more accurate than those retrieved from RASS above 5 km and below 0.5 km, where a bias in the low-altitude RASS measurements degraded the retrieved temperatures. In between, the RASS temperature errors were less than 1°C. At the 750 and 700 mb mandatory levels, the RASS errors matched the reported precision of rawinsonde measurements. Although the combined system errors were smaller than individual system errors at all altitudes above 0.5 km, the differences among the three systems were too small to conclude that the combined system performance was significantly better than that of either system alone.

KEY WORDS: RASS, microwave radiometry, temperature retrieval

1. INTRODUCTION

Knowledge of atmospheric temperature as a function of height is necessary for accurate weather forecasts, which directly impact aviation, agriculture, and human safety. Currently, these temperature soundings are obtained at 12-h intervals with balloon-borne instrument packages called rawinsondes. More frequent observations are needed to monitor the evolution of local weather phenomena, whose lifetimes fall between rawinsonde launch times.

The remote sensing community has responded to this need by developing sensors that measure atmospheric properties (e.g., microwave emission) from which temperature soundings can be derived. This paper shows results of a comparison among temperature soundings obtained with three ground-based remote sensing systems colocated with a rawinsonde launch site at Stapleton International Airport in Denver, Colorado during July and August of 1988. Rawinsonde temperature measurements served as the comparison standard.

2. MEASUREMENTS

The three remote sensing systems compared were a radio- acoustic sounding system (RASS), a microwave radiometric system, and a combined RASS/radiometric system.

The RASS system consisted of a 915 MHz Doppler radar and a vertically-pointing, 50-watt acoustic source. The radar, which normally measures profiles of wind speed and direction, was instead used to track the speed of an acoustic wavefront transmitted by the acoustic source (May et al., 1989). The sound speed obtained can be related back to virtual temperature (List, 1963). Large acoustic attenuation at the frequencies involved limited the altitude range of the RASS virtual temperature measurements to about 2 km AGL (May et al., 1988). Vertical wind velocity can also degrade the RASS measurements.

The six-channel microwave radiometric system measures radiation emitted by the atmosphere from the zenith direction. Two channels, 20.6 and 31.65GHz, respond to changes in water vapor and cloud liquid water content. The other four channels, 52.85, 53.85, 55.45, and 58.80 GHz, respond pri-

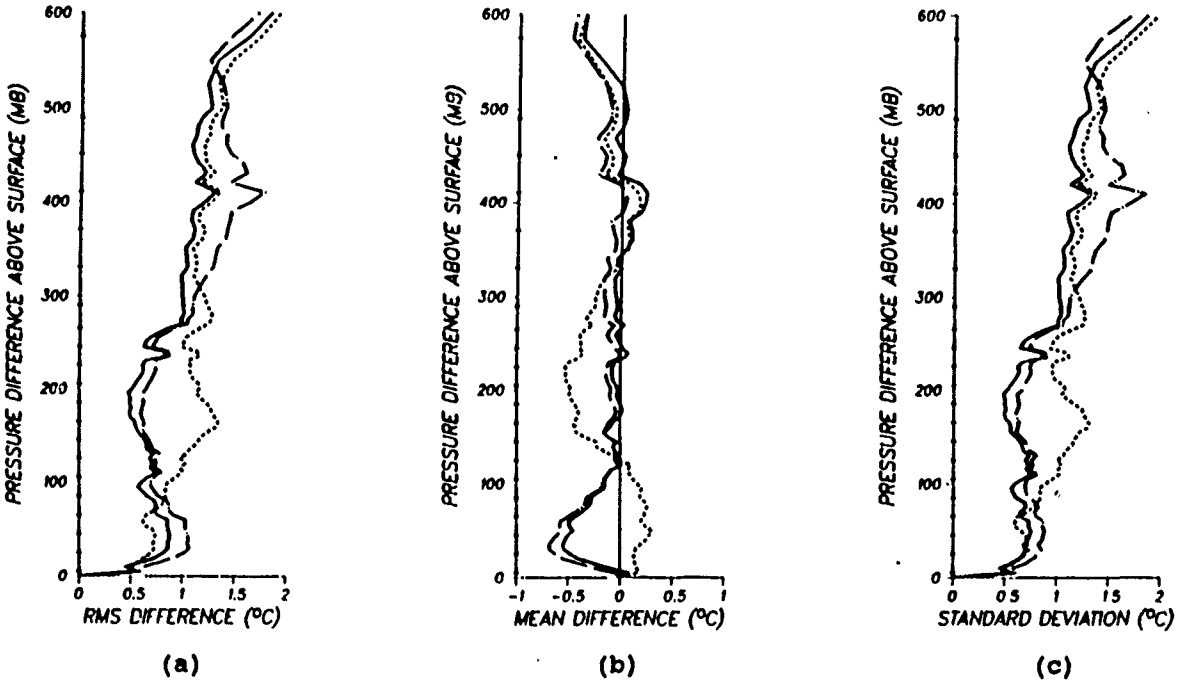


Figure 8. Comparison of differences between rawinsonde-measured temperatures and those retrieved from a radiometric (---), RASS (-•-) and combined RASS/radiometric (——) system. Dots on y-axis represent every fourth level compared.

marily to temperature changes, although some water sensitivity exists in the lower two channels. Rapidly changing oxygen absorption over this frequency range permits detection of temperature changes at differing altitude ranges for each channel, ranging from the surface to about 7 km AGL.

Rawinsonde temperature soundings were provided by the National Weather Service and the National Center for Atmospheric Research (Lauritsen et al., 1987). Surface temperature measurements, provided with the soundings, were considered part of all three remote systems.

4. RETRIEVAL TECHNIQUE

Since neither RASS nor radiometers measure temperature directly, a linear statistical retrieval technique (Strand and Westwater, 1968) was used infer the temperature soundings from the measurements. This technique uses a history of rawinsonde data from the instrument locality to establish a linear relationship between measurements simulated from the rawinsonde data and atmospheric temperatures at selected height or pressure levels. The coefficients obtained by regressing the temperatures on the simulated measurements are then applied to the actual measurements to produce a temperature estimate at each level.

For this comparison, 85 tropospheric pressure levels were chosen to preserve the resolution of both measurement systems. Although the RASS and radiometric measurements represent atmospheric conditions below about 2 and 7 km, respectively, applying the retrieval technique to measurements from the two systems and the combined RASS/radiometric system permitted three-way comparisons throughout the troposphere.

4. DISCUSSION OF RESULTS

During the experiment, 22 sets of simultaneous rawinsonde, radiometer, and RASS measurements were obtained.

The root mean square (rms) difference, mean difference, and the standard deviation of the difference between rawinsonde temperature soundings and those retrieved from each of the three systems appear in Fig. 1 as a function of pressure difference above the surface (AGL). The mean surface pressure for this set of soundings was 842 mb, with a standard deviation of 4 mb. Subtracting the vertical coordinate given in the figure from 842 gives an approximate value for the absolute pressure at the level in question. The dots along the vertical axis, showing every fourth level retrieved, illustrate the decreasing resolution with height chosen for the retrieved temperature levels. The heights of the RASS measurements used in this study fell between approximately

20 and 250 mb AGL, although most of them fell below about 180 mb AGL. In all three graphs, the dashed curve represents the radiometric system, the dot-dash curve represents the RASS system, and the solid curve represents the combined system.

The radiometric system temperatures agree better with the rawinsonde at levels above and below the RASS measurement altitudes. Within the RASS measurement region, the RASS temperatures are dramatically closer to the rawinsonde temperatures. In fact, comparisons at mandatory levels revealed that the difference between RASS and rawinsonde at 750 and 700 mb was within the rawinsonde temperature measurement precision reported by Hoehne (1980).

The combined system appears to outperform both single systems everywhere above the first 60 mb above the surface. Below this level, Fig. 1b shows a large bias in the RASS temperatures that persists in the combined system temperatures. This bias, which is probably due to ground clutter affecting the radar, has since been corrected (R. G. Strauch, 1989).

Multivariate statistical tests indicated that the accuracies of the three systems shown in Fig. 1 are not significantly different ($\alpha = 0.5$). This similarity in performance raises some economic questions. The RASS and radiometric systems cost about the same, so a combined system costs about twice as much as either system alone. Further, plans already exist for the deployment of a network of wind-profiling Doppler radars in the central United States (Chadwick and Hassel, 1987). Adding the capability of RASS operation to existing radars would cost considerably less than a six-channel radiometric system. A RASS system involving these radars (404.37 MHz) would suffer less acoustic attenuation than the one used in this experiment, which should result in more accurate temperature estimates at higher altitudes.

ACKNOWLEDGMENTS

This work was supported in part by the U.S. Customs Service under interagency agreement 1X890084, and we are grateful for the assistance of Mr. Rajiv Singh and Mr. Thomas Henneberger.

I heartily thank the following groups of people for their contributions to the research presented here:

R. H. Beeler, M. T. Decker, and J. R. Jordan of the WPL Thermodynamic Profiling Program provided the radiometer measurements. M. J. Falls compiled the rawinsonde data used in the retrieval process.

R. G. Strauch, P. T. May, K. P. Moran, and D. A. Merritt of the WPL Wind Profiler Research Program provided the RASS measurements.

Rawinsonde data were provided by R. McBeth and J. Murphy of the National Center for Atmospheric Research and by the National Weather Service.

REFERENCES

Chadwick, R. B. and N. Hassel, 1987. Profiler: the next generation surface-based atmospheric sounding system. Third International Conference on Interactive Information and Processing Systems for Meteorology, Oceanography, and Hydrology, Jan. 12–16, 1987, New Orleans, La. American Meteorological Society. Boston, Mass: 15–21.

Hoehne, W. E. 1980. Precision of National Weather Service upper-air measurements. NOAA Tech. Memo. NWS T&ED-16.

Lauritsen, D., Z. Malekmadani, C. Morel, and R. McBeth. 1987. The Cross-chain Loran Atmospheric SoundingSystem (CLASS). Sixth Symposium Meteorological Observations and Instrumentation, Jan. 12–16, 1987, New Orleans, La. American Meteorological Society, Boston, Mass., 340–43.

List, R. J. 1963. Smithsonian Meteorological Tables. Washington, D. C.: Smithsonian Institution. 527 pp.

May, P. T., R. G. Strauch, and K. P. Moran. 1988. The altitude coverage of temperature measurements using RASS with wind profiler radars. Geophys. Res. Lett. 15 (no. 12): 1381–84.

May, P. T., R. G. Strauch, K. P. Moran, and W. L. Ecklund. 1989. Temperature sounding by RASS with wind profiler radars: a preliminary study. Submitted to IEEE Trans. Geosci. Remote Sensing.

Strand, O. N. and E. R. Westwater. 1968. Minimum rms estimation of the numerical solution of a Fredholm integral equation of the first kind. SIAM J. Numer. Anal. 5: 287–295.

Strauch, R. G. 1989. Personal Communication.

MEASURING WINDS AND RAINFALL FROM SPACE—THE RADAR WIND SOUNDER (RAWS)

R. K. Moore and W. Xin

Radar Systems and Remote Sensing Laboratory
University of Kansas Center for Research, Inc.
2291 Irving Hill Road
Lawrence KS 66045-2969, USA
913/864-4835
FAX: 913/864-7789, TELEX: 706352, OMNET: KANSAS.U.RSL

ABSTRACT

Measurement of winds aloft throughout the world is of great importance because computerized global weather models use these winds as major inputs. The scarcity of such measurements over many parts of the world led to the development of the laser wind sounder (LAWS) for use on Eos. LAWS depends on the Doppler frequency shift of optical-wavelength backscatter from atmospheric aerosols, but it cannot be used where clouds interfere. We have conceived a radar wind sensor (RAWS) to fill these gaps in LAWS coverage by measuring the Doppler shift of backscatter from cloud and rain drops.

RAWS must have enough sensitivity to measure Doppler shift of cloud drops, so it will also have more than adequate sensitivity for measuring rain rates around the world. With the scan pattern proposed, RAWS will also obtain multiple looks at backscatter from the ocean, so it can be used as a surface-wind-vector scatterometer as well.

The proposed system will use a circular scan pattern. This permits observation of each point from two directions. By adding a second beam at a different depression angle, we can obtain two more measurements. These four measurements of the hydrometeors allow some overdetermination of the wind vector. Four looks will also be obtained for the surface-wind-vector scatterometer.

The major obstacles to use of RAWS are ambiguity in Doppler frequency measurement, need to rotate a large antenna (8-m diameter used in simulations), and need for relatively high power (1-5 kW peak) to measure cloud backscatter. The system, of course, cannot on a single satellite obtain coverage of the globe on a continuous basis at all points, but it can obtain much more information about winds in the tropics and polar regions than is now available.

INTRODUCTION

Meteorologists need global measurements of winds to improve modeling of atmospheric circulations. Atlas and Korb [1981] stated these needs as:

1. "Wind measurements are more important than temperature soundings wherever the winds are not in balance with the mass field. This means that they are required to faithfully predict the smaller scale systems at all latitudes and all scales in the tropics. For the larger scales in mid latitudes, temperature data are probably more important provided they are accurate to about ± 1°.

2. Even relatively crude satellite wind data (e.g., CTWs) show major real differences (from analyses lacking such data) at the higher levels in the tropics and Southern Hemisphere, and smaller but significant differences at higher latitudes.

3. Forecast simulations using wind data in the tropics and surface wind data over the oceans show significant increases in predictive skill.

4. GCM simulation studies show that an rms wind error of 2 m/s is equivalent to an rms temperature error of about 1° C outside of the tropics. Thus, a wind system that could achieve such accuracy would be equivalent to the best that is possible by any passive temperature measurement system now available or under consideration."

The Laser Atmospheric Wind Sensor (LAWS), one of the primary facility instruments for the Eos (Earth Observing System) polar platform, aims at providing winds aloft as called for. A radar wind sensor (RAWS) would serve the important goal of filling gaps created by heavy cloud cover and precipitation in the coverage of LAWS and passive temperature sensors. These areas, such as typically cloudy tropical oceans and the wintertime North Atlantic, are important regions where the RAWS can improve measurements significantly. Moreover, the same instrument can be used for ocean-surface wind-vector measurement by scatterometry, thereby filling another gap without the need for a separate instrument. The RAWS would also allow rainfall measurement.

APPROACH

The Laser Atmospheric Wind Sensor (LAWS) planned for Eos will monitor winds by measuring the Doppler shifts of lidar returns from aerosols in cloud-free areas [Curran, 1987]. Since the lidar will be scanned in a circular pattern, each point within the circle will be viewed from two different directions, so two components of the velocity can be measured.

Our proposed Radar Wind Sensor (RAWS) would use a similar scanning pattern to measure the winds in the cloudy areas that cover much of the globe. Water drops of small mass respond quickly to horizontal wind

forces, so their motion should trace the wind field. The radar would measure wind by determining the Doppler shift in the backscatter from rain and cloud particles moved by wind. More than 90% of the rms wind fluctuation should be measurable from the speed of water droplets having diameter under 3 mm (Doviak & Zrnic, 1984).

Frequencies in X band or K_u band are appropriate for precipitation, but clouds require higher frequencies because cloud drops are very small. More power is needed for clouds than for rain alone, but this will be available in future unmanned spacecraft with robust power sources. The power required is comparable with that for a SAR. A scaled-down RAWS that would work only in areas with rain exceeding 0.5 mm/hr and would not profile winds in clouds could be built with a power level of the order of 100 W peak and a much smaller antenna.

A radar with enough sensitivity to measure Doppler shifts in rain would also be sensitive enough to measure rain rate and surface winds over the oceans. Thus the RAWS could provide capabilities to:

a. Supplement LAWS by measuring mean Doppler frequency to determine winds aloft in cloudy regions.

b. Provide rainfall-rate measurements by measuring signal strength in rain.

c. Provide measurements in lieu of or in addition to those associated with a specialized ocean-surface wind-vector scatterometer such as the SCATT facility instrument planned for Eos.

THE RAWS CONCEPT

The RAWS would use a circular scan pattern like that of LAWS and the one proposed for a surface-wind scatterometer (Kirimoto and Moore, 1985; Moore et al., 1988). With a single scanning beam, each point on the surface or in the cloud would be seen from two directions. However, use of another beam at a different angle of incidence, as in Kirimoto and Moore, would permit measurements from four directions for each point. The scan pattern on and near the surface would take the form of concentric circles as shown in Fig. 1. Of course, forward motion of the satellite converts these into spirals. We could achieve the dual beam by dual feeds in a reflector antenna or by a suitably designed array. We propose that the angles of incidence on cloud and surface be 30° and 35°. The four looks at each point provide redundancy to improve accuracy.

The Doppler frequency is given by

$$f_D \approx 2(u_s - u_h)\cdot R/R,$$

where u_s is the spacecraft velocity vector and u_h is the hydrometeor velocity vector, while R is the vector from the spacecraft to the volume observed. The scalar product is different for each viewing angle. Hence one may solve three simultaneous equations to obtain the three components of u_h, presuming that u_s is known. The fourth measured Doppler frequency results in overdetermination of u_h that can improve the estimate of its value. Thus, both the wind and the fall rate of raindrops may be determined.

The range resolution will be adequate to allow profiling the winds to a height resolution better than 1.7 km. This depends on vertical beamwidth of the antenna

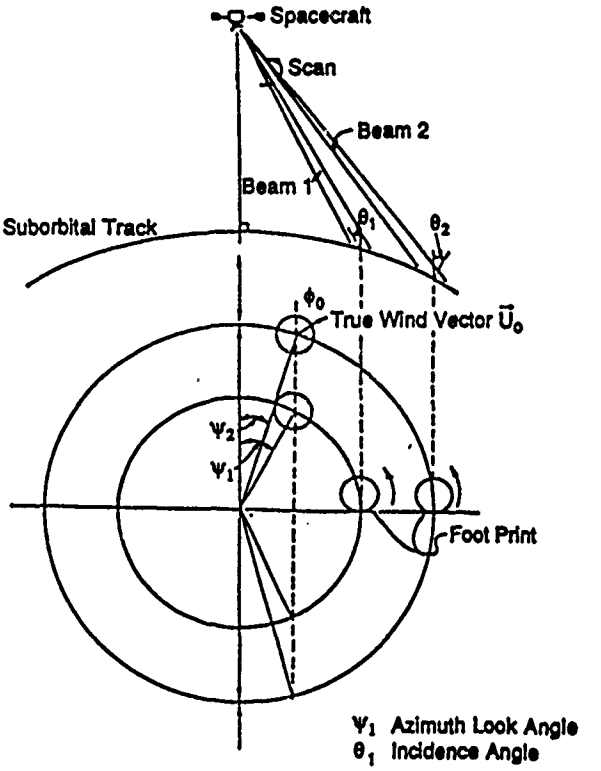


Fig. 1 Scanning pattern for RAWS. Each point within the inner circle is observed from four directions (± Ψ1 and ± Ψ2). Points outside the inner circle are seen only from two directions. The ground pattern is a spiral because of satellite motion.

and maximum angle away from vertical (35° in the reference design). For 35° at the surface, an 8-m antenna at 35 GHz gives about 1300-m vertical resolution and at 10 GHz about 2300 m, assuming use of a 2 μs pulse. For hydrometeors where attenuation does not eliminate use of 35 GHz, therefore, 1300-m resolution can be achieved at 35°.

Parameters proposed for RAWS in our preliminary studies are shown in Table 1.

Table 1. Preliminary System Parameters for RAWS

Frequency (GHz)	10	35.5
Orbit altitude (km)	400	
PRF (Hz)	4000	
Peak transmitted power (W)	1000	
Chirp gain	12.5	
Compressed pulse width ($s)	2	
Antenna aperture (m high x m wide)	4 x 8	
Antenna gain (dB)	56	67
Beamwidth (mrad vert x mrad hor)	7.5x3.75	2.10x1.06
Footprint (km radial x km cross)	1.8x2.2	0.50x0.62
Vertical resolution (km)	2.3	1.3
Doppler (satellite motion)(max MHz)	0.29	1.02
Bandwidth due to beamwidth (kHz)	1.88	
Radial velocity resolution (m/s)	1	

TECHNICAL CHALLENGES

The most difficult technical challenges in the RAWS design are:

1. Obtaining adequate sensitivity for clouds,

2. Determination of Doppler centroid,

3. Provision of a suitable antenna scan.

POWER AND SENSITIVITY

The key power problem is for clouds. A radar such as this would be useful if it only measured winds in rain, but its value would be much greater if winds in the interior of clouds could be measured. If we have enough sensitivity for clouds, the signal-to-noise ratio (SNR) will be high for rain (neglecting attenuation, which we also considered), and very high for surface backscatter.

We calculated power requirements using standard cloud models for a few simple examples. Figure 2 shows the SNR for a water cloud with mean drop radius of 10 μm vs depth in the cloud, using the parameters of our reference design. Clearly 10 GHz is not useful for clouds. The SNR exceeds 0 dB at 35.5 GHz for depths

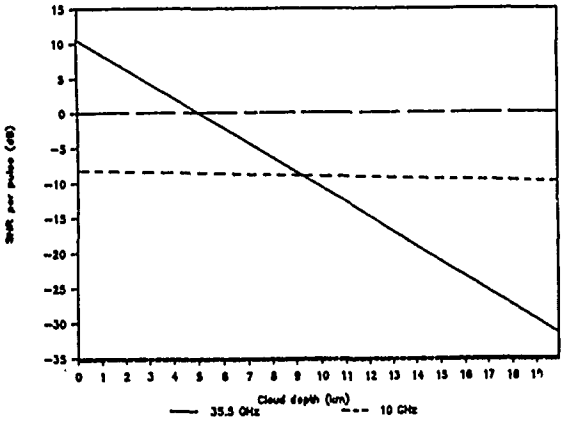


Fig. 2 Calculated SNR for RAWS with cloud having 100 drops/cm^3 and density of 0.95 g/m^3.

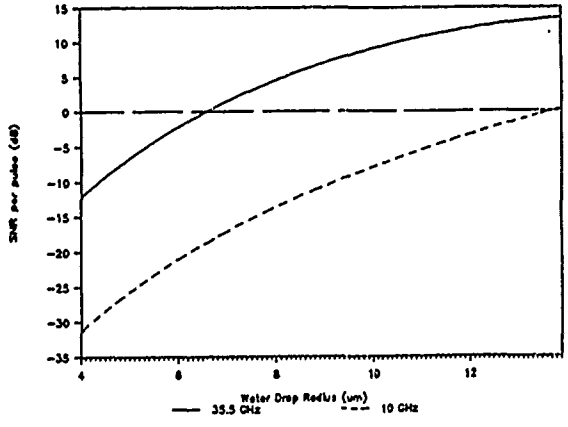


Fig. 3 Calculated SNR for RAWS at depth of 1 km in cloud at 0° C and 100 drops/cm^3 vs mean drop radius.

up to 5 km into the cloud, which should be useful. For thinner clouds with smaller drops, the RAWS would not work without excessive transmitter power. Figure 3 illustrates this by showing the effect of mean drop radius on SNR.

Even though 10 GHz is not useful for clouds, it is required for heavy rain, where the attenuation is significantly more. This was demonstrated in the JPL TRMM-radar study [Im et al., 1987].

DETERMINATION OF DOPPLER CENTROID

To measure the Doppler centroids for clouds and rain we must have a reference. The Doppler centroids for surface echoes can provide this. Otherwise the small Doppler shifts due to wind would be lost in changes in the Doppler shifts due to spacecraft attitude changes. Hence, RAWS must measure both surface and hydrometeor echoes and use the difference in their centroids to estimate the wind velocity.

The Doppler frequencies associated with the turbulent motion and the spread of the beam form a spectrum of considerable width. The RAWS must measure the centroid of this spectrum and remove the effect of spacecraft motion. Methods for measuring the Doppler centroid for rain echoes with ground-based radars have been widely treated [Doviak and Zrnic, 1984]. Preliminary simulations indicate that the pulse-pair and traditional FFT method both have promise.

At 30° and 10 GHz the center frequency of a stationary surface echo would vary from -250 kHz through zero to +250 kHz during a circular scan. Since orbital parameters are well known, a local oscillator can be programmed to estimate the center frequency. Even so, the tracking problem would not be trivial, since the center frequencies for the surface and rain spectra must be located to within about 0.3% of their bandwidth or .027% of their absolute value. Hence, tracking of the surface echo is essential to compensate for small errors in orbital and attitude determinations.

One approach we are considering would use a CW signal near the 10-GHz pulse-radar frequency. Because the bandwidth for this would be narrow, it could use very low power. Except for the very heaviest rainfall, this signal would be dominated by surface return, so it could serve to measure the effect of pointing and position uncertainties. For this we are considering a third-order phase-locked loop and a Kalman filter.

Another problem in Doppler-frequency measurement is ambiguity. To avoid range ambiguities, the PRF cannot exceed 4 kHz. At 35 GHz, one needs a sampling rate of 12 kHz to satisfy the Nyquist sampling criterion. We are considering several approaches to solve this problem, which is common in ground-based weather radars [Doviak and Zrnic, 1984]. One method is the staggered-PRF approach used in MTI radars [Ludloff and Minker, 1985]. Another approach is to transmit successive pulses on adjacent, but independent, frequencies. Other solutions being studied are phase diversity [Zrnic and Mahapatra, 1985; Sachidananda and Zrnic, 1986] and polarization diversity.

ANTENNA

RAWS calls for a large antenna both to achieve suitable veritcal resolution at 35° and to keep power requirements within practical ranges. Antennas as large as 15-m diameter have been designed for use with frequencies up to K_u band [Campbell & Belvin, 1985]. A large array may also be used.

Once an antenna has been selected, one must consider a scan rate that will allow adequate dwell time to achieve the required measurement precision. We have

considered this problem, but do not discuss it here for reasons of brevity. The other part of the scan problem is to achieve the four looks at each point in a grid suitable for forecasting. To combine the scan rate problem with this geometric problem will require analysis yet to be done.

CONCLUSION

Preliminary analysis and simulation show that a radar on a low-earth-orbit satellite can be used to measure winds aloft, rain rates, and at the same time measure surface scattering needed to obtain estimates of surface winds at sea. This radar wind sensor (RAWS) would be a useful complement to the laser sensor (LAWS) planned for Eos. Combining the two would allow sampling the wind fields both in the clear areas accessible to LAWS and the important cloudy ones accessible only to RAWS.

With a peak power of 1 kW, RAWS can measure in fairly dense clouds (10 μm mean drop radius) and in essentially all rains. It requires a large antenna, but not one so large as to be infeasible. Major technical problems include several dealing with Doppler-centroid measurement and tracking the surface echo and others related to the scan pattern. The latter were not treated here.

REFERENCES

Atlas, D. & C. L. Korb (1981), "Weather and Climate Needs for Lidar Observations from Space and Concepts for Their Realization," Bull. Amer. Met. Soc., vol. 62, p. 1270.

Campbell T. G. & W. K. Belvin (1985), "The Development of the 15-meter Hoop column Deployable Antenna System with Final Structural and Electromagnetic Performance Results," Digest IGARSS'85, pp. 658.

Curran, R. J., ed. (1987), Earth Observing System Instrument Panel Report: LAWS Laser Atmospheric Wind Sounder, NASA, Washington DC, vol. IIg.

Doviak, R. J. & D. S. Zrnic (1984), Doppler Radar and Weather Observation, Academic Press: New York.

Kirimoto, T. & R. K. Moore (1985), "Scanning Wind-Vector Scatterometers with Two Pencil Beams," Proc. Conf. Frontiers of Remote Sensing of Oceans and Troposphere, Shoresh, Israel, NASA Conf. Pub. 2303, 89, 1984, & RSL TR 0176-1.

Im, K. E., F. Li, D. Rosing, & W. J. Wilson (1987), Phase A Conceptual Design for a Rain Radar on the Tropical Rain Mapping Mission: Final Report, Jet Propulsion Lab, Pasadena CA, Feb. 6.

Ludloff, A. & M. Minker (1985), "Reliability of Velocity Measurement by MTD Radar," IEEE Trans. on Aerosp. & Elec. Sys., pp. 522-528.

Moore, R. K., F. K. Li & R. G. Kennett (1988), "Performance of A Scanning Pencil-Beam Spaceborne Scatterometer for Ocean Wind Measurements," Digest IGARSS'88, Edinburgh, Scotland, Sep., pp. 563-564.

Sachidananda, M. & D. S. Zrnic (1986), "Recovery of Spectral Moments from Overlaid Echoes in a Doppler Weather Radar," IEEE Trans. Geosc. & Rem. Sens., vol. GE-24.

Zrnic, D. S. & P. Mahapatra (1985), "Two Methods of Ambiguity Resolution in Pulse Doppler Weather Radars," IEEE Trans. Aerospace Elec. Sys., vol. AES-21.

BOUNDARY LAYER CHARACTERISTICS PREVAILING OVER CALCUTTA

J. Das, A.K. De and A. Ganguli

Electronics & Communications Science Unit
Indian Statistical Institute
Calcutta 700035, INDIA.

I. INTRODUCTION

Because of the varied manifestations and applications, the dynamics of air and water-vapour mixture have received lot of attentions. Over the last two decades, numerous research efforts were tried to show that both large and small scale wave motions have played a significant role in determining the circulation and structure of the middle atmosphere. Theoretical and observational studies have performed to establish that the gravity waves and tidal motions are able to transport momentum and energy over considerable distance (Hines, 1960). Such motions may produce a variety of important effects as they can attain a large amplitudes in the middle atmosphere. Recent use of very high resolution microwave and acoustic radars have shown the existence of a wide variety of fine scale turbulent scatter layers. The scatter layers often undulate in response to stable gravity waves.

Gravity waves have been observed to cause short term tilting of turbulent layers (Hooke and Hardy, 1975) and specularly reflecting layers. In both the cases, the effect is a wavelike deformations of atmospheric layers with a period of a few minutes accompanied by a varying amplitude and frequency. Experiment with constant pressure balloons (instrumented) have provided significant results on the morphology of gravity waves and turbulence as well as their consequences for mixing in the stratosphere. The direct evidence of turbulence due to gravity wave motions in the lower atmosphere is provided by observations of Lilly and Lester (1974). They noticed that the wave amplitudes as well as turbulence intensities increased to a maximum near 17 kms, corresponding to a level of minimum mean wind.

Analysis of sodar echograms for a period of one year (April 1986 - March 1987) shows the occurrence of wave-motions (simple/complex) in planetary boundary layer at Calcutta. These oscillations are mostly occur in the boundary layer under conditions of nocturnal radiative inversion, ascending inversion and frontal wind shear. The scope of the present paper is limited over studying the various types of waves observed on sodar echograms under varying atmospheric conditions at Calcutta. An attempt has been made to correlate some of the characteristic features of a few days wave perturbations with that of the surface data prevailing during that time. Spectral analysis and autocorrelation techniques have used to calculate the frequency and periodicity.

II. SODAR SIGNATURES OF BOUNDARY LAYER BEHAVIOURS

The acoustic sounding technique provides interaction of acoustic waves with the wind velocity and thermal turbulent structure of the lower atmosphere. Remote sensing technique that offers continuity in both space and time has been extensively used to probe the microstructure of the planetary boundary layer. The continuously operated monostatic sodar at ISI indicates the various micrometeorological and meso-meteorological conditions surrounding the sodar site. Percentage of occurrence of various atmospheric structures over one year period (April 1986 - March 1987) is shown in fig.(1). The stable structures occur normally at night and is about 66% and the unstable structure which occur mostly after sunrise is about 28%. The mixed structures of plumes capped by an inversion, representing the morning transition, is about 4% only.

Fig.1. Percentage occurrence of various atmospheric structures.

The morning echogram shows a surface based layer due to solar heating of the ground. Slowly build-up convective wind develops a mixed boundary layer which erodes the morning stable inversion upwards from the surface. With continued solar heating, the height and frequency of occurrence of convective plumes increase and become maximum around the noon. The plumes disappear completely within dusk and short-ranged strong echoes having abrupt but almost uniform upper limits begun to appear. This type of stable structure were mostly observed within the height range of 200-300 m above ground. Such inversions are formed largely by the transfer of heat down towards the ground which is being cooled radiatively. During night predominantly horizontal echo regions (shear echoes) impervious to vertical transport are observed. The night time atmospheric structures have a high degree of variabilities showing presence of ground-based layer of variable thickness, the appearence of highly variable multiple layers, the wave-motions within and above the inversion and turbulence structure. All these contribute towards the maximum atmospheric variability at night.

The formation of waves enable one in studying the dynamics of boundary layer. Waves and perturbations [fig.2] have been noticed many a times on the sodar echograms under stable atmospheric conditions. These undulations are mostly quasiperiodic in

structure of simple and complex type and have large periods. The amplitude of these waves are generally around 80 meters peak to peak. These are observed both at the top of the surface based layer or superposed on the elevated layer occuring mostly during the late night. The wave motions may characterise calm or light wind conditions [fig.2(b)], or with herringbone type [fig.2(c)] and for strong wind shear [fig.2(d)].

Fig.2. Sodar structure showing wave motions in a stable condition.

III. SODAR CHARACTERISTICS OF WAVE PERTURBATIONS

Gravity waves are the low frequency transverse atmospheric waves. These waves owe their existence, mathematically, to the presence of gravity terms in the equation of motion. Although our knowledge of the climatological conditions of gravity waves is very limited, there is enough evidence to suggest that they are present for a substantial span of time in the atmosphere. Analysis of sodar echograms observed over Calcutta, during the period of April '86 - March '87 reveal some characteristics of these wave perturbations and we have observed 12 occasions of wave like structure superimposed on the stable nocturnal radiative inversion. These perturbations over Calcutta are mostly observed around mid-night and in early morning hours and are always associated with ground-based or elevated inversions. Also it has observed that they are generally formed during the winter months when the strong nocturnal radiative inversions are developed and in premonsoon season when a large number of thunderstorm/norwesters are usually prevail over this coastal station.

Of the different categories of oscillating phenomena recorded on sodar echogram the oscillations of wave motions in the atmospheric boundary layer can be classified as due to (i) nocturnal radiative inversion, (ii) ascending inversion, and (iii) a front that occurs during the passage of a dustorm, thunderstorm, tornado and severe fog cloud etc. For meteorological significance, it is necessary for gravity waves to propagate horizontally without rapidly transmitting their energy to a great heights. Theoretically this requires more than a low level inversion. Rather, a thermally stable lower troposphere must be bounded above by a statically unstable region that reflects vertically propagating energy. Further, there must be a level in the unstable region where the wind velocity is approximately equal to the phase speed of the ducted wave. Several emperical studies have related gravity wave like phenomena to thunderstorm initiation. The most convincing example has been given by Uccellini(1975).

Regarding the statistics of the gravity wave like motions [fig.2] it has been observed that out of 12 cases, about 30% are noticed on thunderstorm days in this region. Most of these perturbations are associated with single frequency but there are few cases when multiple frequencies are involved. On thunderstorm days, the amplitude and time-period (12-15 minutes) are of greater value. Generally the periodicity vary from 4 to 15 minutes in this region. Usually these waves are appeared on sodar echogram for a small duration typically 3-4 cycles only. Some of these structures are correlated with the surface data for wind velocity and pressure. Results of their power spectral density and autocorrelation analysis are discussed in the next section.

IV. ANALYTICAL STUDIES OF WAVE PATTERNS

The results of analytical studies of wave perturbations of a few days records are given in the Table I. Values of the wave periods noted from sodar echograms, autocorrelograms and power spectral density plots are listed in the above table for comparision. Also the results of the meteorological conditions at the nearest sodar site and fluctuating components of surface wind velocities and surface pressure at the nearest point are included in the same table. Perturbations of pressure (P) and wind velocity ( ) attain large values during the time of wave perturbations on the above days.

For analysis purposes we have taken the sodar structure for 14th Feb., 7th May, 28th May and 19th Nov. of 1986. The plot of power spectral density, autocorrelation, surface wind and pressure fluctuations for the wavemotions of the above dates are shown in fig.(3)-(5). The wave perturbations on 7th May 1986 are noticed around 0500 IST and the height range of the propagating wave lies between 100-250 m. The nature of the wave is a complex one and spectral analysis reveals multiple frequencies. The spectral plot of the above dates are shown in fig.(3). The dominant frequencies as calculated from the spectral analysis are 600 μHz and 1200 μHz. The periodicity of the order of 54 minutes. A thunderstorm was recorded on the midnight and prior to the above oscillations. The morning radiosonde data at India Meteorological Department, Dum Dum which is about 1.5 km from sodar site indicates a sharp increase in dry bulb temperature and a fall of humidity of about 5%. The surface records of pressure and wind velocity as shown in fig.(5) indicate an enhancement of pressure by 2 mb and wind velocity by 5 kmph.

Fig.3. Spectral plot of wave perturbations on different dates.

The examination of the echogram in fig.2(c) shows that on 18-19 Nov. 1986 data, prior to wave perturbations which occurred from 0430 to 0530 IST there was no ground based inversion. The peak to peak amplitude of wave perburbations corresponds to about 50 m only and are superposed on a very shallow inversion layer. Wave motions are almost sinusoidal in nature and the peaks are vary distinct. The spectral analysis as shown in fig.(3) also indicates a single frequency of 3.5 mHz and the periodicity as calculated from the autocorrelation indicates a period of 4.2 minutes. The corresponding radiosonde data for morning flight (0515 IST) shows a slight increase of atmospheric temperature and slight fall of humidity at the higher altitude. The surface records of pressure and wind velocity indicate an increase in pressure of about 1 mb and sharp fluctuations of wind velocity from 4 to 10 kmph.

Fig.4. Auto correlation plot of sodar structures on different dates.

A similar analysis has been made for the sodar structure on 14th Feb. 1986 and 28th May 1986. The dominant frequencies as calculated are 3.1 mHz and 720 μHz respectively and the periodicity are 5 minutes and 5.4 minutes. On 28th May a severe thunderstorm was recorded at night and the wind speed attained is 60 kmph. According to Beer (1974) uniform sinusoidal oscillations lying within the narrow bands of periods (5-10 munites) should be essentially tropospheric gravity waves, irregular fluctuations with periodicities 15-60 minutes may be due to convection and influence of the meteorological

disturbances. Also these oscillations in the early morning hours may be followed a reversal of land and sea breeze in this region. The sea breeze starts in the early morning and persists only for 3 to 4 hours over this region sometimes noticed during premonsoon.

Fig.5. Plot of surface wind and pressure fluctuations.

The analysis also shows that during the period of wave undulations there is a considerable increase of ground pressure and wind velocity. This suggests peopagating waves in the atmosphere. Hooke et.al (1973) concentrated their attention on one gravity wave went recorded by both microbarographs and acoustic sounders and concluded that the generation of this gravity wave could be attributed to a numerically unstable wind shear layer in the planatary boundary layer.

REFERENCES

1. Beer, Tom, (1974) : Atmospheric waves, Adam. Hilger, London.
2. Gage, K.S. (1979) : J. Atmos. Sci.,V.36,1950-1954.
3. Hines, C.O. : Canadian J. Phys., V.38, 1441-1479.
4. Hooke, W.H., Hall F.F.(Jr.) and Gossard, E.E. (1973) : Boundary Layer Meteorol,V.5,No.2,200-210.
5. Hooke, W.H. and Hardy, K.R. (1975) : J. Appl. Meteorol, V.14, 31-88.
6. Uccellini, L.W. (1975) : Mon. Western Rev., V.103, 497-513.
7. Vanzandt, T.E., (1982) : Geophys. Res. Letter, V.9, 575-578.

Authors are greatful to DST for funding the project and render their thanks to Mrs. S De Bhowmick, Mr.S Chakravorty and Mr.J Gupta for secretarial assistance.

TABLE 1 : Summary of observed and computed data for the four days wavemotions. (April 1986 to March 1987)

Type of data	Date	14 February	07 May	28 May	19 November
	Time (IST)	0245-0345	0500-0630	2330-2430	0430-0500
SODAR DATA	Weather condition	Clear sky,mild wind condition	Thunderstrom with rain on 6th night then cloudy.	Thunderstorm around 2300 IST, mild wind speed.	Clear sky with strong wind.
	Surface wind velocity	04 kmph	11 kmph	15-20 kmph	10 kmph
	Velocity fluctuation(Δv)	almost nil	Sharp,5-11 kmph	Sharp, 6-15 kmph	Sharp, 4-10 kmph
RADIOSONDE & SURFACE DATA	Pressure fluctuation (ΔP)	+0.8 mb	+2.5 mb	+2.5 mb	+1.0 mb
	Surface temperature change	-1.0°C	+1.0°C	-1.5°C	+0.5°C
	Surface humidity change	+6%	-5%	-3%	-5%
	Dominant frequency	3.1 mHz	0.45 & 1.2 mHz	0.36 & 0.68 mHz	4.1 mHz
COMPUTED VALUES	Periodicity	5 minute	About 54 minute	About 28.6 minute	4.2 minute
	Nature of wave	Periodic	Complex	Complex	Periodic

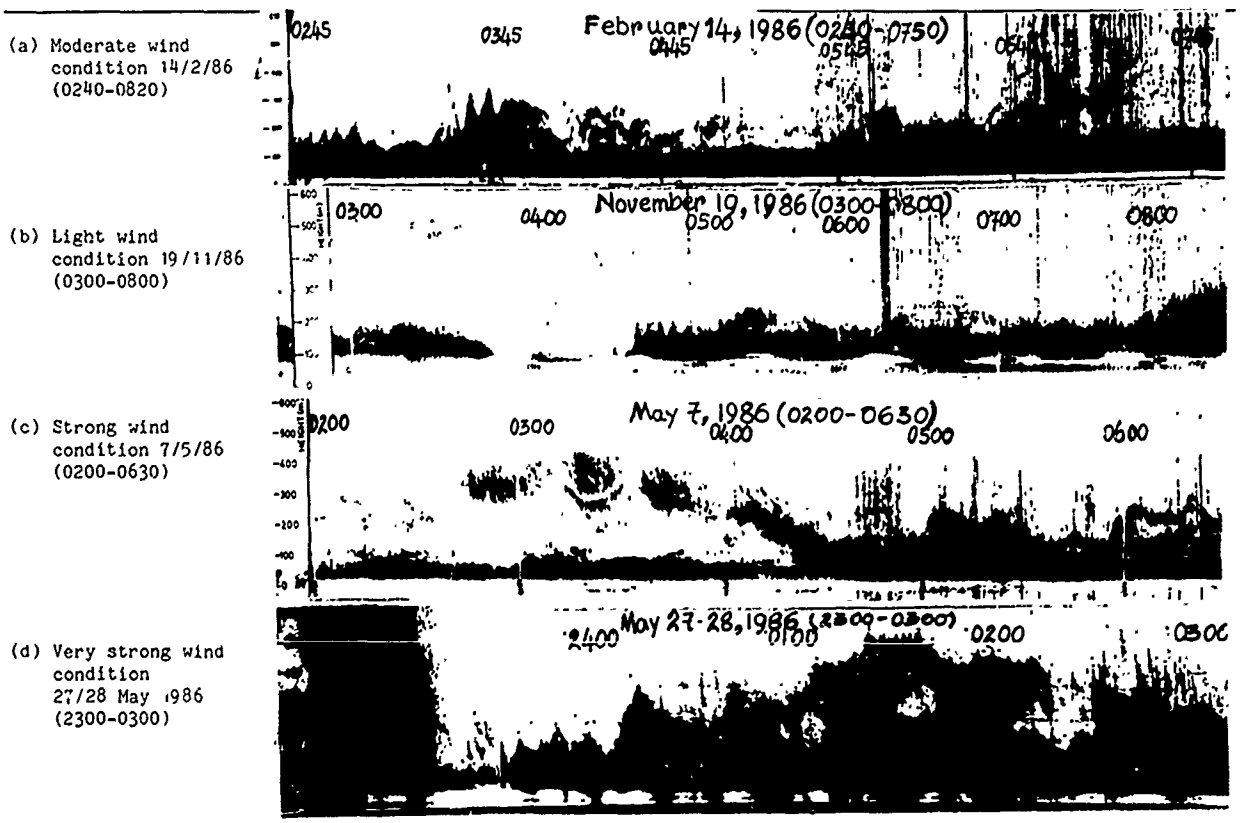


Fig. 2 Sodar structure showing wave motions in a stable condition.

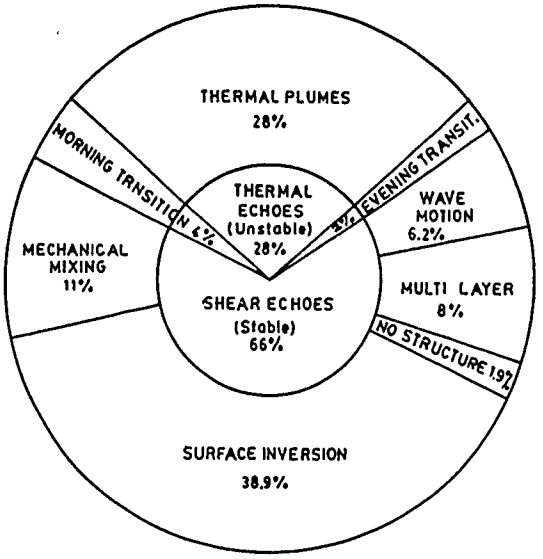


Fig.1 Percentage occurrance of various atmospheric structure (April 1986 - March 1987)

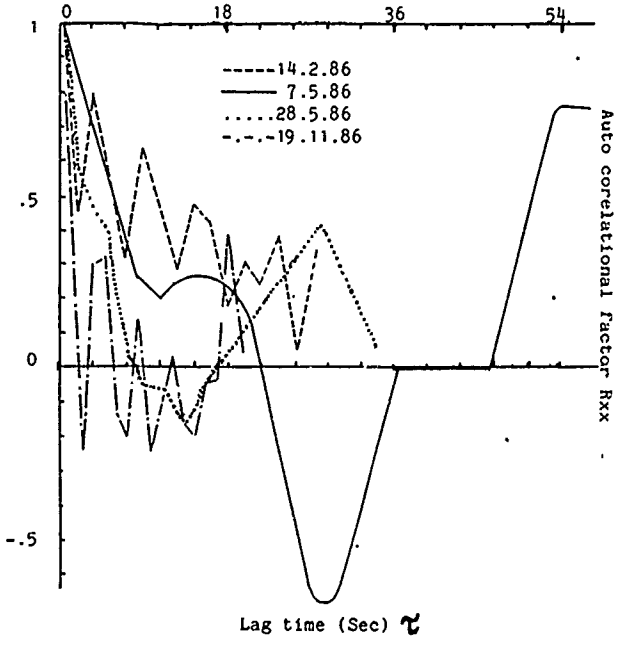


Fig.4 Auto correlation plot of sodar structure on different dates.

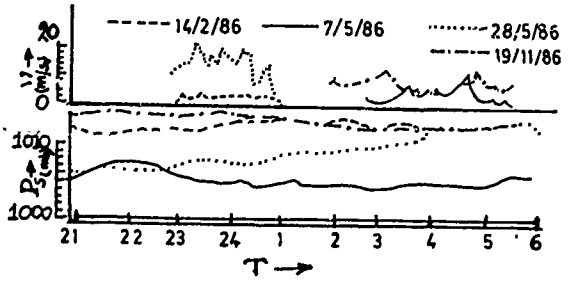
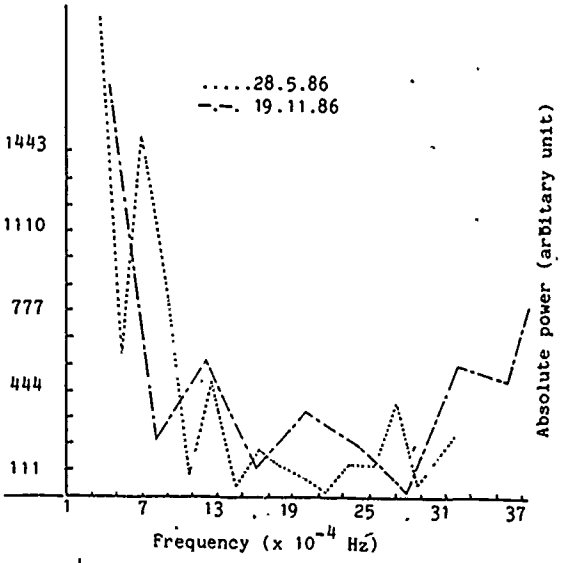


Fig.5 Plot of surface wind velcciti (v) and pressure (P_s) fluctuation.



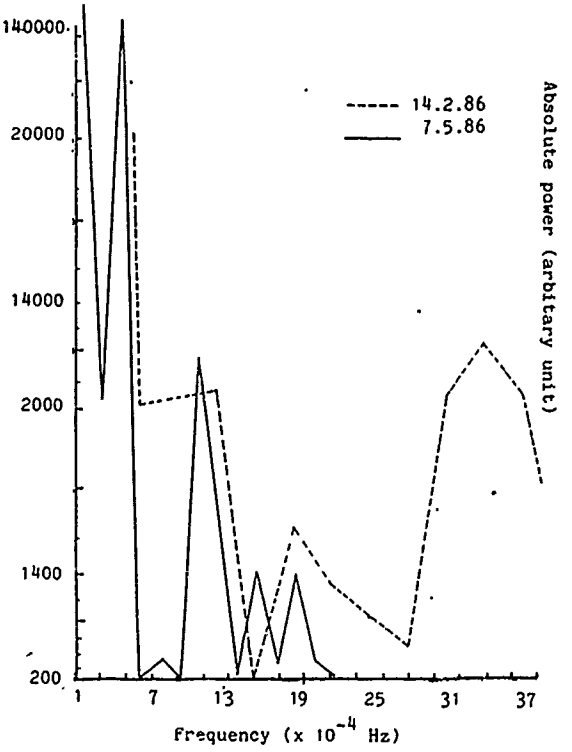


Fig.3 Spectral plot of wave perturbution on different dates.

Dissection of Multimodal Directional Ocean Wave Spectra from LEWEX

Thomas W. Gerling and Robert C. Beal
The Johns Hopkins University
Applied Physics Laboratory
Laurel, Maryland 20707, USA

Abstract

The Labrador Extreme Wave Experiment (LEWEX) was a comprehensive international experiment whose purpose was to compare directional ocean wave spectra obtained from a variety of sources. Over a seven day period, from 13 through 19 March 1987, directional spectra were estimated simultaneously and coincidently from as many as 12 shipboard and aircraft sensors. Subsequently, using a common hindcast wind field, nine wind wave models (two first generation, four second generation, and three third generation) have all yielded additional estimates of the directional spectrum. At some epochs, more than 40 simultaneous estimates now exist at adjacent wave model grid points. Spectral estimates from a number of LEWEX participants have now been merged into a single common data base, and the preliminary results show a number of interesting patterns which may further stimulate debate concerning not only the improved accuracy of recent third generation wave models, but also the potential role of spaceborne synthetic aperture radar for updating and improving wave model predictions. Three sets of meteorological conditions which occurred during LEWEX have proven to be especially valuable for revealing subtle differences among both the wave models and the wave sensors (both in situ and remote): 1) a rapidly evolving wave field associated with a well developed low pressure system that passed through the region on the 13th and the 14th, 2) two directly opposing wave systems with nearly identical dominant wavenumbers that were incident upon the site on the 15th and 16th, and 3) a remotely-generated swell system that encountered a regenerating wind field enroute to the site on the 17th and 18th.

Introduction

LEWEX was an international effort to assess methods of measuring and modelling the directional aspects of wind generated ocean waves, especially their evolution in the presence of rapidly turning winds (Bales, et al., 1987). The main data gathering period of LEWEX occurred from 13 through 19 March 1987, and was supported by a large number of ship and aircraft based estimates of the directional wave spectrum. In addition, directional spectra were estimated by a total of six first, second, and third generation wave models (using their own separately determined wind fields) and then hindcasted by these same six models plus three others (using a common hindcast wind field).

The LEWEX site was located in the southern Labrador Sea (Figure 1), centered on two adjacent "first generation" model grid points (50N, 47.5W, and 50N, 45W). The Canadian research vessel QUEST was located at 47.5W from 15 through 19 March. The Dutch research vessel TYDEMAN was located at 45W from 14 through 18 March. The TYDEMAN joined the QUEST on 19 March for sensor intercomparisons in a common sea. Both vessels obtained estimates of the directional wave spectrum using drifting buoys, moored buoys, and standard navigation radars.

Independent estimates of directional spectra were also obtained with three aircraft radar sensors: the NASA surface contour radar (SCR), the NASA radar ocean wave spectrometer (ROWS), and the CCRS C-band synthetic aperture radar (SAR). Aircraft estimates of the spectrum were obtained by at least one of the three sensors on six of seven days between 13 and 19 March. All three sensors obtained estimates over both grid points on 14 and 17 March, and two of the three obtained estimates over both grid points on 16 and 18 March. On 13 and 19 March, only the SAR obtained estimates, and only over the western grid point.

Several numerical wind wave models were providing routine predictions of directional wave spectra in the North Atlantic during March 1987: the "first generation" U.S. Navy Global Spectral Ocean Wave Model (GSOWM), the "second generation" British Meteorological Office (BMO) model, and the "third generation" European Centre for Medium Range Weather Forecasts (ECMWF) WAM model. In addition, a tailored forecast for the LEWEX region was provided by Oceanweather, Inc., of Cos Cob, Conn., using the ODGP (Ocean Data Gathering Program) first generation wave model. Subsequent to the experiment, the HYPA model in Spain has been driven by the BMO wind fields, and the French Met Office has driven their VAG model with the ECMWF wind fields. Thus, at least six wave models have been driven by four independent estimates of the wind field.

Also subsequent to LEWEX, a new "common" wind field was created by Oceanweather, Inc., using a subjective blend of all the available ship reports and buoys. This common LEWEX wind field was used to drive each of the above six models and three additional ones. a second generation NOAA model, a third generation NASA/GSFC model, and a third generation BIO (Bedford Institute of Oceanography) model. The

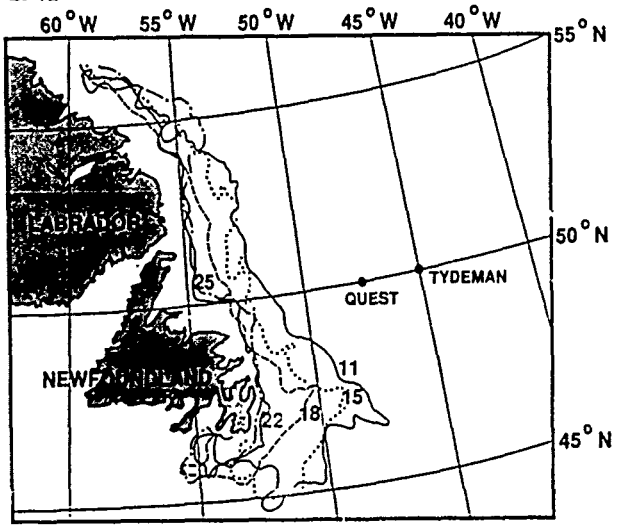


Figure 1: LEWEX ship locations and ice edge, March 1987.

BIO model was run with and without modelling of the ice in the region of LEWEX. The results included here incorporate this modelling; the spectra calculated without consideration of the ice edge were nearly identical to those calculated with the WAM model. The common wind field was not necessarily the "correct" wind field, since it was based on only a sparse set of surface measurements Nevertheless, it closely approximated the actual wind field dynamics that occurred during LEWEX, and it has been most useful for revealing subtle differences among the models. Nine wave models have been driven with the LEWEX common winds.

Although the data gathering period of LEWEX began on 13 March, the common wind field began at 1200Z on 9 March to allow sufficient startup time for each of the models. For safety reasons, the aircraft and ship sensors (except for two moored buoys, one each at the QUEST and the TYDEMAN) were constrained to measure only during daylight hours. Furthermore, operational wave forecasts are usually issued at six hour intervals centered on 00Z, 06Z, 12Z, and 18Z. The bulk of the LEWEX measurements were generally centered on 12Z (0830 Atlantic Standard Time) each day, with 18Z considered a backup choice. The backup was utilized on only one occasion, on 16 March, when inclement weather caused a delay of the NASA (SCR and ROWS) flight, and a cancellation of the CCRS (SAR) flight.

Virtually all of the wave activity during LEWEX was the result of three approximately sequential wave generating events, the first passing to the north around 12 March, the second passing directly through during 13 and 14 March, and the third spending much of its energy just to the south during 17 and 18 March. These three events produced a small number of fairly distinct wave systems which evolved systematically through the seven day LEWEX measuring period. A complete summary of the LEWEX data set is given in (Gerling, 1989).

A Typical Set of LEWEX Spectra

Figure 2 shows a typical three-hourly-interval sequence of LEWEX spectra from one of the nine models which were driven with a common wind field. Full directional wavenumber spectra are displayed in spectral energy density units of m^4, individually normalized to the range of 0 to 255 (the range of one byte integers), and contoured in eight equally spaced intervals. The peaks indicate the direction towards which waves propagate. All spectra are plotted in linear wavenumber out to $2\pi/100$ radians per meter (100 meter wavelength), with reference circles at 100 meters, 200 meters and 400 meters. Prior to contouring, an interpolation scheme has been employed which increases the resolution of the spectra so that smoother and more accurate contours can be obtained. This method is designed to preserve the total energy of the high resolution spectra within the original coarse binning and therefore does not introduce distortion such as broadening of spectral peaks. When the spectra are effectively undersampled by the model binning, this undersampling is revealed by a correlation between the contours and the bin boundaries.

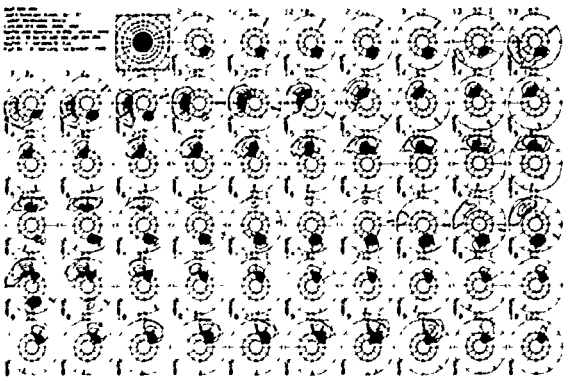


Figure 2. Three-hourly-interval sequence of ODGP spectra from 12Z 12 March through 12Z 19 March, 1987 at 50N, 45W.

Parameterization of Multimodal Spectra

The sequence of spectra shown in Figure 2 can be reduced to a more compact form by displaying only the significant peaks or "spectral modes" The vector plot shown in Figure 3 displays the location of spectral modes for the temporal sequence of Figure 2 on a graph whose horizontal coordinate is time and whose vertical coordinate is log frequency. The vertical scales are wave period on the right and wavelength on the left. The two scales are related through the dispersion relationship $k = \frac{(2\pi f)^2}{g}$ The location of the mode in polar coordinates can be obtained from this graph, with the vector direction indicating the angular component The vector length is proportional to a strength estimate w_i:

$$w_i = SWH \times \sqrt{P_i}$$

Each spectral mode has a corresponding value P_i and $\sum P_i = 1$ for any particular spectrum. The specific calculations are described below. These vectors may correspond to wave systems generated from distinct, remote sources. This will certainly be the case when the spectral masses corresponding to the modes are completely separate. When there is overlap, the interpretation is not as obvious, but the algorithm here determines the location of all modes, presuming that they might represent distinct wave systems. If two modes are very close, each location estimate may be distorted by the other mode's presence and the partitioning of the total wave height variance between the two systems will be difficult. When they become too close, only one mode may be discernible and this algorithm will only track that one, even if the spectrum is really the result of the superposition of several wave systems. An averaged location estimate for each mode is also calculated but is not used here. This particular estimate exhibits more variability than the mode and

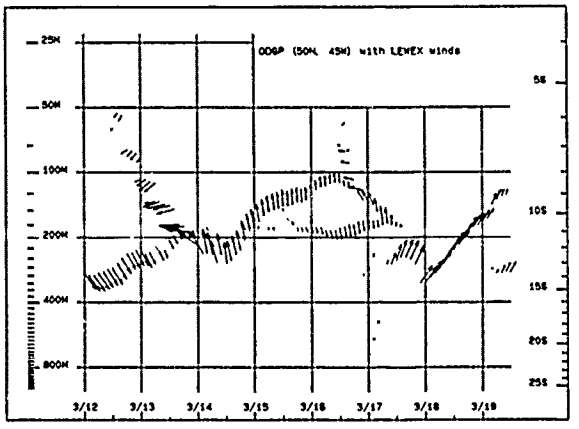


Figure 3: Vectors indicating the direction, period and strength associated with all spectral modes from the sequence of Figure 2.

generally results in a shorter wavelength estimate.

To examine whether a spectrum $S(f,\theta)$ is multimodal, first determine whether the support of the spectrum is connected The idea of connectedness of the support is really the most direct way to determine if an arbitrary spectrum is unimodal. The reason for this is that a unimodal spectrum may not have a unique maximum value, but instead might be topped with a "plateau", in which several contiguous locations all share the same maximal value. A nearest neighbor scheme to determine a unique maximum is complicated in this situation. The regions which determine connectedness are also used in the calculation of several of the quantities discussed below. If the support is connected, then first find if a level l exists such that the support of S thresholded at l, S^l, is not connected. If there is no such l, then the spectrum has only one mode, and its location will be calculated and assigned a strength of 1. However, if such an l exists, then find the lowest and call it l_0. Express the support of S^l as the union of a minimal number of connected regions R_{i_0}:

$$\bigcup_{i_0=1}^{n_0} R_{i_0}$$

For each of these regions, calculate a quantity M_{i_0} which will ultimately be used to estimate the strength of a component wave system recursively:

$$M_{i_0} = \int_{R_{i_0}} S(f,\theta)\, f\, df\, d\theta$$

Also calculate

$$P_{i_0} = \frac{M_{i_0}}{\sum M_{i_0}}$$

Now, for each of the regions R_{i_0}, perform the same procedure as just described: first determine if a level l exists such that the support of S restricted to R_{i_0} and thresholded at l, $S^l_{R_{i_0}}$, is not connected.

If there is no such l, then this portion of the spectrum has only one mode, and its location will be calculated and assigned a strength of P_{i_0}. However, if such an l exists, then find the lowest and call it l_{i_0}. Express the support of $S^{l_{i_0}}_{R_{i_0}}$ as the union of a minimal number of connected regions $R_{i_0 i_1}$:

$$\bigcup_{i_1=1}^{n_{i_0}} R_{i_0 i_1}$$

For each of these regions, calculate as before a quantity $M_{i_0 i_1}$:

$$M_{i_0 i_1} = \int_{R_{i_0 i_1}} S(f,\theta)\, f\, df\, d\theta$$

and also calculate

$$P_{i_0 i_1} = P_{i_0} \times \frac{M_{i_0 i_1}}{\sum_{i_1} M_{i_0 i_1}}$$

The algorithm continues now to consider each of the regions $R_{i_0 i_1}$ and assigns a strength of $P_{i_0 i_1}$ to a unique mode but if there is none, goes on to calculate $l_{i_0 i_1}$ which in turn determines

$$S^{l_{i_0 i_1}}_{R_{i_0 i_1}}$$

and its support

$$\bigcup_{i_2=1}^{n_{i_0 i_1}} R_{i_0 i_1 i_2}$$

and the corresponding strength estimates

$$P_{i_0 i_1 i_2} = P_{i_0 i_1} \times \frac{M_{i_0 i_1 i_2}}{\sum_{i_2} M_{i_0 i_1 i_2}}$$

and so on. These recursively calculated strength estimates correspond to a finer and finer fractionation of wave height variance and are computed analogously to conditional probabilities. The procedure results in a hierarchically structured description of the spectrum, a "tree structure", which lends itself relatively easily to "pruning", or removal of weak or spurious peaks from the description of noisy spectra. If a mode is judged to be insignificant, the corresponding "branch" is "pruned" and the statistics used are the ones calculated at the previous level in the tree. Finally, when the algorithm finishes, there will be a set of weights $P_{i_0 i_1 i_2 \cdots i_n}$ which sum to 1 and can be used to partition the wave height variance among the various wave systems represented in the spectrum.

At the beginning of the description of this algorithm, it was assumed that the support of $S(f,\theta)$ was connected. If it is not, the identical procedure is followed except the l_0 at the beginning is now 0.

As noted above, weighted estimates of location are also calculated, although they have not been displayed in this report. They are calculated using the regions R_I, where I represents an index in the hierarchical description:

$$f_I^{av} = \frac{\int_{R_I} S(f,\theta) f^2 df\, d\theta}{\int_{R_I} S(f,\theta) f\, df\, d\theta}$$

$$\theta_I^{av} = \arctan \frac{\int_{R_I} S(f,\theta)\sin(\theta) f\, df\, d\theta}{\int_{R_I} S(f,\theta)\cos(\theta) f\, df\, d\theta}$$

These estimates can show more variability than the modal estimates because the region R_I can change significantly when a new peak emerges.

Dissection of the ODGP Wave Spectra

The calculations described above have been performed for many of the LEWEX spectral sequences. The parameters obtained have then been clustered using a partially automated algorithm which separates slowly varying time sequences of parameters corresponding to distinct wave systems. For nearly all of the data sources examined, the LEWEX wave field can be explained as the joint evolution of five distinct wave systems. Several models which have been run using the LEWEX common winds embedded in their own wind field show more. These five systems are evident in Figure 3. The first is swell from the Labrador Sea present from 0Z 12 March through 14 March. Nearly all models agreed on this system. The sec-

ond is a rapidly evolving system which is consistently aligned with the local wind direction. This system represents a newly generated wind sea until 14 March, when the wind direction changes by about 60 degrees, and wind sea and swell merge. The next system represents northerlies present from about 15 March through 17 March. The fourth system is a freshly generated wind sea beginning during the latter half of 16 March and persisting through about 12Z 17 March. The final system represented is distant swell from the southwest which arrived at the Tydeman site on 17 March and is then the only component of the spectrum for several days. Peak period and direction for each wave system, as well as an estimate of the associated strength, are encoded in the vectors in Figure 3. Figure 4 depicts the total SWH for the ODGP model at 45W throughout the LEWEX time period. Figure 5 shows the partitioned SWH curves for the five LEWEX wave systems. The sum of the squares of the five curves of Figure 5 is equal to the square of the curve in Figure 4.

Natural comparisons between data sources can be made in terms of each of these wave systems. These comparisons are more revealing than those using averaged quantities such as total SWH, and average period and direction for a number of reasons. First, the five LEWEX wave systems are tied to distinct meteorological events. Differences between the buoy and model wave systems can indicate some information regarding the development and time evolution of these meteorological systems which were only sparsely observed. Since these generating regions are distinct, measures which average the entire directional spectrum mix this information and are not so useful in this purpose. Moreover, all the data sources agree on some systems but differ on others. Again it is useful to have a method of reducing spectra which will not mix all this information.

References

Bales, S.L., Beal, R.C., and Freeman, N.G., The Labrador Extreme Waves Experiment (LEWEX) Science Plan: March 9-27, issued 9 March, 1987.

Gerling, T.W., Common Processing and Display of LEWEX Directional Ocean Wave Spectral Estimates from 12 March through 19 March 1987, JHU/APL Technical Report S1R-89U 911, April, 1989.

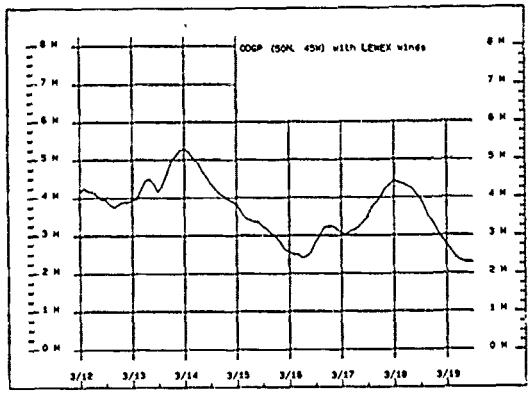


Figure 4: Significant wave height for ODGP.

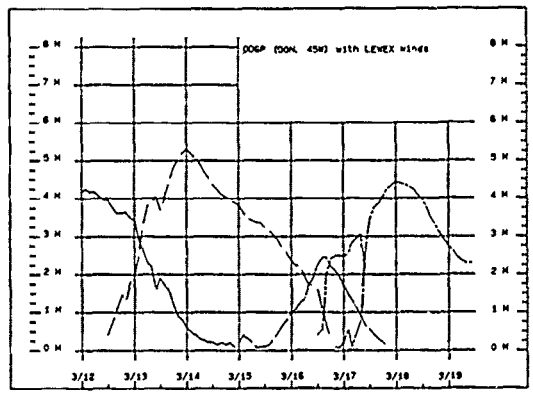


Figure 5: Partitioned significant wave heights for the five LEWEX wave systems for ODGP.

COMPARISON OF SIMULATED AND MEASURED SAR IMAGE SPECTRA DURING THE LABRADOR EXTREME WAVES EXPERIMENT

Claus Brüning and Liana F. Zambresky
European Centre for Medium Range Weather Forecasts
Shinfield Park, Reading, Berks. RG2 9AX, UK
Tel: +44-734-8760000, Fax: +44-734-869450, Tlx: 847908

During the Labrador Extreme Waves Experiment (LEWEX) in the North Atlantic in March 1987, a C-band SAR was flown on a Canadian CV-580 aircraft. SAR imagery of the ocean wave field was acquired at different flight altitudes. The measured SAR image spectra exhibit a strong dependence on flight altitude, indicating that the SAR imaging mechanism strongly depends on the R/V ratio as predicted by velocity bunching theory (R denotes the range of the target and V the platform velocity). We have simulated 2-dimensional SAR image spectra from hindcast wave spectra by using velocity bunching theory. The wave spectra were calculated from the wind field by applying a third generation wave forecast model, the so called WAM model.

Examples are presented showing that at a low flight altitude a unimodal wave spectrum is imaged into a unimodal SAR image spectrum, whereas at a high flight altitude the same wave spectrum is imaged into a bimodal image spectrum. Other examples show that a bimodal wave spectrum is imaged into a bimodal SAR image spectrum when the aircraft flies at a low altitude and into a unimodal image spectrum when it flies at a high altitude. These peculiarities of the SAR imaging are well predicted by our proposed SAR imaging theory of ocean waves.

We have compared the measured SAR image spectra with the simulated ones and it is shown that both spectra usually agree quite well. We consider this as a further experimental confirmation of the proposed SAR imaging theory of ocean waves. It also shows that under certain conditions it is not straightforward to invert measured SAR image spectra into ocean waves spectra. The inversion requires a detailed knowledge of the SAR imaging mechanism.

The Effect of Sampling Variability on the Estimation of Wave Number and Propagation Direction from SAR Wave Spectra

Frank Monaldo

The Johns Hopkins University Applied Physics Laboratory
Johns Hopkins Road, Laurel, Maryland 20707-6099
U.S.A.

Abstract

Even if a synthetic aperture radar (SAR) image of the ocean surface were a perfectly linear, noiseless representation of ocean surface wave height, a spectrum computed from the imagery would be but a single realization of the ensemble-mean ocean spectrum. In this paper we address two questions. (1) What statistics adequately describe the spectra of SAR ocean wave images? (2) Given these statistics, with what precision can ocean peak wave number and propagation direction be determined? It is shown that for typical SIR-B (Shuttle Imaging Radar-B) wave imagery and proper smoothing, precisions of about 2×10^{-3} rads/m and 3° in wave number and angle, respectively can be achieved.

Keywords: SAR, wave spectra, sampling variability

1 Introduction

A computed Fourier spectrum of an ocean surface wave field is a single realization of the ensemble-mean, wave height-variance spectrum. For a homogeneous wave field, the average of individual spectra approaches this ensemble-mean spectrum as the number of spectra included in the average increases. However, the fact that the mean spectrum, calculated from a finite number of spectra, will be randomly distributed about the ensemble-mean is referred to as "sampling variability". As a consequence of sampling variability, the statistical precision with which spectral parameters, for example dominant wave number and propagation direction, can be extracted is limited.

For ocean wave data obtained by more conventional instrumentation, the statistical properties of the corresponding spectra are well understood. Assuming that the phases of the Fourier coefficients computed from a Fourier transform are uniformly distributed between 0 and 2π, it can be demonstrated that the spectral energy density at any particular place in the wave number plane is randomly distributed about the ensemble mean value with a χ^2 distribution having $2N$ degrees of freedom (DOF). Here N is the number of individual spectra averaged to produce the mean spectrum [Donelan and Pierson, 1983].

The spectra computed from SAR imagery could potentially involve additional complications. First, there is considerable controversy as to the extent that the SAR image intensity can be linearly related to either ocean wave height or slope [Monaldo and Lyzenga, 1988; Alpers et al., 1986]. This paper does not attempt to address this question. We are addressing the inherent precision of SAR image spectra here, not how accurately they represent ocean wave spectra.

Second, SAR images are contaminated by multiplicative speckle noise. If $I(x, y)$ represents the SAR image that would be obtained if no speckle noise were present as a function of the spatial coordinates x and y, then $\tilde{\nu}_n \cdot I(x, y)$ is a corresponding n-look, speckled SAR image. The term $\tilde{\nu}_n$ represents a random variable having χ^2-distribution with n-degrees of freedom. To avoid confusion, please note that a χ^2-distribution can be an appropriate distribution for either a spectrum or the speckle statistics of a SAR image. It is possible, however, that multiplicative speckle-noise in a SAR image may alter the spectral statistics in such a way that the χ^2-distribution is no longer applicable to SAR image spectra.

In 1982, Goldfinger analytically demonstrated that if the spectrum of a speckled image is white, then the noise in this spectrum could be described by a χ^2-distribution. However, it is not clear whether the spectra of ocean wave imagery are sufficiently broad that Goldfinger's result is applicable.

In this paper, we will use actual SIR-B SAR ocean image spectra to verify that indeed a χ^2-distribution adequately describes the sampling variability of such spectra. Using this distribution and a computer simulation, we will estimate the limitations on the precision with which wave number and propagation direction can be estimated.

2 Statistics of SAR Image Spectra

We tested the hypothesis that the χ^2-distribution is applicable to image spectra by examining 25 spectra from SIR-B datatake 91.5 acquired off the southern coast of Chile. The spectra were obtained by computing the two-dimensional Fourier transforms of 25 contiguous 6.4 km by 6.4 km, 4-look SAR wave images. The individual images were sampled every 25 m by 25 m, so that each image was composed of 256 by 256 pixels. Each of the values in the complex transforms, computed from the images, was multiplied by its complex conjugate to produce the corresponding raw SAR image intensity-variance spectrum.

We designate each one of these spectra as $S_i(k_r, k_a)$, where i varies from 1 through 25, k_r represents range (perpendicular to the satellite ground track) wave number, and k_a represents azimuth (parallel to the ground track) wave number. The mean

spectrum is computed by

$$S_{mean}(k_r, k_a) = \frac{1}{25}\sum_{i=1}^{25} S_i(k_r, k_a). \quad (1)$$

This mean spectrum is an estimate of the ensemble-mean spectrum.

At each wave number bin, i.e., specific pair of values k_r and k_a in each of the raw spectra, the ratio of the associated spectral value to the corresponding value in the mean spectrum was computed and is designated $\tilde{z}$. For example, let k_{r0} and k_{a0} represent a specific pair of wave number values, let the value of the mean spectrum, $S_{mean}(k_{r0}, k_{a0})$, equal 100, and let the corresponding value in the third spectrum $S_3(k_{r0}, k_{a0})$, be equal to 134, then $\tilde{z}$ equals 1.34. The distribution of these $\tilde{z}$-values were computed over all wave number bins and all individual spectra.

If the χ^2-distribution is applicable to SAR image spectra, then the distribution of $\tilde{z}$'s ought to be χ^2 distributed with two DOF. This special case of the χ^2 distribution is also known as an exponential distribution. Figure 1 presents a histogram of the computed $\tilde{z}$'s and a theoretical exponential distribution. Note that the two agree very well. Although this agreement does not constitute formal proof, we do claim this circumstantial evidence is sufficiently strong so that we may proceed on the assumption that the χ^2-distribution describes the sampling varaibility of image spectra calculated from speckled, 4-look SAR images.

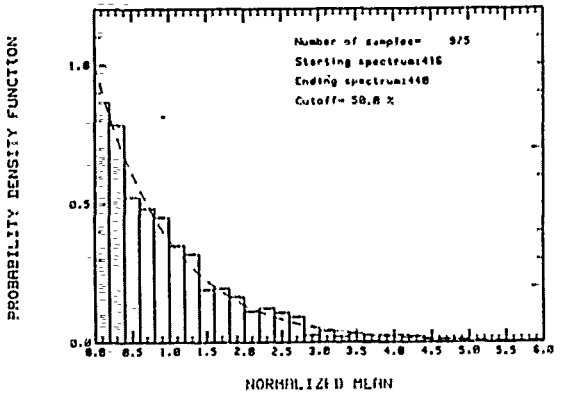


Figure 1 The dashed curve is a theoretical exponential distribution. The $\tilde{z}$ values are calculated by dividing individual spectral values with the corresponfing spectral value in the mean of 25 spectra. The histogram of these values is shown.

The DOF associated with the average of N spectra is equal to $2N$. However, we have in the past [Beal et al., 1986] reduced sampling variability by smoothing within individual spectra rather than averaging separate spectra. This smoothing allowed us to examine the spatial variability of spectra in a way that would not be possible if we simply computed averages from individual spectra covering large areas.

The smoothing is performed by convolving a Gaussian-shaped kernel, $\kappa(k_r, k_a)$, with an individual raw spectrum. The kernel is given by

$$\kappa(k_r, k_a) = \frac{1}{\sqrt{2\pi}k_\sigma} \exp\left[\frac{-(k_r^2 + k_a^2)}{2k_\sigma^2}\right] \quad (2)$$

where $k_\sigma = m\Delta k$ and Δk is the wave number bin size. For our spectra Δk equals 9.86×10^{-4} rads/m. Typically, in the past, we have used an m equal to 3.5. An individual smoothed spectra is calulated from the following equation:

$$S_{i,smooth}(k_r, k_a) = S_i(k_r, k_a) * \kappa(k_r, k_a) \quad (3)$$

where the convolution operation is designated as $*$.

This spectral smoothing raises two questions. First, it is not immediately obvious how to compute the effective DOF of a spectrum obtained from internal smoothing as opposed to averaging separate spectra. This DOF value permits the comparison of the sampling variability of SAR image spectra with spectra from other instruments, which frequently have their sampling variability specified in terms of DOF. Second, a way must be found to estimate the precision with which wave number and propagation direction can be extracted given a specified level of smoothing.

The first question was addressed by applying various smoothing kernels to the 25 raw SIR-B spectra discussed earlier. With no smoothing applied the $\tilde{z}$-values appeared to be exponentially (χ^2 with two DOF) distributed. In general, for any χ^2 distribution the DOF $= 2\mu^2/\sigma^2$, where μ and σ^2 are the mean and variance of the distribution.

Since the distribution of $\tilde{z}$ values calculated from 25 smoothed spectra is an estimate of the true distribution, we used the mean and variance calculated from the $\tilde{z}$ values to estimate the effective DOF. The relationship between spectral smoothing and DOF was determined by smoothing each of the 25 SIR-B spectra with $\kappa(k_r, k_a)$ functions having m-values varying between 0 and 3.5, computing the mean and variance of the resulting $\tilde{z}$ distribution, and inferring the effective DOF. Figure 2 provides a plot of DOF versus m. Note that for $m = 0$, i.e., no smoothing, DOF = 2 and for $m = 3.5$ (the smoothing which was generally applied) DOF = 100.

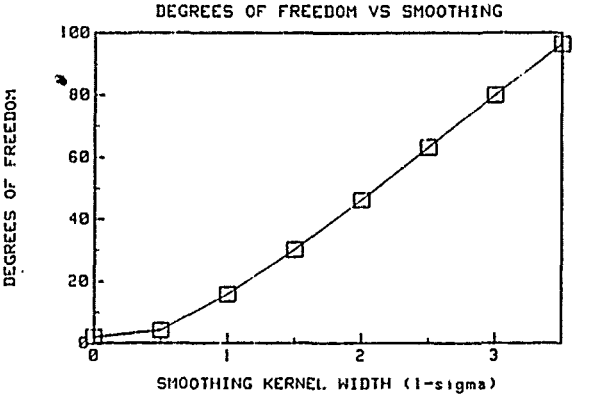


Figure 2 The DOF of SAR image spectra versus the size of the smoothing kernel.

3 Wave Number and Angle Precision

Although we can estimate the DOF associated with smoothed spectra, we can not directly infer the precision with which wave number and propagation direction can be determined from this value. This precision is dependent not only on the DOF but on the shape of the spectral peaks as well. For example, the position in the wave number plane of a extremely narrow peak is well determined regardless of the amount of sampling variability. The position of a broad peak, on the other hand, might be substantially affected by small levels of sampling variability.

Estimates of wave number and propagation angle precision were obtained via a computer simulation. Figure 3 is an example of a singled smoothed SIR-B image spectrum from data take 91.5. The center represents zero wavenumber. The outer

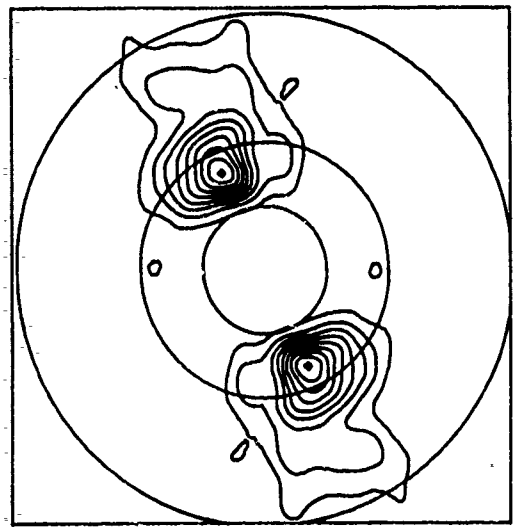


Figure 3 A smoothed SIR-B image spectrum calculated from imagery acquired off the coast of Chile, data take 91.5. Contours represent linear levels of image intensity-variance. The outer circle represents a wavelength of 100 m. The range direction is vertical.

circle represents $k = \sqrt{k_r^2 + k_a^2} = 2\pi/(100\text{ m})$. The horizontal axis lies along the k_a-direction, the vertical axis lies along the k_r-direction. The linearly spaced contours represent levels of spectral intensity-variance. We choose here to concentrate our attention on the spectral peak at about 200 m wavelength.

We asserted this smoothed spectrum to be our ensemble-mean spectrum. We then generated 25 simulated raw spectra by multiplying the ensemble-mean spectrum at every wave number bin with a exponentially distributed random variable having mean and variance equal to one. Various levels of smoothing were then applied to these 25 simulated spectra producing eight sets of 25 spectra, with m varying from 0 to 4. For each of these sets the ensemble-mean spectrum and its dominant wave number and direction are known *a priori*.

The dominant wave number and propagation direction for each of the spectra were computed using two different methods. The maximum-value method simply assigned the dominant wave number and propagation direction by finding the location in the wave number-plane of the maximum spectral value. In the second method, position of the "center-of-mass" of the spectral peak used as a measure of dominant wavelength and direction.

The variation from spectrum to spectrum, within every particular set of 25 spectra, of the extracted wave number and propagation direction represents the noise on the measurement of these quantities solely as a result of sampling variability for the particular level of smoothing applied.

Figure 4 represents the root-mean-square (rms) variation in wave number and propagation direction as a function of m (the level of smoothing) using the maximum-value method. The error bars represent the 95% confidence limits of the determination of the rms values based on a sample size of 25. The line shown is a linear fit to the data, showing the gradual decrease in variability as a function of m. At $m = 3.5$, the level traditionally employed, the precision of is 2.3×10^{-3} rads/m in wavenumber and 2.8° in propagation direction.

Figure 5 similar plots obtained using the center-of-mass method. Note that at low levels of smoothing the precision associated with the center-of-mass method is considerably better than the maximum-value method. However, at higher levels of smoothing the differences in methods grow smaller. At $m = 3.5$, the center-of-mass method achieves precisions of 1.8×10^{-3} rads/m in wave number and 2.5° in propagation direction.

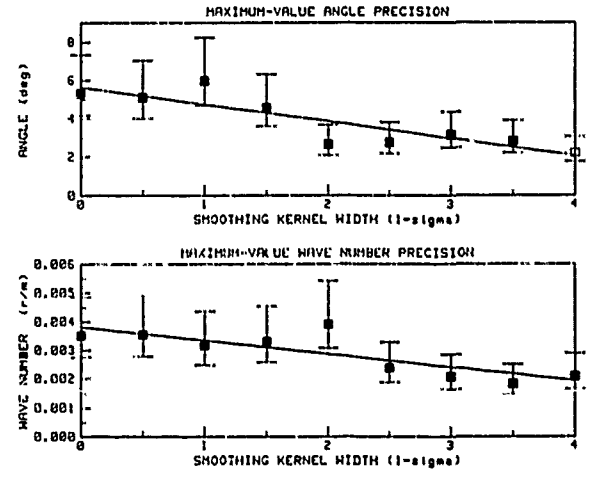


Figure 4 The rms variation in wave number and propagation angle associated with sampling variability as a function of smoothing kernel size. The error bars represent 95% confidence intervals on the estimates of this variability. Wave numbers are propagation angles are obtained with the maximum value method.

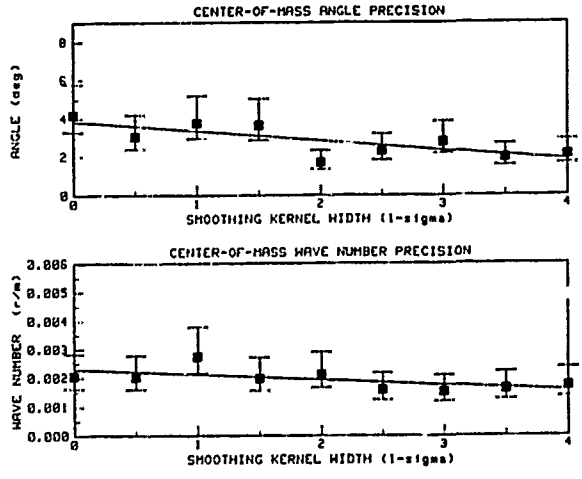


Figure 5 The rms variation in wave number and propagation angle associated with sampling variability as a fuction of smoothing kernel size. The error bars represent 95% confidence intervals on the estimates of this variability. Wave numbers and propagation angles are obtained with the center-of-mass method.

4 Conclusions

We have reached two major conclusions concerning the impact of sampling variability on SAR wave image spectra. (1) The traditional χ^2-distribution, which describes the sampling variability of spectra in general, also applys to the spectra of 4-look SAR images in spite of the fact that the imagery is contaminated with multiplicative, speckle noise. (2) For a typically shaped peaks found in SIR-B image spectra, the precisions of wave number and propagation direction associated with sampling variability are about 2×10^{-3} rads/m and 3°, respectively.

Acknowledgements. I would like to thank Professor Willard Pierson of the City College of New York and Mr. Robert Beal of The Johns Hopkins University Applied Physics Laboratory for posing enlightened questions to me concerning the nature of sampling variability as it applies to SAR image spectra.

5 References

1. Alpers, W.R., C. Brüning, and K. Richter, "Comparison of simulated and measured synthetic aperture radar imaging spectra with buoy derived ocean wave spectra during the Shuttle Imaging Radar-B mission", IEEE Trans. Geosci. and Remote Sensing, Vol. 88, pp. 559-566, 1986.

2. Beal, R.C., T.W. Gerling, D.E. Irvine, F.M. Monaldo, and D.G. Tilley, "Spatial variations of ocean wave directional spectra from Seasat synthetic aperture radar", J. Geophys. Res. Vol. 91, pp. 2433-2449, 1986.

3. Donelan, M. and W. Pierson, "The sampling variability of estimates of spectra of wind-generated gravity waves", J. Geophys. Res. Vol. 88, pp. 4381-4392, 1983.

4. Goldfinger, A., "Estimation of spectra from speckled images", IEEE Trans. Aerosp. Electron. Syst., Vol. AES-18, pp. 675-681, 1982.

5. Monaldo, F.M. and D.R. Lyzenga, "Comparison of Shuttle Imaging Radar-B ocean wave imagery with linear model predictions", J. Geophys. Res., Vol. 93, pp. 15,374-15,388, 1988.

Resolution of Ocean Wave Propagation Direction In Single-Pass Airborne SAR Imagery

Paris W. Vachon

Data Acquisition Division
Canada Centre for Remote Sensing
2464 Sheffield Road
Ottawa, Ontario K1A 0Y7

R. Keith Raney

Canada Centre for Remote Sensing
and RADARSAT
2464 Sheffield Road
Ottawa, Ontario K1A 0Y7

Abstract

There is an inherent 180° ambiguity in the derived wave propagation direction when using conventional spectral analysis techniques on standard SAR image products. In this paper, we successfully use three different techniques to resolve this ambiguity in propagation direction on the basis of a single pass of airborne SAR data. We take advantage of the fact that the SAR is characterized by a large time-bandwidth product, and therefore, that a sequence of individual looks extracted from the Doppler spectrum represents images of the scene as collected at a sequence of discretely delayed intervals of time. The techniques utilized include correlation-based motion analysis of a pair of looks, spectral weighting based upon a pair of looks and the ocean wave dispersion relation, and a full three-dimensional spectral analysis.

Key Words: SAR, Ocean, Wave, Spectrum

1 Introduction

A major problem in the extraction of ocean wave spectral information from SAR image data using conventional spectral analysis techniques, is that there is an inherent 180° ambiguity in the derived wave propagation direction. In the sections below, we consider three techniques which allow resolution of this ambiguity in propagation direction on the basis of single pass airborne SAR data. We take advantage of the fact that any practical SAR is characterized by a large time-bandwidth product. Therefore, a sequence of individual looks extracted from the Doppler spectrum represents images of the scene as collected at a sequence of discretely delayed intervals of time (Raney and Vachon, 1988). Thus, a seven-look image, for example, represents a time sequence of seven ocean wave images.

The temporal separation between looks i and j is given by

$$\tau = \frac{R \sin \gamma_{ij}}{V} , \qquad (1)$$

where R is the nominal scene range, V is the platform velocity, and γ_{ij} is the angular separation between the two looks. For typical airborne SAR parameters and processing, we achieve a time sequence of seven spatial images, each separated by times on the order of one second.

The first technique we consider is correlation-based motion analysis, previously used successfully to establish pack ice motion in AVHRR imagery (Ninnis *et al.*, 1986). The second and third techniques are borrowed from Marine Radar ocean wave analysis. The second technique uses dispersion relation-based phase weighting of undetected spectra derived from a pair of looks to resolve the propagation direction (Atanassov *et al.*, 1985). The third technique involves a full three-dimensional spectral analysis of the individual look imagery time series (Young *et al.*, 1985).

These approaches to resolving the wave propagation direction are demonstrated using C-band SAR imagery of ocean waves penetrating the marginal ice zone. The data was collected by the Canada Centre for Remote Sensing (CCRS) CV-580 SAR (Livingstone *et al.*, 1987) during the Labrador Ice Margin Experiment (LIMEX) '87 (McNutt *et al.*, 1988), and processed to seven individual looks on the CCRS C-Sharp ground-based SAR processing system. The look-summed image of interest is shown in Fig. 1. The relevant SAR/processor parameters for our scene are listed in Table 1.

The "waves in ice" case is used for these demonstrations because the ice cover acts like a natural low-pass filter, effectively removing the high frequency ocean wave components. This eliminates the complication of a finite scene coherence time as is present in the general open ocean case (Raney, 1981).

In the sections below, we outline these three techniques and present some results.

2 Correlation Motion Analysis

Consider the ocean surface displacement in response to a single sinusoidal ocean wave component:

$$\eta(x, y, t) = A \cos(K_x x + K_y y - \Omega t + \phi) , \qquad (2)$$

where x is the azimuth direction, y is the ground range direction, A is the amplitude, ϕ is the relative phase angle,

$$K = \sqrt{K_x^2 + K_y^2} = \frac{2\pi}{\Lambda} \qquad (3)$$

is the magnitude of the wavenumber vector, Λ is the wavelength, and $\Omega = \sqrt{gK}$ is the wave frequency, which is assumed to obey the deep water dispersion relation. This is a simplification for the case of "waves in ice" for which the dispersion relation is somewhat more complicated (Liu and Mollo-Christensen, 1988). Superposition applies to the case of multiple wave components.

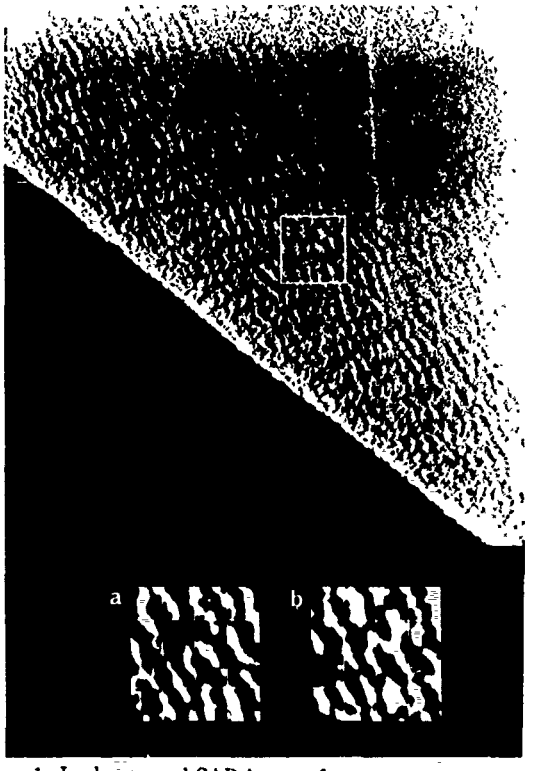


Figure 1: Look-summed SAR image of ocean waves penetrating floating ice as used in this study. Subscenes a) and b) were extracted from looks 3 and 5 at the indicated location. The individual look subscenes have been smoothed for clarity and overlaid with a fixed grid. The relative displacement of certain wave features is apparent.

Table 1: Relevant SAR/processor parameters for image of Fig. 1.

date		87-03-21
time		16:46 GMT
nominal location		N47.3° W52.5°
platform velocity	V	129 m/s
platform height	h	6065 m
platform track angle		160°
PRF/V		2.57 Hz-s/m
radar wavelength	λ	5.6 cm (C-band)
look direction		left
polarization		HH
sensitivity time control	STC	land
minimum incidence angle	θ	45°
number of looks	N	7
separation between looks	γ	0.521°
azimuthal pixel spacing	δx	6.0 m
ground range pixel spacing	δy	6.0 m

Assume that the spatially and temporally varying surface height of equation (2) leads to an envelope of reflectivity which propagates noncoherently and keeps pace with the actual wave crests. The details of the mechanism of generation of this pattern are irrelevant. However, suffice it to say that the translating patterns exist due to the combined effects of velocity bunching, tilt cross-section modulations, and hydrodynamic straining of the Bragg scale waves.

Consider now two successive look images of the wave system of equation (2) separated by τ seconds as calculated from the SAR/processor parameters and equation (1). It can be shown that over the course of this time interval, the relative vector displacement $\Delta\vec{x}$ of the two images will be given by

$$\Delta\vec{x} = \frac{\Omega}{K}\left(\frac{\vec{K}}{K}\right)\tau = \vec{C}_0\tau\,, \tag{4}$$

where $C_0 = \Omega/K$ is the phase velocity of the wave component of interest. The vector displacement may be explicitly measured by using cross-correlation-based motion analysis techniques.

The earlier image was broken up into a series of 64-by-64 pixel filter windows, here denoted $f(i,j)$, where i is the sample number along x and j is the sample number along y. The later image was broken up into a series of 84-by-84 pixel search windows, here denoted $s(i,j)$, which were centred on the corresponding filter window. The cross-covariance function is

$$r_{fs}(i,j) = \frac{E\left[f(i',j')\,s(i'+i,j'+j)\right] - m_f m_s}{\sqrt{\sigma_f^2\,\sigma_s^2}}\,, \tag{5}$$

where $E[\,]$ denotes the expected value operator, m_f and m_s are the means, and σ_f^2 and σ_s^2 are the variances of the filter window and the overlapping portion of the search windows respectively. The required cross-correlation function was calculated using FFT techniques. Only unbiased correlation values were retained.

The resulting cross-covariance maxima described the peak misregistration reasonably well. Fig. 2 contains a plot of displacement vectors as derived by the technique outlined above and using look numbers 4 and 7. Displacements with low correlation values were discarded. Assuming that the waves are propagating into the marginal ice zone, the propagation direction has been correctly resolved.

Fig. 3 contains a plot of the magnitude of the measured displacement vectors D as a function of slant-range R. The least-

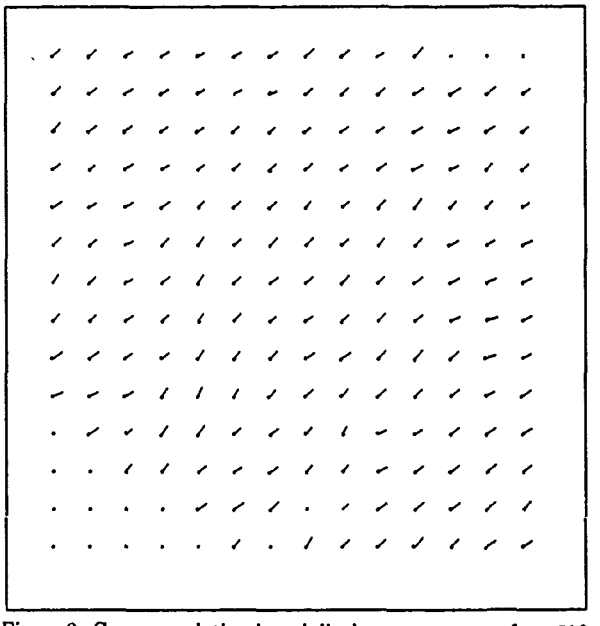


Figure 2: Cross-correlation-based displacement vectors for a 512-by-512 pixel subscene extracted from looks 4 and 7 of our waves in ice data set. The "dot" is at the base of the vector.

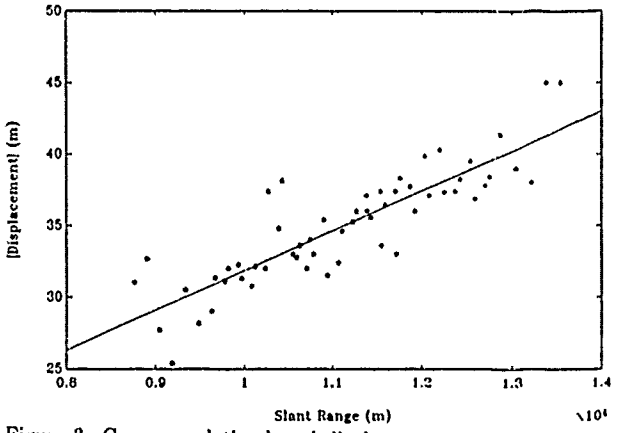


Figure 3: Cross-correlation-based displacement D versus slant-range R for looks 4 and 7.

squares fit line to this data is given by

$$D = 2.8(10)^{-3}\, R + 4.0\,. \tag{6}$$

For look numbers 4 and 7 and $\Lambda = 160$ m, the predicted slope from equation (1) is $C_0 \sin\gamma_{ij}/V = 3.3(10)^{-3}$ with 0 intercept. Our fitted line matches these predicted values rather well.

3 Dispersion Relation Weighting

Consider now the spatial Fourier transform of the wave of equation (2):

$$\begin{aligned}\tilde{\tilde{\eta}}(u,v,t) &= \iint A\,\cos(K_x x + K_y y - \Omega t + \phi) \\ &\quad \cdot \exp\{-j(ux+vy)\}\,dxdy\,,\end{aligned} \tag{7}$$

where the $\tilde{}$ represents the Fourier transform operation. The Fourier transform is

$$\begin{aligned}\tilde{\tilde{\eta}}(u,v,t) &= 2\pi^2 A\left[\exp\{-j(\Omega t-\phi)\}\delta^2(u-K_x)\,\delta^2(v-K_y)\right. \\ &\quad \left. + \exp\{j(\Omega t-\phi)\}\delta^2(u+K_x)\,\delta^2(v+K_y)\right]\,,\end{aligned} \tag{8}$$

and the power spectrum is proportional to the magnitude-squared of this result. The spectrum has a 180° ambiguity in propagation direction since

$$|\tilde{\tilde{\eta}}(u,v,t)|^2 = |\tilde{\tilde{\eta}}(-u,-v,t)|^2\,. \tag{9}$$

If we have spectral estimates from two successive times t_1 and $t_2 = t_1 + \tau$, it can be shown that the positive propagating component of the ambiguous spectrum $\Delta_+(u,v)$ may be separated from the negative propagating component $\Delta_-(u,v)$ via [Atanassov *et al.*, 1985]

$$\begin{aligned}\Delta_+(u,v) &= \{2[1+\cos(2\Omega\tau)]\}^{-1}\cdot\{|\tilde{\tilde{\eta}}(u,v,t_1)|^2 + |\tilde{\tilde{\eta}}(u,v,t_2)|^2 \\ &\quad - 2\,\mathrm{Re}[\tilde{\tilde{\eta}}(u,v,t_1)\,\tilde{\tilde{\eta}}^*(u,v,t_2)\exp(j\Omega\tau)]\}\,.\end{aligned} \tag{10}$$

An analogous expression exists for the ambiguous component of the power spectrum. In developing equation (10), it is assumed that the dispersion relation is known and is an even function of K.

The procedure of equation (10) was implemented in software, and applied to look numbers 3 and 5 for a 128-by-128 pixel region having a temporal separation of about $\tau = 1.5$ seconds. The deep water open ocean wave dispersion relation was assumed, and there was no effort to account for scanning distortion. The results are shown in Fig. 4. No preprocessing of any sort was applied to the two image subscenes. The technique has clearly succeeded in resolving the correct direction of wave propagation.

4 Three-Dimensional Analysis

Consider now the three-dimensional (two space, one time) Fourier transform of the wave of equation (2):

$$\begin{aligned}\tilde{\tilde{\tilde{\eta}}}(u,v,\omega) &= \iiint A\,\cos(K_x x + K_y y - \Omega t + \phi) \\ &\quad \cdot \exp\{-j(ux+vy+\omega t)\}\,dxdydt\,.\end{aligned} \tag{11}$$

The power spectrum is proportional to

$$\begin{aligned}|\tilde{\tilde{\tilde{\eta}}}(u,v,\omega)|^2 &= 16\pi^6 A^2\left[\delta^2(u-K_x)\,\delta^2(v-K_y)\,\delta^2(\omega+\Omega)\right. \\ &\quad \left. + \delta^2(u+K_x)\,\delta^2(v+K_y)\,\delta^2(\omega-\Omega)\right]\,.\end{aligned} \tag{12}$$

Note that the three-dimensional spectrum has an inherent asymmetry since

$$|\tilde{\tilde{\tilde{\eta}}}(u,v,\omega)|^2 = |\tilde{\tilde{\tilde{\eta}}}(-u,-v,-\omega)|^2\,. \tag{13}$$

We can use this asymmetry to advantage by integrating over negative frequencies to give

$$\int_{\omega<0} |\tilde{\tilde{\tilde{\eta}}}(u,v,\omega)|^2\,d\omega = 16\pi^6 A^2 \delta^2(u-K_x)\,\delta^2(v-K_y)\,, \tag{14}$$

in which the direction of propagation is completely resolved.

A three-dimensional spectral analysis has been applied to our "waves in ice" data set. The individual spectral bins, each separated by frequency interval

$$\Delta f = \frac{\Delta\omega}{2\pi} = \frac{1}{7\,\tau}\,, \tag{15}$$

are given in the first part of Fig. 5. Note that the frequency bins are rather coarse. The spectrum integrated over all frequencies, negative frequencies only, and positive frequencies only are given in the second part of the figure. Although much of the spectral energy lies in the 0 Hz bin of the spectrum, the wave propagation direction has been correctly resolved.

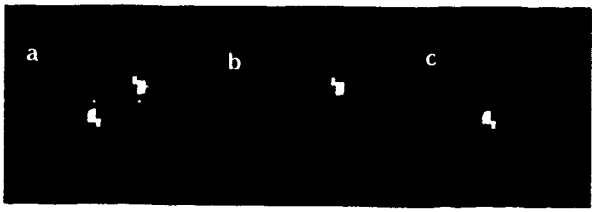


Figure 4: Spectral analysis results of dispersion relation based phase weighting of spectra derived from looks 3 and 5. a) The ambiguous spectrum $|\tilde{\tilde{\eta}}(\vec{K})|$, b) the correct spectral lobe $\Delta_+(\vec{K})$, and c) the ambiguous spectral lobe $\Delta_-(\vec{K})$. The cross-hairs are centred on zero wavenumber. The outer circle corresponds to 100 m wavelengths while the inner circle corresponds to 50 m wavelengths.

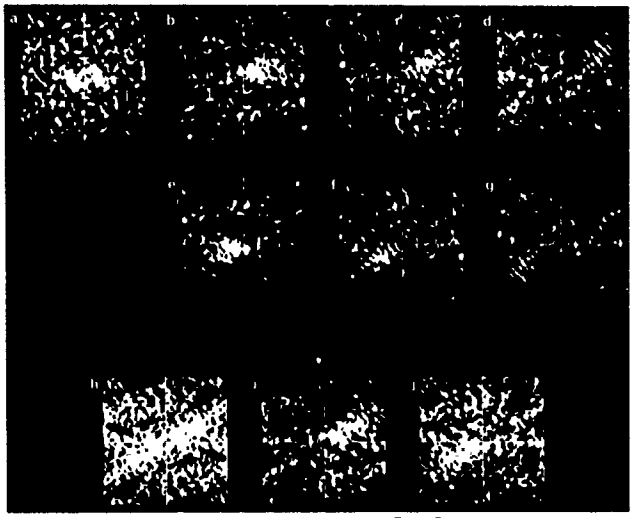


Figure 5. Three-dimensional spectral analysis on subscenes from the seven individual looks. The first seven frames represent the frequency bins centred on a) 0 Hz, b) −0.19 Hz, c) −0.38 Hz, d) −0.57 Hz, e) 0.19 Hz, f) 0.38 Hz, and g) 0.57 Hz. The final three frames are the results of integrating over h) all frequencies, i) negative frequencies only, and j) positive frequencies only.

5 Discussion

Each of the three techniques for resolving the ocean wave propagation direction in SAR data were employed with some success. The correlation analysis was least satisfying for several reasons. The output is a single displacement vector pointing in the nominal direction of average scene motion. Substantial optimization efforts were required to achieve reasonable performance. It is doubtful that this technique could be used at an operational level. However, using two pairs of looks with different temporal separations may make the technique more robust.

The dispersion-relation-based phase weighting approach seems to provide excellent performance for the case of "waves in ice". The disadvantage of this approach is the necessity to explicitly know the dispersion relation. This technique could be extended to provide enhanced spectral reliability by applying it to pairs of looks and averaging the results.

The most elegant and general approach is via the three-dimensional spectral analysis. Explicit knowledge of the wave dispersion relation is not required. However, general purpose SAR systems tend to have short integration times (of seconds), and therefore, inadequately sample the temporal motions of the waves within the scene.

A key aspect of these approaches is the useful extension to the case of a spaceborne SAR system. It is unlikely that spaceborne SARs designed to date will support the required level of temporal separation between looks, particularly considering their larger resolution cell size. However, a SAR could be specifically designed to image ocean waves and have as a feature the ability to provide such a look image time series (Raney and Vachon, 1989). The design of such a system is the subject of ongoing research.

6 Acknowledgements

The correlation-based motion analysis software was supplied by M. Collins (CRESS, York University) and was modified to suit the SAR case by F. Gildenhaar (CCRS). The suitability of the 3-D spectral analysis and the dispersion relation spectral weighting was suggested by W. Rosenthal (GKSS). P. McLean (CCRS) performed the SAR processing. The efforts of the LIMEX/LEWEX experimental teams lead to the successful collection of the SAR data.

References

1. Atanassov, V., W. Rosenthal, and F. Ziemer, "Removal of ambiguity of two-dimensional power spectra obtained by processing ship radar images of ocean waves," *J. Geophys. Res.*, Vol. 90, No. C1, pp 1061–1067, 1985.

2. Liu, A.K., and E. Mollo-Christensen, "Wave propagation in a solid ice pack," *J. Physical Oceanography*, Vol. 18, No. 11, pp 1702–1712, 1988.

3. Livingstone, C.E., A.L. Gray, R.K. Hawkins, R.B. Olsen, R.A. Deane, and J.G. Halbertsma, "CCRS C-band airborne radar—system description and test results," *Proc. 11th Canadian Symposium on Remote Sensing*, Waterloo, Ontario, June 22–25, 1987, pp 503–517.

4. McNutt, L., S. Argus, F. Carsey, B. Holt, J. Crawford, C. Tang, A.L. Gray, and C. Livingstone, "LIMEX '87: The Labrador Ice Margin Experiment, March 1987—A pilot experiment in anticipation of RADARSAT and ERS 1 data," *EOS*, Vol. 69, No. 23, pp 634–635, 643, 1988.

5. Ninnis, R.M., W.J. Emery, and M.J. Collins, "Automated extraction of pack ice motion from Advanced Very High Resolution Radiometry imagery," *J. Geophys. Res.*, Vol. 91, No. C9, pp 10,725–10,734, 1986.

6. Raney, R.K., "Wave orbital velocity, fade, and SAR response to azimuth waves," *IEEE J. Oceanic Eng.*, Vol. OE-6, No. 4, pp 140–146, 1981.

7. Raney, R.K., and P.W. Vachon, "SAR imaging of ocean waves from an airborne platform. focus and tracking issues," *J. Geophys. Res.*, Vol. 93, No. C10, pp 12,475–12,486, 1988.

8. Raney, R.K., and P.W. Vachon, "Implications from LIMEX/LEWEX for spaceborne SAR sensing of ocean wave spectra", *Proc. IGARSS '89*, Vancouver, Canada (this issue), 1989.

9. Young, I.R., W. Rosenthal, and F. Ziemer, "A three-dimensional analysis of marine radar images for the determination of ocean wave directionality and surface currents," *J. Geophys. Res.*, Vol. 90, No. C1, pp 1049–1059, 1985.

On the Focussing Issue of Synthetic Aperture Radar Imaging of Ocean Surface Waves

Werner Alpers[1], Claus Brüning[2], Jens Schröter[3]

[1]Universität Bremen, Fachbereich Physik, 2800 Bremen 33, FRG
[2]European Centre for Medium Range Weather Forecasts, Reading, UK
[3]Alfred-Wegener-Institut für Polar- und Meeresforschung, 2850 Bremerhaven, FRG

Abstract

Experimental evidence exists that the image contrast of synthetic aperture radar (SAR) imagery of ocean waves can, under certain conditions, be enhanced by defocussing the SAR processor. Several theories have been advanced recently to explain the focus sensitivity of ocean wave SAR imagery. In all these theories it is assumed that the phase or group velocity of the long waves enters into the SAR imaging mechanism of single-look imagery.

In this paper we show that also velocity bunching theory yields a focus sensitivity. In this theory, we have no need for recourse to the "phase velocity hypothesis". The focus sensitivity of SAR imagery of ocean waves results from an interplay of two motion-induced effects, which are affected by changes in focus setting: velocity bunching and azimuthal resolution.

The focussing theory developed in this paper is applied to ocean wave SAR imagery acquired by an airborne L-band SAR during the TOWARD experiment which was conducted off the coast of California in 1984. Two-dimensional SAR simulations studies carried out with variable focus settings show a reasonably good agreement with observed values.

Key words: synthetic aperture radar, ocean wave imaging, focussing theory

1. Introduction

The velocity bunching theory, together with the two-scale wave model, is a well-established theory to describe the imaging of ocean surface waves by synthetic aperture radar (for a review, see Hasselmann et al., 1985). However, it has been questioned whether this theory is capable of explaining the focus sensitivity of airborne SAR imagery of ocean waves. Experimental evidence seems to exist that the image contrast in SAR imagery of ocean waves can be enhanced when defocussing the SAR processor by an amount which is of the order of the phase velocity of the dominant ocean wave. Since the phase velocity of the long ocean wave does not enter into the velocity bunching theory, this experimental fact was considered by some investigators (see, e.g., Kasilingam and Shemdin, 1988) as experimental evidence that velocity bunching theory is not the correct SAR imaging theory of ocean surface waves.

Clearly, in multi-look SAR imagery the velocity of the modulation pattern, which is equal to the group velocity of the long ocean waves, is detectable. There is a wave pattern shift between different incoherent looks. When dealing with multi-look SAR imagery the image contrast can be enhanced by compensating individual looks for wave movement before look summation. As already pointed out by Raney and Schuchman (1978), this SAR image contrast enhancement is based on noncoherent tracking of waves and applies to any incoherent imaging device. But most investigators do not mean this image contrast enhancement when referring to focussing (Raney and Vachon, 1988; Lyzenga, 1988; Kasilingam and Shemdin, 1988; Schult et al., 1988). The focussing issue concerns primarily the focussing of one-look SAR imagery of ocean waves.

In the present paper, we show that velocity bunching theory can also yield image contrast enhancement when defocussing the SAR processor. In this theory the increase of SAR image contrast is caused by the fact that a defocus of the SAR processor modifies velocity bunching as well as the azimuthal resolution. The parameter setting where maximum image contrast is encountered can occur at velocities which are of the order of the group velocity of the dominant ocean wave.

2. SAR Imaging Theory of Ocean Waves

In this paper we make the hypothesis that the SAR imaging of ocean waves can be described by the velocity bunching theory as described in the papers by Alpers and Rufenach (1979), Swift and Wilson (1979) and Alpers et al. (1981). This theory contains the following elements that contribute to the SAR image formation: (a) velocity bunching, (b) degradation in azimuthal resolution due to a finite scene coherence time and orbital accelerations, and (c) cross section modulation due to tilting and hydrodynamic interaction.

In the velocity bunching theory it is assumed that the facets remain "physically distinct" during the SAR integration time. The radar cross section of a facet may vary during the integration time due to tilting and hydrodynamic modulation of the Bragg waves, but this has only a small effect on the imaging mechanism, which can be neglected to first order (see Hasselmann et al., 1985). The propagation velocity of the radar cross section modulation pattern, which in the case of Bragg scattering is equal to the group velocity of the long waves, do not affect the Doppler history of the backscattered radar signal and thus have no place in this SAR imaging theory.

When the SAR processor is not perfectly focussed, then, for stationary scenes, the image contrast is always reduced in comparison to the perfect focus case. However, this does not necessarily hold when SAR is employed for imaging ocean surface waves. In order to prove this we calculate the expression for the ensemble-averaged SAR image intensity, $I(\underline{x})$, for the case of a defocussed SAR processor. Mathematically, this amounts to replacing the platform velocity V by V - ΔV in the matched filter function, where ΔV is a focus adjustment parameter. The expression for $I(\underline{x})$ reads in the case of a defocussed SAR processor

$$I(\underline{x}) = B \int_{-\infty}^{+\infty} \int_{-\infty}^{+\infty} dx_o dy_o \, \Delta(y - y_o) \, \frac{\sigma(\underline{x}_o)}{\rho'_{aN}(\underline{x}_o, \Delta V)} \cdot$$

$$\cdot \exp\left\{ - \frac{\pi^2}{\rho'^2_{aN}(\underline{x}_o, \Delta V)} \left[(x - x_o - \frac{R \, u_r(\underline{x}_o)}{V(1 - 2\Delta V/V)} \right]^2 \right\} \quad (1)$$

Here $\sigma(\underline{x}_o)$ denotes the average normalized radar cross section of the facets, $u_r(\underline{x}_o)$ the orbital velocity of the facets in range direction, $\Delta(y)$ the SAR impulse response function in cross-track direction, R the range of the facet, and B a constant, which depends on SAR parameters. In deriving (1) we have assumed that $\Delta V/V \ll 1$.

The degraded azimuthal resolution for N incoherent looks appearing in (1) is given by

$$\rho'_{aN} = (1 - 2\Delta V/V)^{-1} N \rho_a \cdot$$

$$\cdot \left[1 + \frac{\pi^2 T^4}{N^2 \lambda_i^2} \left[a_r(\underline{x}_o) - 2\frac{V}{R} \Delta V \right]^2 + \frac{T^2}{N^2 \tau_s^2} \right]^{1/2} \quad (2)$$

Here λ_i denotes radar wavelength, ρ_a the one-look azimuthal resolution, T the coherent SAR intergration time, N the number of incoherent looks, a_r the orbital acceleration of the facets in range direction and τ_s the scene coherence time. This time is mainly determined by the spread of the facet velocities within the SAR resolution cell.

From these equations it can be seen that for $\Delta V \neq 0$ not only the azimuthal resolution ρ'_{aN} is modified, but also velocity bunching. As can be seen from the exponent in the integral (1), a de-tuned SAR processor modifies the azimuthal shift δx to

$$\delta x = \frac{R}{V(1 - 2\Delta V/V)} \, u_r(\underline{x}_o) \quad (3)$$

If the factor 1 - 2ΔV/V is smaller than unity, then the azimuthal shift is increased. As a consequence, also velocity bunching can increase and thus enhance the SAR image contrast. On the other hand, a de-tuned processor leads to a degradation in azimuthal resolution and thus to a decrease in image contrast. In general, there exists a $\Delta V \neq 0$ where the image contrast is maximum. In this case the SAR image is optimally focussed.

However, the degradation in azimuthal resolution due to a defocussed processor is slightly more complex than discussed above. While the average azimuthal resolution is always degraded by a mis-matched processor, this does not necessarily hold for all facets on the ocean surface. If certain facets on the ocean surface have a given non-vanishing radial acceleration $a_r(\underline{x}_o) \neq 0$, then it can be achieved by changing the focus setting ΔV that the second term in the last bracket of (2) vanishes. In this case ΔV has the value

$$\Delta V = \frac{R}{2V} \, a_r(\underline{x}_o) \quad (4)$$

Locally, this leads to a finer azimuthal resolution. If the SAR image intensity is particulary large for these facets, then such focus setting with $\Delta V \neq 0$ may also contribute to an increase of the SAR image contrast. Note that ΔV as given by (4) can attain values that are of the order of the phase velocity of long ocean waves.

3. SAR Simulations

Equation (1) contains the focus adjustment parameter ΔV in a very intricate way. In order to determine the focus setting at which the SAR image contrast is largest, numerical techniques have to be applied. In the spectral domain we have to find the focus setting at which the peak of the SAR image spectrum attains its maximum value. The numerical analysis is carried out by using Monte Carlo simulation techniques as described in detail in Alpers (1983) and Alpers and Brüning (1986).

The parameters used in the numerical simulations are close to the ones encountered during the Tower Ocean Wave and Radar Dependence (TOWARD) experiment on Oct. 31st, 1984 (Shemdin, 1988). The SAR employed was an L-band (1.225 GHz) SAR flown on a CV-990 aircraft. The slant range resolution was 8 m and the azimuth resolution (4 looks) 13 m. A swell propagating parallel and antiparallel to the flight direction was imaged during flight legs 1 and 2, respectively. The SAR simulations were performed with a peak wavelength of 116 m, a significant wave height of 1.24 m, a radar-ocean wave modulation transfer function (RAR MTF) of $|M^{RAR}| = 4$ (modulus) and $\eta^{RAR} = -135^{\circ}$ (phase), a scene coherence time τ_s of 0,1 s and an incidence angle of 35°. Fig. 1 shows the results of the Monte Carlo simulations for the flight altitude of 11780 m where the aircraft velocity was 273 m/s for leg 1 and 218 m/s for leg 2. Fig. 2 shows the results for the flight altitude of 6240 m where the aircraft velocity was 229 m/s for both flight legs.

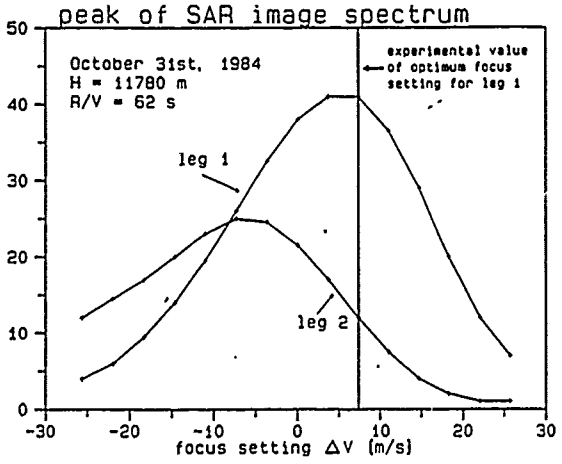


Fig. 1

Theoretical focussing curves for flight leg 1 (altitude: 11780 m, direction parallel to the wave direction) and flight leg 2 (altitude: 11720 m, flight direction antiparallel to the wave direction) on Oct. 31st, 1984. The curves show the peak value of the variance spectrum of the SAR image intensity in arbitrary units as a function of the focus adjustment parameter ΔV. The line indicates where optimum focus was found experimentally by De Witt et al. (1988).

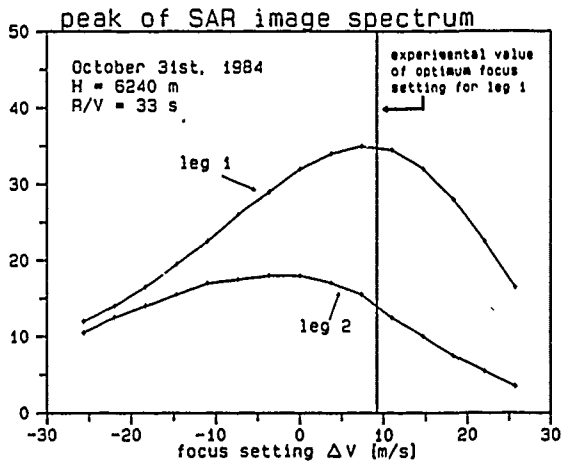


Fig 2

Same as Fig. 1, but for flight altitude 6240 m.

In both figures the experimental values of optimum focus settings found by De Witt et al. (1988) for flight leg 1 are inserted (vertical lines). The theoretical and experimental values of optimum focus settings agree reasonably well. Note that also the experimental values depend on flight altitude.

4. Discussion

From the above analysis it is evident that velocity bunching theory can account for focus sensitivity of airborne SAR imagery of ocean waves. The focus setting where optimum focussing is encountered is not related to the azimuthal component of the group velocity of the dominant ocean wave. In the case considered here the group velocity was 6.8 m/s.

Focussing depends on the wave propagation direction relative to the platform velocity, the range-to-velocity (R/V) ratio, the platform velocity V, and the wave spectrum. The sensitivity is greater for swell than for wind waves. Furthermore, focussing depends on the radar-ocean wave modulation transfer function (RAR MTF) as well as on scene coherence time. Quantitative results on focus sensitivity can be obtained from theory only if good estimates of these parameters are available.

Acknowledgement:

This research was supported by the Bundesministerium für Forschung und Technologie, Bonn (grant BMFT 01 QS 86174), and by the Office of Naval Research, Washington (grant N0014-83-G-0126, monitored by Hans Dolezalek).

References

Alpers, W., "Monte-Carlo simulations for studying the relationship between ocean wave and synthetic aperture radar image spectra", J. Geophys. Res., 88, pp1745-1759, 1983.

Alpers, W.R., and C. Brüning, "On the relative importance of motion-related contributions to the SAR imaging mechanism of ocean surface waves", IEEE Trans. Geosci. Remote Sens., GE-24, pp873-885, 1986.

Alpers, W.R., and C.L. Rufenach, "The effect of orbital motions on synthetic aperture radar imagery of ocean waves", IEEE Trans. Antennas Propag., AP-27, pp685-690. 1979.

Alpers, W.R., D.B. Ross, and C.L. Rufenach, "On the detectability of ocean surface waves by real and synthetic aperture radar", J. Geophys. Res., 86, pp5481-6498, 1981.

De Witt, R.J., F.S. Henyey, and J.A. Wright, "Comparison of computed SAR focus-setting curves with model predictions", La Jolla Institute, Center for Studies of Nonlinear Dynamics, La Jolla, Cal., 92039, report LJI-88-P-484, 32 pp., 1988.

Hasselmann, K., R.K. Raney, W.J. Plant, W.R. Alpers, R.A. Schuchman, D.R. Lyzenga, C.L. Rufenach, and M.J. Tucker, "Theory of synthetic aperture radar ocean imaging. A MARSEN view", J. Geophys. Res., 90, pp4659-4686, 1985.

Kasilingam, D.P., and O.H. Shemdin, "Theory for synthetic aperture radar imaging of the ocean surface: with application to the tower ocean wave and radar dependence experiment on focus, resolution, and wave height spectra", J. Geophys. Res., 93, pp13837-13848, 1988.

Lyzenga, D.R., "An analysis representation of synthetic aperture radar image spectrum of ocean waves", J. Geophys. Res., 93, pp13859-13865, 1988.

Raney, R.K., and R.A. Shuchman, "SAR mechanisms for imaging ocean waves", Proceedings of the 5th Canadian Symposium on Remote Sensing. Can. Remote Sens. Soc., Victoria, B.C., 1978.

Raney, R.K., and P.W. Vachon, "Synthetic aperture radar imaging of ocean waves from an airborne platform: focus and tracking issues", J. Geophys. Res., 93, pp12475-12486, 1988.

Schult, R.L., F.S. Henyey, and J.A. Wright, "Imaging of ocean waves by SAR", La Jolla Institute, Center for Studies of Nonlinear Dynamics, La Jolla, Cal., 92039, report LJI-87-P-467, 39 pp., 1988.

Shemdin, O.H., "Tower Ocean Wave and Radar Dependence Experiment: an overview", J. Geophys. Res., 93, pp13829-13836, 1988.

Swift, C.F., and L.R. Wilson, "Synthetic aperture radar imaging of moving ocean waves", IEEE Trans. Antennas Propag., AP-27, pp725-729, 1979.

THE USE OF AMBIGUITY FUNCTIONS FOR SAR MEASUREMENT OF OCEAN WAVES

Andrew S. Milman
Albert O. Scheffler

Radar Science Laboratory, Advanced Concepts Division
Environmental Research Institute of Michigan
PO Box 8618, Ann Arbor, Mi 48107

Abstract

This paper discusses the use of the ambiguity function of the complex SAR image for measuring ocean waves. This can reveal features that are not present in the real SAR image. The analysis leads to expressions relating this ambiguity function to the spectrum of the ocean surface height in a manner that can avoid many of the approximations usually needed to model SAR imaging of ocean waves.

1. Introduction

We describe a new method of analyzing the process by which synthetic aperture radars (SAR's) image ocean waves. We consider the complex image, rather than the real image, of the surface, and calculate its *ambiguity function* (defined below). The ambiguity function of the complex SAR image is interesting because it contains information about waves with periods shorter than the coherent integration time of the SAR and it extends the usual methods of focusing ocean waves to include range-traveling, as well as azimuth-traveling, waves.

For a scene that is large enough so that all the waves, both long and short, can be considered as random, the ambiguity function can be related by a simple integral expression to the ocean surface spectrum. We discuss this ambiguity function in the balance of this paper. Many of the usual approximations that are needed to describe SAR imaging of ocean waves are not needed here; this simplifies the theory considerably. Some aspects of this treatment of SAR imaging of ocean waves are:

(1) No separation of the surface height into large- and small-scale waves is needed.

(2) The surface is divided, in general, into random and non-random components, not large- and small-wavelength. Here, we treat only the case where all waves are random.

(3) No assumption is made about the coherence length of the surface reflectivity (e.g., see Hasselmann *et al.*, 1985); it is calculated from the surface spectrum.

(4) We also can calculate the coherence time, which plays an important role in the theory.

(5) The complex image, rather than the real image, is used: there is important information in the phase of the complex image.

(6) The use of the ambiguity function allows imaging any plane in the space formed by K_x, K_y, and Ω, which are the x- and y-components of the water wavenumber and the wave frequency, respectively.

2. Ambiguity Function

The *ambiguity function* of the complex image (J. Bennett and D. Lyzenga, 1988, private communication) is defined as follows. Let $I'(x_\iota,y_\iota)$ be the complex image for the image point (x_ι,y_ι), where x and y are the range and azimuth coordinates, respectively, and the subscript ι denotes coordinates in the complex image. Then define $I(x_\iota,y_\iota) = \exp(\pi i x_\iota^2) I'(x_\iota,y_\iota)$. The ambiguity function is

$$\chi_I(K_x,K_y;\xi_\iota,\eta_\iota) = \int_{-\infty}^{\infty}\int_{-\infty}^{\infty} I(x_\iota-\xi_\iota/2,y_\iota-\eta_\iota/2)\, I^*(x_\iota+\xi_\iota/2,y_\iota+\eta_\iota/2)\, e^{-2\pi i(K_x x_\iota+K_y y_\iota)}\, dx_\iota dy_\iota \qquad (1)$$

where * denotes complex conjugation; ξ_ι and η_ι are offsets in the image. This is equal to the Fourier transform of the real image in the special case that $\xi_\iota = \eta_\iota = 0$. While the complex image may be uncorrelated for $\eta_\iota \neq 0$, we shall include this term explicitly.

The main result of this theory can be expressed as follows. We start by writing the complex reflectivity as

$$r(\mathbf{x},t) = e^{-2ik_{em}R(\mathbf{x},t)}\, g(\mathbf{x},t) \qquad (2)$$

where $R(\mathbf{x},t)$ is the distance from the SAR to a point on the surface $z(\mathbf{x},t)$, and k_{em} is the electromagnetic wavenumber. The coefficient $g(\mathbf{x},t)$ is a complex number whose phase is constant, and therefore irrelevant; $|g(\mathbf{x},t)|^2$, the magnitude of the reflection coefficient, is given by Wright (1968) and others as the tilted-Bragg model. The phase of the complex reflectivity, aside from an unimportant constant, is determined by the instantaneous distance from the SAR to each point on the ocean surface. The ambiguity function of the reflectivity is defined by

$$\chi_r(K_x,K_y,\Omega;\xi,\eta,\tau) = \int_{-\infty}^{\infty}\int_{-\infty}^{\infty} r(x-\xi/2,y-\eta/2,t-\tau/2)$$

$$r^*(x+\xi/2,y+\eta/2,t+\tau/2)\, e^{-2\pi i(K_x x+K_y y+\Omega t)}\, dxdydt \qquad (3)$$

Using equation (2), this can be expressed as

$$\chi_r(K_x,K_y,\Omega;\xi,\eta,\tau) = \int_{-\infty}^{\infty}\int_{-\infty}^{\infty}\int_{-\infty}^{\infty} a^2(x-V_o t)g_o^2$$

$$e^{-\pi^2\Lambda(x,y,t;\xi,\eta,\tau)}\, e^{-2\pi i(K_x x+K_y y+\Omega t)}\, dxdydt \qquad (4)$$

where $a^2(x)$ is the antenna power pattern in the azimuth direction; g_o^2 is the mean radar cross-section; and V_o is the speed of the SAR. The function Λ is given by

$$\Lambda(\mathbf{x},t;\underline{\xi},\tau) = \int_{-\infty}^{\infty}\int_{-\infty}^{\infty}\int_{-\infty}^{\infty} S(\mathbf{x},t;\mathbf{k},\omega)$$

$$\Big[\, 4\sin^2[\pi(\mathbf{k}\cdot\underline{\xi}-\omega\tau)]\left[\cos^2\theta + \sin^2\theta\,\cos^2\psi\right] +$$

$$4\pi i\,\cos\theta\,(\underline{\gamma}\cdot\mathbf{k})\sin[\pi(\mathbf{k}\cdot\underline{\xi}-\omega\tau)]\Big]\, d^2\mathbf{k}\, d\omega \qquad (5)$$

where $S(\mathbf{x},t;\mathbf{k},\omega)$ is the surface height spectrum, which might depend on the location $\mathbf{x}$ and time t; $\mathbf{k}$ and ω are the water wavenumber and frequency; $\underline{\gamma} = (\gamma_x,\gamma_y)$ is the derivative of $g(\mathbf{x},t)$ with a change in surface slope; and $\underline{\xi} = (\xi,\eta)$ and τ are the spatial and temporal lags. To derive this result, we did not need to assume that the surface is spatially white, i.e., that $\Lambda(\mathbf{x},t;\underline{\xi},\tau)$ is proportional to $\delta(\xi)\delta(\eta)$. We also did not need to assume that the surface velocities are small; we have included the effects of acceleration and higher-order terms.

For an image that contains only random waves,

$$\chi_I(K_x,K_y;\xi_t,\eta_t) = \int_{-\infty}^{\infty} \chi_r(K_x-\xi,K_y,\xi-\xi_t;\xi,\eta_t,K_x)d\xi \qquad (6)$$

in the limit of a long integration time. The frequency of the water waves being imaged, Ω, is related to the lag ξ_t by

$$\Omega = \frac{2V_o}{\lambda_{em}R_o}\,\xi_t \qquad (7)$$

where λ_{em} is the electromagnetic wavelength and R_o the mean distance to the ocean surface. This is similar to the result of Lyzenga (1988), although he did not allow for a spatial lag ξ_t.

3. Relation to Surface Spectrum

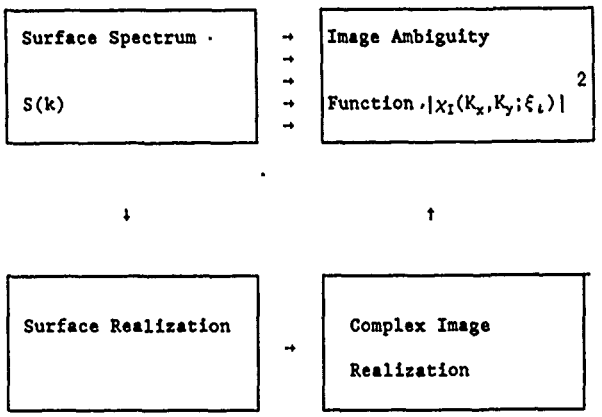


Figure 1. Surface spectrum and image ambiguity function.

The surface height spectrum $S(\mathbf{k})$ and the ambiguity function of the complex image $|\chi_I(K_x,K_y;\xi_t)|^2$ are related conceptually as we show in Figure 1. One way to calculate $|\chi_I|^2$ would be to start with S, generate some random realizations of the surface, generate a complex SAR image for each one, and finally calculate χ_I according to equation (1) and determine the squared magnitude. This is shown by the single arrows. The problem with doing this, aside from the amount of computer time that might be used, is that there are some approximations needed to calculate the complex image from a realization of the ocean surface; these involve somehow dividing the surface into large- and small-scale, or random and non-random, components. How this division is done will affect, sometimes strongly, the results of the calculations. In addition, it seems that some large-scale (or non-random) waves will affect the supposedly random component of the surface, making this division of scales inappropriate.

It is also possible to evaluate an ensemble average of $|\chi_I(K_x,K_y,\xi_t)|^2$. This would require calculating fourth-order statistics, however; it would be hard to interpret the results, so we shall not pursue this possibility.

Another method is to use equation (6) to calculate χ_I directly from S (multiple arrows). When this is done, waves on all scales are treated equally as being random; consequently, there is no scale separation. This will prove to be a great advantage because the computations are simplified. More importantly, there is a conceptual simplification because we do not have to argue about what are the most appropriate approximations to use; except for the assumed linear dependence of the radar cross-section on the surface slope, the relationship between χ_I and S can be expressed exactly.

It is also possible to calculate the change in χ_I when the spectral amplitude of the ocean surface height is changed by $\delta S(\vec{k}_o)$. This effectively linearizes equation (6), and allows us to determine the effect on the complex image of changing the surface height spectrum at some wavenumber k_o.

4. Results

There are some simple results of this SAR imaging theory. The most important may be that the image ambiguity function is proportional to

$$\exp\left\{-\left[\frac{R_o\lambda_{em}}{2V_o}\frac{K_x}{\tau_c}\right]^2\right\},$$

where τ_c is the coherence time of the reflectivity, in agreement with other authors (e.g., Rufenach and Alpers, 1981). This sets a lower limit to the wavelength of the azimuth-traveling waves that can be imaged.

A wave with a particular wavenumber k_o will produce harmonics in the complex image, as was found by Lyzenga (1988). This can confuse the interpretation of SAR images.

Let T be the coherent integration time; the spectral resolution is determined by a weighting function whose width is inversely proportional to T. Therefore, the spectral resolution is improved by using a long coherent integration time, which in turn requires an antenna with a large beamwidth.

Waves in range or azimuth with periods shorter than T, which would be smeared out in the real image, contribute to the ambiguity function of a single complex image; therefore we can estimate the amplitudes of these waves.

The correlation length and correlation time are both proportional to λ_{em}; SAR imaging of ocean waves would seem to be done best by use the longest electromagnetic wavelength possible.

Acknowledgement

This work was partially supported by the Office of Naval Research under contract N00014-81-C-0692.

References

Hasselmann, K., R. K. Raney, W. J. Plant, W. Alpers, R. A. Shuchman, D. R. Lyzenga, C. L. Rufenach, and M. J. Tucker, 'Theory of synthetic aperture radar ocean imaging: a MARSEN view,' J. Geophysical Res., 90(C3), 4659-4686 (1985).

Lyzenga, D. R., 'An analytic representation of the synthetic aperture radar image spectrum for ocean waves,' J. Geophysical Res., 93(C11), 13,859-13,865 (1988).

Rufenach, C. L. and W. Alpers, 'Imaging ocean waves by synthetic aperture radars with long integration times,' IEEE Tr. Antennas & Propagation, AP-29, 422-428 (1981).

Wright, J. W., 'A new model for sea clutter,' IEEE Trans. Antennas and Propagation, AP-16, 217-223 (1968).

SAR MEASUREMENTS OF OCEAN WAVES AND OTHER SURFACE PHENOMENA DURING THE GOASEX EXPERIMENT

David J. Napolitano and John F. Vesecky [1] Frank Gonzalez [2] Steve Peteherych [3]

[1] Center for Radar Astronomy
STAR Laboratory, Electrical Engineering Dept.
Stanford University, Stanford CA 94305-4055
Tel: 415-723-3475 or 2669 FAX: 415-723-0010

[2] NOAA
Pacific Marine Environment Laboratory
Seattle WA
Tel: 206-526-6803

[3] Atmospheric Environment Service of Canada
Downsview, Ontario M3H 5T4

Abstract

Synthetic aperture radar (SAR) is an important remote sensing tool for observation of the ocean over a large range of scales from 10's of meters to 1000's of km. Unique properties of SAR include self-illumination and cloud penetration because of the microwave frequencies employed. To evaluate SAR for observing the ocean surface, the SEASAT satellite was launched in 1978. To validate SAR ocean measurements, the Gulf of Alaska Experiment (GOASEX) was set up to provide measurements of the ocean surface and atmosphere within a region imaged by SAR. This data set is important because it contains a significant portion of the surface truth data collected during the SEASAT mission. With this data set, direct comparison between SAR and buoy ocean measurements can be made since the surface truth data was obtained *in situ* at the SEASAT overpass time. In this paper we compare SAR and surface observations of ocean waves. Further, we assess several models for SAR imaging of the ocean developed by Alpers, Harger, and Rotheram and Macklin. Lastly, we utilize the SAR for examining windrows on the ocean surface.

Keywords : SAR, Wave Spectra, Windrows

Introduction

The SAR data set consists of 11, 100 km by 100 km images digitally processed at the Jet Propulsion Laboratory (JPL) in 1987. During GOASEX, surface and atmospheric measurements from 9/13/78 through 9/26/78 were made by pitch roll and wave rider buoys, and CTD's, XBT's and radiosondes by the NOAA research vessel *Oceanographer* at the approximate locations 48° 42' N Lat., 133° 18' W Long. and 46° N Lat., 131° W Long. as shown in Fig. 1. In all cases, the surface measurements are well within the SAR image so a direct comparison can be made. Specifically, for ocean gravity waves we compare the SAR and surface measurements on the basis of dominant directional and omnidirectional wavelength, dominant direction, and multiple wave systems. In addition, we use the two dimensional FFT spectrum of the SAR to estimate waveheight, using the surface observations for calibration.

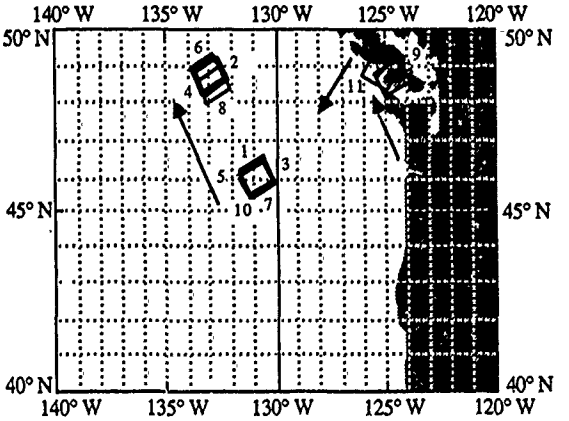
At the present date, there still exists controversy over wave imaging mechanisms, involving investigators such as Alpers, Harger, Plant, Lyzenga, Monaldo, Rotheram and Henye. To help resolve this controversy, we compare these theories on a analytical basis and use the GOASEX SAR and surface data since it provides many linear cases of wave imaging. We compare the unmodified, two dimensional FFT spectrum intensity of the SAR images to buoy spectra to see how well the SAR estimates of the waveheight spectrum fares without further processing.

We find that unmodified, the SAR spectrum provides good estimates of the waveheight spectrum in terms of estimating dominant directional and omnidirectional wavelength, direction, and omnidirectional bandwidth. In some cases, the unmodified SAR image spectrum corresponds closely with corresponding buoy measurements of the waveheight spectrum.

Lastly, the manifestation of periodic streaks approximately parallel to the wind direction are observed in the SAR image intensity spectrum. We will refer to these features as *'windrows'*. Two possible causes are either Langmuir circulations or atmospheric roll vortices. Gerling [2] has used SAR observations of windrows to estimate wind speed and direction. We extend this work, examining ocean surface windrows by comparing SAR observations with surface and subsurface *in situ* measurements of ocean variables.

Fig. 1

SEASAT SAR Image Footprints
Gulf of Alaska Experiment



No.	Rev.	No.	Rev.	No.	Rev.	No.	Rev.
1, 2	1126	5, 6	1212	9	1269	11	1306
3, 4	1169	7, 8	1255	10	1298		

Arrows indicate the SAR velocity vector of the images shown.

Processing

The GOASEX data set was processed in the following manner. The JPL digitally processed SAR images were 100 km x 100 km in extent and, depending on the particular image used, had the following approximate characteristics :

Center Incidence Angle :	23.1°
Antenna Beamwidth :	20.5°
Swath Velocity :	6.76 km/s
Center Resolution (Range x Azimuth) :	22.7 m x 23.6 m
Pixel Size :	12.5 m
Bits per Sample :	5
Surface Track to True North Angle :	28°
Revs. :	1126,1169,1212,1255,1269,1298,1306

In each of the 11 SAR images - 2 each of 1126 thru 1255 - the buoy coordinates were located. Next, a 25.6 km x 25.6 km section, or processing block, was centered at the buoy location in the SAR image. To generate a representation of the wave spectrum over this scale, the processing block was first subdivided into 4 - 12.8 km x 12.8 km sub-images. Then, each of these images were added together to produce a *composite* image. Then a 1024 x 1024 point, 2-dimensional FFT was performed on this composite image using Data General MV/10000, with the square magnitude, or intensity of the result being stored for later comparison with buoy spectra. The resulting spectrum has the property that the amplitude or intensity at a given frequency will have the same value as if the 2048 x 2048 point, 2-dimensional FFT of the entire 25.6 km x 25.6 km section had been performed. Although we have 1/2 the wavenumber resolution of a 2048 x 2048 point transform, we have reduced the computation and I/O time by approximately a factor of 4. Alternatively, one could transform each 1024 x 1024 point sub-image seperately, and add the complex-amplitude spectrum together - coherent addition. However, although it gives identical results, it incures ≈ 4 times the computation time due to the added 2-D FFT's. Presently, the computing time is 25 min. for the transformation process described. Therefore, the resulting wavenumber resolution is that of a single transformed 1024 x 1024 point image, and is given by

$$\Delta k = 2 * \frac{1}{25.6\ \text{km}} * .001\ \frac{\text{km}}{\text{m}} = 4.91 \times 10^{-4}\ \text{m}^{-1}$$

After computation of the raw spectrum intensity, convolution with a 9 x 9 (81-point), 2-D normalized gaussian *smoothing filter* with a 61% width of 3.48×10^{-3} m^{-1} was applied in the wavenumber domain to create a smoother spectrum (trading off resolution) to aide in locating the wavelength and direction of the dominant waves. The results were displayed using color-graphics software written by us for the Data General GDC-1000 Workstation.

Wave Imaging Theory

Comparison - There have been several theories proposed regarding the SAR ocean wave imaging process, including those of Alpers, Ross, and Rufenach [1], Harger [3], [4], and Rotheram [7], with each predicting similar effects in some cases, but different effects, in others, *ie.* velocity bunching. Upon analyzing each author's expressions for the *Sea Surface, Scattering Model, Received Signal, SAR Processing, and Transfer Function*, we found the expressions to be similar in form despite the different mathematical formalism used.

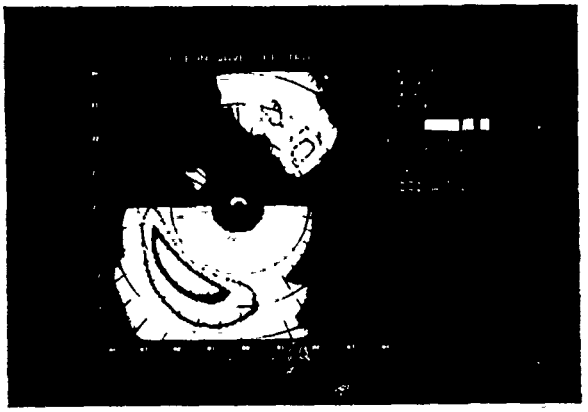
SAR and Buoy Observations

The following is a sample of some of the results of our investigation. In each of the figures, the horizontal axis (-.04 m^{-1} to .04 m^{-1}) is along the azimuth, or along-track direction, and the vertical axis (-.04 m^{-1} to .04 m^{-1}) is along the range, or cross-track direction. In comparing SAR and buoy measurements, one must remember that the buoy *does not* give the *'true'* spectrum. It corrupts the true spectrum in its own fashion, especially with regards to beamwidth. Therefore, we are comparing two different measuring instruments and trying to infer the input conditions which caused the observed measurements.

Directional Wavenumber Spectra - Fig. 2 shows an enlarged, split-screen view of the SAR and Pitch-Roll buoy wavenumber spectra for rev. 1126, Sept. 13, 1978. Both SAR (top) and buoy (bottom) spectrum peaks have been aligned to the same intensity level for comparison. Since the SAR has a 180° ambiguity in wave direction, the SAR spectrum is actually double-sided, however we display only the top portion so as to compare to the buoy's spectrum. The contour intensity level varies over a 20 dB range. The bright center of each peak is within 1 dB of the peak.The outside edge of the first contour corresponds to -3 dB from the peak and the second outside edge -6 dB from the peak. intensity profile, or slice along the peak directions. Here the intensity is over a 10 dB range.

Comparing the spectra, some interesting features can be noted. First, the SAR has a main peak at a wavelength and direction of 278 m, 220.6°, and a secondary peak at 228 m, 248° of lower intensity. The buoy peak is at 257 m, 231.5°. The two peaks are resolved in the SAR spectrum since the SAR exhibits a much narrower beamwidth than the buoy. However, note the slight deformation of an otherwise, angularly symmetric buoy response about the dominant wave direction. This is due to the buoy responding to the secondary peak in the wave spectrum. In addition, note the similar change in intensity with wavenumber around the origin in both the SAR and buoy spectrums. The inset in the lower right hand corner shows the profile of the spectra along their respective peak directions and the shift of the buoy's response to higher wavenumbers relative to the SAR's response.

Fig. 2



Wavenumber Spectrum Profiles- Fig. 3 shows the SAR and buoy wavenumber spectrums which were computed from rev. 1126 shown in fig. 2 by taking a slice of data in |k| along the respective peak directions of the SAR and buoy 2-D spectrums. Note closely the distinctive features in each curve. Even though the SAR's response is shifted from the buoy's response, one can identify similar *kinks* in each which appear to approximately line up if the SAR's curve were shifted. This is remarkable considering the fact that these curves were derived from entirely different measurement techniques, one remote sensing and one *in situ*. When compared on a omnidirectional basis, the SAR and buoy spectrums virtually overlayed on top of one another. Another feature of the spectra in fig.'s 2 and 3 is the dynamic range of the SAR spectra at high wavenumbers. This dynamic range is due to high quality digital imaging of the SAR data from John Curlander at JPL and careful image processing therafter. A comparison of these results with the spectra of Vesecky *et al.* [8] derived from optically imaged SAR data demonstrates the improvement.

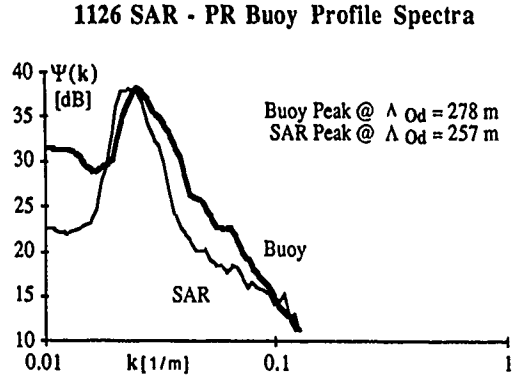


Fig. 3

Measurement of Ocean Parameters - In fig.'s 4,5 we show the correlation between SAR and buoy estimates of ocean parameters for the orbits in which waves were detected. The parameters shown are peak directional wavelength (along the peak direction) and wave direction relative to the SAR's azimuth, or along track direction. When the omnidirectional wavelength was used in the comparison, the results were ≈ as good as the directional results. The straight line is at 45° indicating a one to one relationship. In each case, the SAR estimate is in close agreement with the buoy estimate. Included in each plot are the average percentage and rms error.

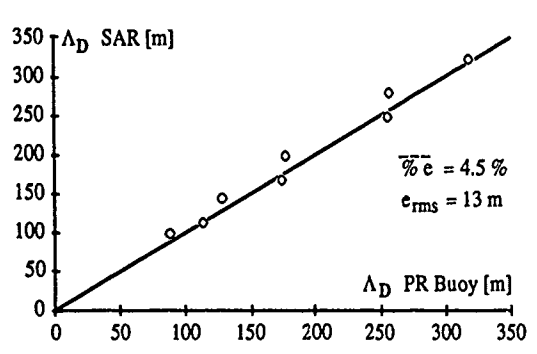


Fig. 4

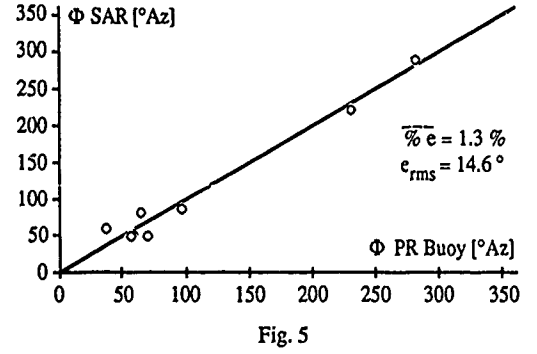


Fig. 5

Windrows

Background - As mentioned in the introduction, we have observed what appear to be periodic, wind-aligned *streaks* in the SAR image spectrum which we will refer to as *windrows*. The theory governing windrows is still in the process of development even though many observations have been reported in the literature [6]. Understanding these windrows is important since 1) they may be related to Langmuir circulations and thus, the formation of the ocean mixed layer and 2) they may be related to atmospheric wind flow, thus provide the means through which SAR can measure the surface wind field and other parameters, such as the mixed-layer depth. Leibovich [6] discusses the dynamics of *Langmuir circulations*, which are roll vortices beneath the water surface causing these windrows on the ocean surface with spacings of 2 m to several 100 m. This circulation is shown schematically in fig. 6 where the arrows below the ocean surface indicate fluid motion. Current research includes the study of the spacing, penetration depth, formation times, and laboratory experiments on these circulations. A number of researchers, including Langmuir [5], have thought the windrow spacing to be proportional to the penetration depth of the cellular pattern which they mark. Originally it was thought L/D = 2. However, measurements of L/D have produced values in the range .66 - 1.66. Also, the orientation of these windrows is nearly always aligned with the wind direction, with the maximum angular difference between streak pairs ≤ 20°.

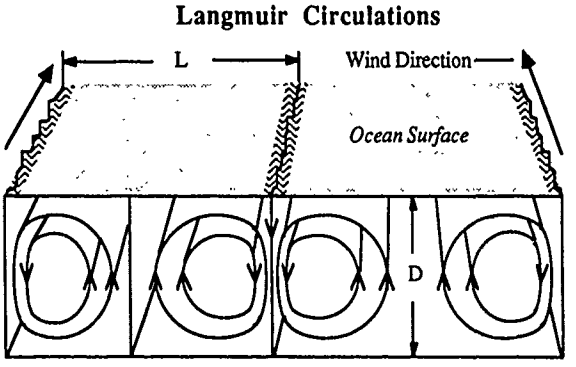


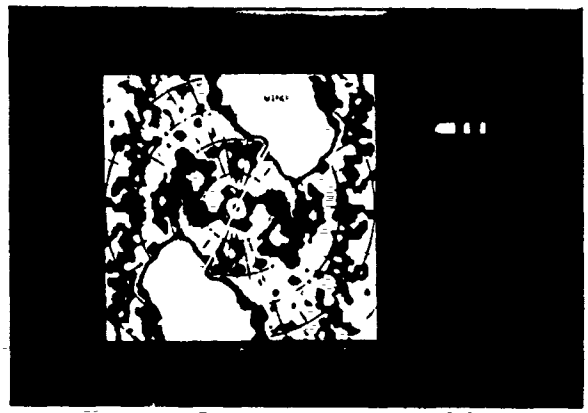
Fig. 6

Another mechanism for these windrows observed in SAR image intensity spectrum may be *atmospheric roll vortices*, where the axes of the roll vortices are aligned with either the geostrophic wind or average planetary boundary layer (PBL) wind direction. In this case, the windrow spacing L is in the kilometer range, with D being the planetary boundary layer height. Again, it is thought L/D to be a constant ≈ 2. Gerling [2] has used the SEASAT SAR radar backscatter data to estimate wind speed and direction using this hypothesis, with the SASS (SEASAT-A Scatterometer System) measured wind field as a reference.

Observations - In fig. 7 we show the spectrum of SAR windrow features which we suspect are due to Langmuir circulations. This display was generated by taking the spectrum for rev. 1126 in fig. 2 and shifting the color intensity scale to lower values to enhance less intense features. This results in the main peaks shown before to appear saturated. Notice the two new peaks, which appear at locations 474 m (.0133 m^{-1}), 141° and -39°, straddle the measured wind direction indicated by the line at 62°. The deviation of these peaks from the wind direction is 79°, thus the average streak direction (perpendicular to the peak direction) is within 11° of the wind direction. Since the wind direction is important in this measurement, SASS wind speed and direction were used as a check on the buoy estimates. The inset at the lower right corner shows a profile or slice along the peak direction over a 10 dB range. In all of the revs. studied, each

displayed similar directional behavior albeit at different wavelengths. However, in some cases the average streak directions were not quite aligned with the wind direction, with differences ≈ 50°. Also, notice the asymmetry in the low wavenumber region around the origin. By going back to the raw spectrum data and averaging over 9 points, one observes small peaks on the order of 1 km in wavelength, and approximately at the same angle as the other windrow peaks previously discussed. These are the features Gerling [2] used to estimate wind direction.

Fig. 7



Next, to see what relationship exists between SAR measured windrow parameters and *in situ* ocean measurements, we obtained XBT (Expendable Bathythermograph) data which provides the temperature as a function of depth over 0 - 800 m at the desired place in the SAR swath. A graph containing three of these curves fitted with a spline, over a time period straddling the SEASAT overpass time, was created and the mixed layer depth measured as shown in fig. 8. The additional XBT curves, however had different spatial locations relative to the curve corresponding to the SAR overpass time. From fig. 8, the mixed layer depth was ≈ 50 m, which corresponds to fig. 7. The spacing L = 474 m was given above. Computing L/D for this rev. yields

$$\frac{L}{D} = \frac{474\ m}{50\ m} = 9.48 >> 2$$

Therefore, this ratio does not seem to fit the current Langmuir circulation theory [5].

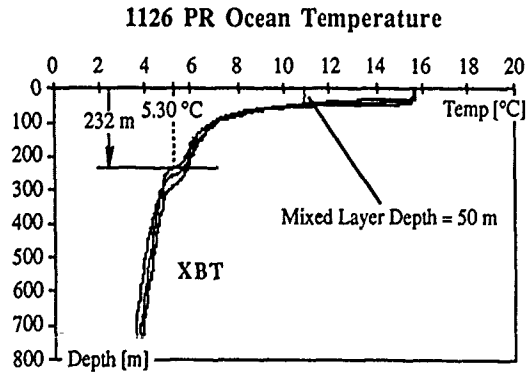


Fig. 8

However, instead of the mixed layer depth, suppose we use a depth D of 232 m, noted in fig. 8. This depth was found by locating where the second derivative of the XBT curve, corresponding to the SAR overpass time, crosses zero, ie., an inflection point. With the same windrow spacing, L/D becomes

$$\frac{L}{D} = \frac{474\ m}{232\ m} = 2.04$$

which is quite close to 2. We used this approach on all the revs. which displayed windrow type features in their spectrums and for which we could obtain XBT data. When the windrow spacing was plotted against this inflection point depth for the revs. in question, the result showed some scatter but was near two. More data for differing ocean thermal structure will be necessary to demonstrate this relationship convincingly.

Conclusions

We find that for the GOASEX data set the following conclusions can be made :

1. SAR and buoy measurements of dominant directional and omnidirectional wavelength, and dominant wave direction are in excellent agreement. On a two-dimensional basis, the SAR image spectrum intensity faithfully reproduces a significant number of features seen in the buoy spectrum. Additionally, the SAR displays features that the buoy does not appear to respond to.

2. Analysis of ocean wave imaging theories by Alpers, Harger, and Rotheram shows them to be similar in many respects, although their mathematical formalism makes them appear quite different.

3. Periodic wind aligned *streaks* are observed in the SAR image spectrum intensity which we have referred to as windrows. Although these features have been observed in a large number of SAR images, more research will be needed to identify the imaging mechanism for these windrows.

Acknowledgments

We would like to thank the follwing people for their help : John Curlander at JPL for the SAR images and ONR (Hans Dolezalek) for financial support.

References

[1] Alpers, W.R., Ross, D.B., Rufenach, C.L., *"On the Detectability of Ocean Surface Waves by Real and Synthetic Aperture Radar"*, JGR, Vol. 86, No. C7, pp. 6481-6498, July 20, 1981.

[2] Gerling, T.W., *"Structure of the Surface Wind Field from the SEASAT SAR"*, JGR, Vol. 91, No. C2, pp. 2308-2320, Feb 15, 1986.

[3] Harger, R.O., *"The Synthetic Aperture Radar Image of Time-Variant Scenes"*, Radio Sci., Vol. 15, pp. 749-756, 1980.

[4] Harger, R.O., *"Synthetic Aperture Radar Systems : Theory and Design"*, 1970.

[5] Langmuir, I., *"Surface Motion of Water Induced by Wind"*, Science, Vol. 87, pp. 119-123, 1938.

[6] Leibovich, S., *"The Form and Dynamics of Langmuir Circulations"*, Ann. Rev. Fluid Mech., Vol. 15, pp. 391-427, 1983.

[7] Rotheram, S., *"Ocean Wave Imaging by SAR"*, Satellite Microwave Remote Sensing, pp. 155-186, 1983.

[8] Vesecky, J.F., Stewart, R.H., Shuchman, R.A., Assal, H.M., Kasischke, E.S., Lyden, J.D., *"On the Ability of Synthetic Aperture Radar to Measure Ocean Waves"*, Wave Dynamics and Radio Probing of the Ocean Surface, pp. 403-421, 1986.

Pier-based Measurements of Ocean Wave Spectra Using Three-Frequency K_u- and L-Band Radars

D. L. Schuler and W. C. Keller

Naval Research Laboratory
Washington, D.C. 20375
U.S.A.

ABSTRACT

A new three-frequency microwave scatterometer technique has been synthesized from theory developed earlier for the related dual-frequency instrument. The new three-frequency technique completely removes many of the limitations on the measurement of ocean wave spectra by this type of microwave instrument. It does so by converting the entire received backscatter return into a new two-channel correlation signal modulated by long ocean waves. This new output signal is directly proportional in strength to the area illuminated by the antenna footprint. Signal detectability is now only limited by system thermal noise and phase stability. Extensive experimental measurements using both L-band and K_u-band radars have been made from a pier-site located on the coast of the Atlantic Ocean at Duck, North Carolina. The measurement campaigns conducted during 1984 and 1988 will be reported on. Ocean wave spectra were compared with non-directional in-situ Baylor Gauge data during all periods.

1. Introduction

There is considerable interest within the oceanographic and remote sensing communities in the building of an instrument which can measure directional sea surface wave spectra. An ideal instrument should be sensitive enough to allow accurate measurements to be made on virtually the entire ocean wave spectrum, even down to wavelengths as short as one meter. In addition, the directivity of such a system should exceed that obtainable from buoy measurements. Finally, the system should be compatible with airborne and spaceborne platforms for large-scale, synoptic, remote sensing of the seas.

2. Origin of the New Technique

The three-frequency scatterometer is a recently conceived microwave remote sensing technique which theoretically has the ability to accurately measure directional ocean wave spectra. The three-frequency technique was developed through a re-examination of the theory behind the related dual-frequency scatterometer. Both of these techniques involve scattering areas which are so large that wave features are not resolved. Wave properties are studied instead by spatially resonating with selected waves over the entire area. A non-resonant, clutter background return forms a significant and undesirable part of the dual-frequency scatterometer output. This unwanted contribution severely limits the minimum signal detectability. The new three-frequency technique actually converts this background term into a part of the wave modulated signal, thus creating a system whose detectable signal range is limited only by system thermal noise. The most obvious benefit from the improved signal quality and dynamic range is that accurate measurements can be made on the weaker, low-amplitude wave systems that occur at both ends of the spectrum. A knowledge of the short wavelength (15 m or less) waves is important to underwater acoustic-system performance, to studies of the early stages of growth of wind-wave systems, and the the detection of surface manifestations of internal waves.

The three-frequency scatterometer concept is closely related to the well-developed dual-frequency technique which has been used to measure both directional wave spectra and surface current flows. Both types of scatterometers operate coherently in the microwave band and transmit carrier signals separated in frequency by some Δf which is in the megahertz range. Return signals due to each of the transmitted signals are separately received and then beat together. The power spectrum of the resultant output when any pair of signals is beat together contains the sum of a broad background spectrum and a sharp resonance line. The background spectrum is the result of a convolution of the Doppler spectra of the two received microwave signals. The sharp "Δk" line, on the other hand, is the result of a resonance between the beat pattern of the two electromagnetic signals and the modulation pattern of the short ocean surface waves (centimetric "Bragg waves") responsible for the backscatter.

The ratio of integrated intensities of the Δk-line and the background spectrum, χ, has been shown to be related to the surface wave slope spectral density $S(\bar{K}_w)$ by the equation

$$\chi = \frac{2\pi^2 \mid m \mid^2 S(\bar{K}_w) coth^2(k_w d)}{A} \qquad (1)$$

where d is water depth, A is illuminated area, and m is the modulation transfer function relating the received power modulation to the slope of the long surface wave. The inverse dependence of χ on A has been observed experimentally and the equation has been shown to account well for observed values of χ. The three-frequency scatterometer was conceived of as a solution to the problem of inferior signal to background performance which characterized the two-frequency method. Calculations indicate that the three-frequency technique will produce signal-to-background ratios 16 dB better than the dual-frequency technique operating from a spaceborne platform.

3. Theory of the Technique

The three-frequency scatterometer concept is most easily understood by considering a microwave system transmitting four carrier frequencies either simultaneously, or in rapid time sequence. Each of the pairs of signals separated by Δf could be processed to obtain a conventional dual-frequency scatterometer output If, however, we allow the separation δf between these two pairs to become equal to Δf then high signal quality three-frequency scatterometer measurements can be made.

The processing of the three-frequency output signals can be carried out quite simply in the time domain. Returns due to the three equally-spaced transmitted frequencies, E_1, E_2, and E_3, are first multiplied in pairs and low pass filtered to yield $E_1E_2^*$ and $E_2E_3^*$ (called P_n and P_1). These resultant products are then multiplied together and low pass filtered to yield $\overline{P_{nl}}$, the desired mean value. Thus, the Fourier transforming techniques employed above to obtain frequency spectra are, in fact, not really required to process three-frequency data to obtain ocean wavenumber spectra. The value $\overline{P_{nl}}$ is important because it is a quantity which, as we will see, can be directly related to wave slope (or height) spectral density for a water wavenumber $K_w = 2\Delta k \cos\theta$. It can be shown (Schuler, et al., 1984) that an expression for a normalized value of P_{nl} may be given by,

$$\frac{P_{nl}}{P} = \frac{2\pi^2 |m(\Delta k)|^2 S(\Delta k) \coth^2(\Delta K d)}{A} \qquad (2)$$

where $\Delta K = \frac{2\pi\Delta f}{c}$ (c is the speed of light),

$m(\Delta k)$ = modulation transfer function,
$S(\Delta k)$ = wave slope spectral density,
A = scattering cell area,
d = water depth,
P = system output when $\delta f = \Delta f = 0$.

Equation (2) indicated that the system output is proportional to wave slope. Wave slope and wave height spectra $F(\Delta k)$ are, however, related by

$$S(\Delta k) = (\Delta k)^2 \cdot F(\Delta k) \qquad (3)$$

Combining Equations (2) and (3) we may solve for the ocean wave height spectra $F(\Delta k)$,

$$F(\Delta k) = \frac{A}{2\pi^2 (\Delta k)^2 |m(\Delta k)|^2 \coth^2(\Delta k d)} \frac{P_{nl}}{P} \qquad (4)$$

The value of $m(\Delta k)$ is, in reality, a complicated function of both geophysical and microwave parameters. Its value, however, has been experimentally determined for many cases (particularly at L-Band and K_u-Band) and Equation (4) may then be used to determine $F(\Delta k)$ from the three-frequency scatterometer outputs.

4. Experimental Verification of the Concept

Extensive experimental measurements using both L-band and K_u-Band radars have been made from an Army pier-site located on the coast of the Atlantic Ocean. Two separate measurement campaigns conducted during 1984, and 1988 will be reported on. Ocean wave spectra compared favorably with non-directional in-situ Baylor Gauge data during all periods.

The ultimate objective of this program was to develop an airborne wave spectrometer. The early L-band measurements were a necessary first step which led to the development of a K_u-band three frequency system which was more suitable for aircraft deployment. This K_u-band system was first successfully tested in 1988.

The prototype L-band version of the three-frequency scatterometer was built and experimentally tested at the Army Coastal Engineering Research Center (CERC) Field Research Facility at Duck, North Carolina. Measurements of wave spectra were carried out during the period 4-8 June 1984 from a site at the end of the pier 1840 feet out in the Atlantic Ocean. Data runs were scheduled to straddle the thirty-four minute data acquisition periods used by CERC for the Baylor gauges mounted along the pier. These Baylor gauges yield a non-directional estimate of wave height (vs wave frequency) and were used as a source of comparative data for the scatterometer. The data were made comparable by converting the Baylor gauge data to a wavenumber spectrum and by correcting the three-frequency scatterometer data for (1) depth differences, (2) directionality, and by (3) assuming a value for the L-band (or K_u-band) modulation transfer function $m(\Delta k)$ which was correct for the environmental conditions that were present.

The series of experiments (4-8 June 1984) was performed at the CERC Duck, NC pier to determine if multiplexing, or parallel processing, of the wave spectral information could be utilized to dramatically reduce the overall data acquisition period. Figure 1 gives an example of a de-multiplexed wave height spectra derived from data taken on 5 June 1984. The wind was from the southeast at 15 kts during the measurements and the significant wave height was 0.57 m. The scatterometer and the Baylor gauge spectra agree in both magnitude and general shape. The 9.75 second (77 meter) dominant wave was detected by both systems.

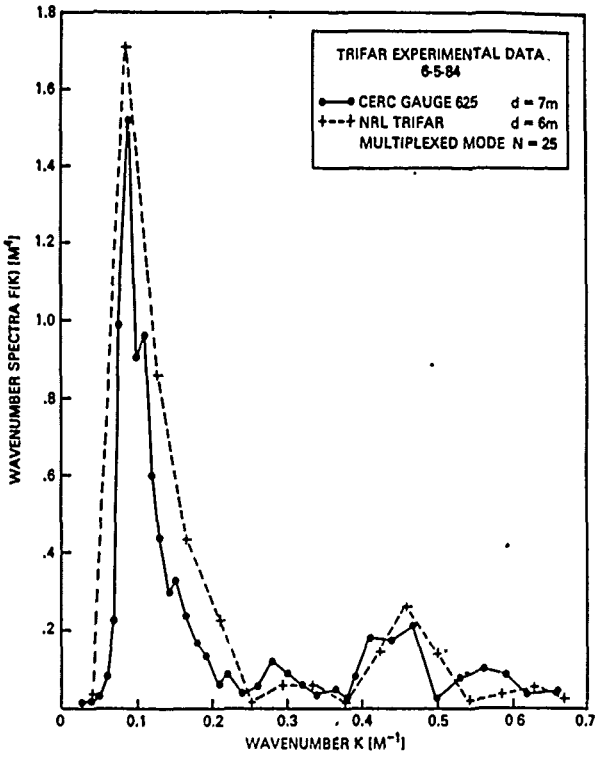
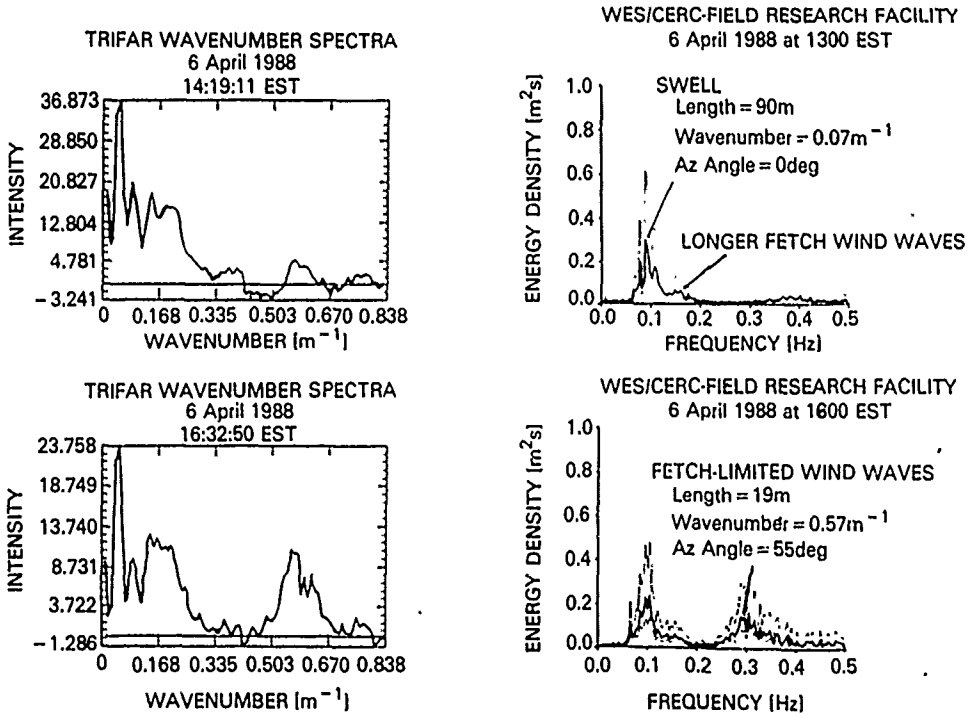


Fig. 1 — Example of de-multiplexed three-frequency scatterometer data (5 June 1984).

In April, 1988 a K_u-band version of the three-frequency wave spectrometer was tested at CERC. This new system operated by transmitting 80 triads of frequency $\{+\Delta F, -\Delta F$ and a separate $\Delta F = 0$ signal$\}$ in 16 ms. These bursts of frequency are repeated and the returns averaged. For the measurements reported on here each wave spectral measurement is an average of 1280 independent returns. Thus, 80 highly averaged, long wave spectral estimates were carried out in approximately 52 seconds. Longer, more stable averages can, of course, be achieved by concatenating sequential data-files of this same size.

Figure 2 shows samples of wave spectra which were measured by the K_u-band system and the CERC Baylor gauge on 6 April 1988 dur-



TEMPORAL DEVELOPMENT OF A FETCH-LIMITED WIND WAVE SYSTEM

Fig. 2 — Comparison of three-frequency K_u-band scatterometer and Baylor Gauge Wave Spectra (6 April 1988).

ing the passage of an atmospheric front. This event did not produce a storm of significant proportions, but did allow measurements to be made on a developing wind-wave system and an existing swell from a different direction. On 6 April shortly after 1400 EST the wind strength increased from ~5 m/s to 12-15 m/s from 160°. Prior to this event the sea had been almost calm with a 90 m visible swell coming from 105°, the result of an earlier storm. By 1600 EST the wind was a reasonably steady 12 m/s and a wind wave system was beginning to develop in this direction. The antenna used (2.5° × 2.5° half-power widths) and the radar pulsewidth used (2 μs) provided an illuminated backscatter footprint of 30 m azimuthally by 300 m radially. Such an elongated footprint does not provide much directional sensitivity in the look direction. Therefore, the sequential plots of wavespectra presented in Figure 2 show peaks associated with both the 90 m swell and the developing windwave spectra. In Figure 2 the long wavelength swell appears at .07 m^{-1}, the windwave spectrum at .35 m^{-1} to .80 m^{-1}. A third spectral peak centered at .167 m^{-1} is also present in the Baylor Gauge records and was produced by waves developing over a longer fetch and propagating nearly parallel to the shore.

5. Conclusions

We have demonstrated through field-experimentation that ocean wave height spectra are measurable using the three-frequency scatterometer technique. The theory that has been developed for the technique is well founded and yields accurate estimates of wave spectra. Furthermore, a multiplexed version of the basic three-frequency scatterometer has been shown to be a practical means of collecting data during acquisition times that are short enough so that naturally occurring dynamic, geophysical changes in the wave field may be studied.

6. References

Schuler, D.L., Plant, W.J., and Eng. W.P., 1981, "Remote Sensing of the Sea Using One- and Two-Frequency Microwave Techniques," Radio Science, Vol. 15, No. 3, pgs. 605-15, May-June 1980.

Schuler, D.L., Plant, W.J., Reeves, A.B., and Eng. W.P., 1984, "Removal of Clutter Background Limitations in Dual-Frequency Scattering from the Ocean: The Three-Frequency Scatterometer," *Int. Journal of Remote Sensing*, Vol. 6, No. 7, pgs. 1091-1112, 1985.

MAPPING BOTTOM TOPOGRAPHY WITH X-BAND SLAR

J. Vogelzang	G.J. Wensink	G.P. de Loor	H.C. Peters	H. Pouwels
Rijkswaterstaat, Tidal Waters Division P.O. Box 20907 2500 EX The Hague The Netherlands	Delft Hydraulics P.O. Box 152 8300 AD Emmeloord The Netherlands	Physics and Electronics Laboratory TNO P.O. Box 96864 2509 JG The Hague The Netherlands	Rijkswaterstaat, North Sea Directorate P.O. Box 5807 2280 HV Rijswijk The Netherlands	National Aerospace Laboratory P.O. Box 90502 1006 BM Amsterdam The Netherlands

ABSTRACT

On January 19th, 1988 an experiment with X-band SLAR has been performed to study the possibilities and limitations for cartographic mapping of sea bottom topography with the aid of imaging radar. The bottom topography of the test area is dominated by sand waves with a height between 2 and 6 m and a crest-to-crest distance of typically 500 m. Simultaneously with the radar measurements wind and wave data were collected on the Noordwijk Tower (MPN) at 20 km from the test area, and the current velocity profile was measured from two ships in the test area. The results of the current measurements agree with the continuity equations. It is shown that the extremes in the radar backscatter are located right above regions with maximum bottom slope, with an error in position of two pixel sizes (30 m), in agreement with simple relaxation models like the model of Alpers and Hennings (1984).

Key words : bottom topography
X-band SLAR
imaging mechanism
sand waves
imaging radar

1. INTRODUCTION

It is known for some time now that under suitable conditions (moderate wind and strong tidal current) the bottom topography of shallow seas can be made visible with imaging radar (Side Looking Airborne Radar, SLAR and Synthetic Aperture Radar, SAR). This phenomenon was discovered in 1969 by de Loor (de Loor and Brunsveld van Hulten, 1978 ; de Loor, 1981) After the SEASAT mission in 1978 the phenomenon provoked considerable interest, both from theoretical and experimental side.

The mapping of sea bottom topography is studied in the Netherlands as a part of the Netherlands National Remote Sensing Program (NRSP). On January 19th an experiment has been performed to study the usability of radar imagery of the sea for cartographic purposes. The experiment was set up in such a way that it is possible to compare the predictions of existing models of the imaging mechanism with the experimental results. In this work we will concentrate ourselves on the positional relation between radar images and bathymetric maps.

2. OVERVIEW OF THE THEORIES

It is now generally accepted that the imaging mechanism consists of three steps :

1. interaction between (tidal) flow and bottom topography produces modulations in the surface current velocity, usually described with the continuity equations (mass conservation).
2. modulations in the surface current velocity cause variations in the wind-wave spectrum as described by the action balance equation.
3. variations in the wave spectrum express themselves as spatial modulations in the backscattered intensity.

Note that if first order Bragg scattering is assumed to be the dominant scattering mechanism, the backscattered intensity is proportional to the wave height spectrum at the Bragg wave number k_B given by

$$k_B = 2\,k\,\sin(\theta) \quad , \qquad (1)$$

where k is the wave number of the incident radar radiation and θ the angle of incidence.

Models of this kind have been proposed by Alpers and Hennings (1984), Shuchman, Lyzenga and Meadows (1985), Holliday, St-Cyr and Woods (1986), van Gastel (1987a) and Hennings (1988). The central point in these models is the second step : the solution of the action balance equation. Van Gastel (1987a) includes the effect of wind input, dissipation and non-linear wave-wave interactions in a rigorous way into the source term of the action balance equation, while the other models (further referred to as simple relaxation models) employ a simple form for the source term.

These simple relaxation models describe SEASAT L-band data well. At X-band, however, these models underestimate the radar backscatter by an order of magnitude at least. This can be explained by nonlinear wave-wave interactions (van Gastel, 1987a) or other scattering mechanisms than first order Bragg (Holliday, 1986 ; Hennings, 1988).

Vogelzang (1989) has compared the simple relaxation models by solving the action balance equation numerically. He has shown that for modulations of 10 % in the current velocity these models produce similar results for the hydrodynamic modulation. In particular, advection can be neglected at X-band for wind speeds exceeding 5 m/s. Therefore the Alpers and Hennings (1984) model should be applicable in this case. According to this model, the radar backscatter is proportional to $-\nabla_{per}(d)/d^2$, with d the depth and $\nabla_{per}(d)$ the gradient of the depth perpendicular to the sand wave crests in the direction of the current

velocity. As a consequence, the extremes in the radar backscatter are located right above regions with maximum bottom slope.

3. DESCRIPTION OF THE EXPERIMENT

For reasons of completeness a short description of the experiment is given here. More information can be found elsewhere (Vogelzang et al., 1988).

The experiment was performed in a test area measuring 5 x 5 km, 20 km North West to the Noordwijk Tower (MPN), 10 km off the Dutch coast. The bottom topography in this area is dominated by sand waves with a height between 2 and 6 m and a crest-to-crest distance of typically 500 m at a depth of about 23 m. The sand wave crests are oriented perpendicular to the coast line and perpendicular to the dominating current direction.

An ordinary bathymetric map of the test area was digitized. Using spline interpolation techniques it is possible to construct maps of depth and bottom slope.

The test area was recorded from four different flight directions (two parallel and two perpendicular to the sand wave crests), to study the effect of the flight direction relative to the bank and wind direction. Two flights were made, at opposing directions of the current velocity, to study the effect of the current velocity direction.

Simultaneously the current velocity profile was measured from two ships, the 'Octans' and the 'Smal Agt'. The position of the 'Octans' was recorded with high precision using a HYPERFIX positioning system. Wind, wave and temperature data were collected at MPN.

The radar measurements were performed using the Dutch Digital SLAR. With this X-band SLAR system and the processing facilities at the National Aerospace Laboratory, it is possible to obtain both geometrically and radiometrically correct images. The pixel size of the resulting images is 15 x 15 m; the radiometric resolution 0.2 dB.

A long radar pulse has been chosen to increase the sensitivity of the radar system. As a consequence, the images are not absolutely radiometrically calibrated. Since we are mainly interested in relative differences, and the images are still relatively radiometrically calibrated, this is not considered a serious drawback.

4. RESULTS AND DISCUSSION

4.1 Current velocity

The current velocity profiles measured during the experiment were processed using Davies' method (Davies, 1988) to obtain the averaged (in the vertical) current velocity. According to the continuity equations, the components of the current velocity parallel and perpendicular to the sand wave crests, U_{par} and U_{per}, are given by

$$U_{par} = C_{par} \quad , \tag{2a}$$

$$U_{per} = C_{per}/d \quad , \tag{2b}$$

with d the depth and C_{par} and C_{per} constants, depending only on time. This means that both the averaged current velocity parallel to the sand wave crests and the discharge perpendicular to the sand wave crests are constant, with the discharge D defined as

$$D = U_{per}\, d\, (1\ m) \quad . \tag{3}$$

The factor (1 m) has been included in (3) to give D the dimension m^3/s.

Figure 1a shows the discharge D for the current measurements from the ships, while figure 1b shows the current velocity parallel to the sand wave crests U_{par}. The error in the magnitude of the averaged current velocity was estimated 10 %, the same as the error in the measured current velocities. The error in the direction of the current velocity was 5°, the error in the sand wave direction also 5°. From figures 1a and 1b one can see that the measurements from the two ships overlap within the error bars, in agreement with the continuity equations, except around 18:30 GMT when the current velocity vanishes. These results agree with those of similar measurements by van Gastel (1987b).

However, due to the large errors in the measurements this can not be regarded as conclusive evidence for the validity of the continuity equations. Moreover, it can not be concluded whether or not the maximum current velocity is located right above the sand wave crests. More precise and elaborate measurements are needed to settle this point.

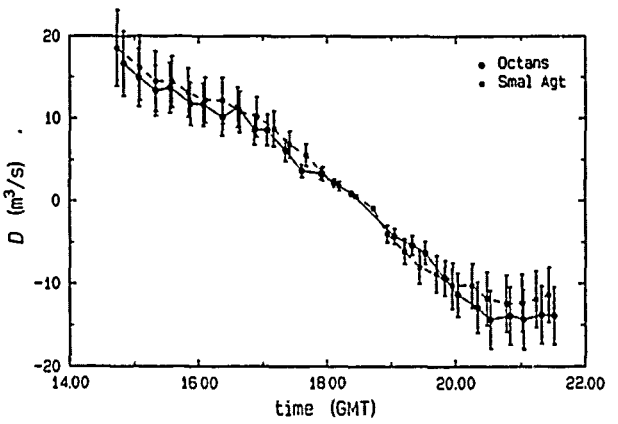


Figure 1a. The discharge D as a function of time.

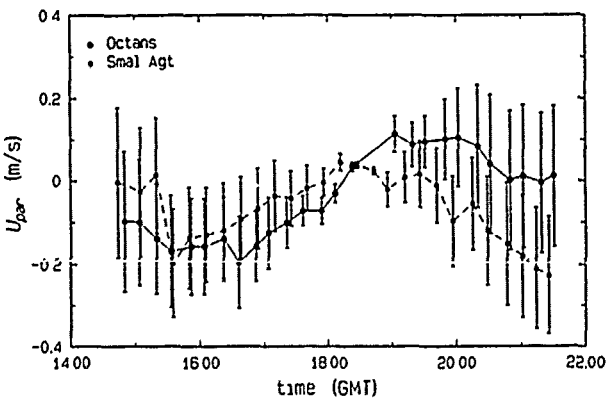


Figure 1b. The current velocity parallel to the sand wave crests, U_{par}, as a function of time.

Image no.	Recording time begin	end	Image dir.	Wind speed	dir.	Current speed	dir.
1	14:46:59	14:49:01	-62°	8.3	165°	0.8	20°
5	20:33:07	20:35:09	-60°	7.4	197°	0.6	-160°

Table 1. Some flight data and ground measurements.

4.2 Relation between radar images and bathymetric maps

A total number of eight radar images has been obtained during the experiment. Two of them, images 1 and 5, have the flight direction parallel to the sand wave crests and the range direction in the upwind/-crosswind direction. These images show the sand waves most clearly and were selected for further study. Some flight data and ground measurements relevant for these images are listed in table 1.

The recording times, written as hh:mm:ss, are in GMT. The image direction roughly corresponds to the flight direction, the difference between the two due to the wind. All speeds are in m/s, all directions in degrees true North.

The quality of the images is rather poor, due to the low sensitivity of the radar system. Moreover, the meteorological conditions were stable by 1 - 2 °C, which may affect the quality of the images in a negative way, because the radar backscatter of the sea decreases with increasing stability. The images were filtered with a 3 x 3 median filter to exhibit the sand waves more clearly.

Using the position data of the 'Octans' it is possible to fix the position in the radar images with an error of two pixel sizes (30 m) at most. Digitized maps of depth and bottom slope with the same scale as the radar images were constructed. Figures 2a and 2b show parts of images 1 and 5, both 129 x 129 pixels, with contour lines of $\nabla_{per}(d)/d^2$. The sand wave crests were taken parallel to the image direction. The ships are visible as orange spots in the lower left corner.

As can be seen from these figures, regions where $\nabla_{per}(d)/d^2$ is minimal correspond with high backscatter (light area's) in figure 2a and regions where $\nabla_{per}(d)/d^2$ is maximal correspond with low backscatter (dark area's) in figure 2b. This is in perfect agreement with the predictions by simple relaxation models.

Note that $\nabla_{per}(d)/d^2$ is to good approximation proportional to the bottom slope since the change in depth is small compared to the change in slope. Therefore the radar images show the bottom slope. The banded structure of the radar images is caused by the fact that the sand waves have an asymmetric profile : they look more like a saw-tooth than a sine. The steeper slopes show up as light (image 1) or dark (image 5) bands, while the gentle slopes act as a background. The images look like each others negative because of the change in current direction. It should be noted that not all features on the radar image are related to bottom structures. This is considered to be due to the quality of the images. The modulation in the radar images is about 4 dB.

To quantify the relation between the radar images and the map of $\nabla_{per}(d)/d^2$, the correlation coefficient between image and map has been calculated. Figure 3

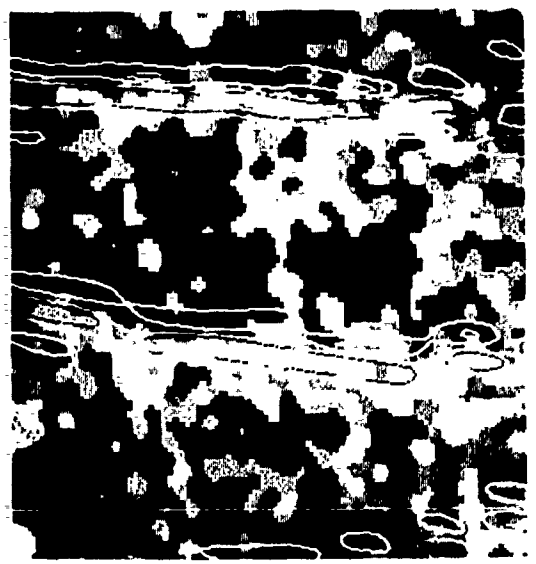


Figure 2a. Part of radar image 1 with contour lines of $\nabla_{per}(d)/d^2$. The image direction is from left to right, the range direction from bottom to top. The values of $\nabla_{per}(d)/d^2$ are 10^{-4} (green), $5\ 10^{-5}$ (orange), $-5\ 10^{-5}$ (yellow) and -10^{-4} (red). The current is directed from top to bottom.

Figure 2b. Part of radar image 5 with contour lines of $\nabla_{per}(d)/d^2$. See figure 2a. The current is directed from bottom to top.

shows the result of this calculation for images 1 and 5. The correlation coefficient C_j, with j the shift of the radar image relative to the map in the range direction, has for both image 1 and 5 a clear minimum at j = 1. This means that the minimum correlation occurs if the radar image is shifted one pixel upward relative to the map of $\nabla_{per}(d)/d^2$. Since the error in position is two pixel sizes, this implies that the extremes in the radar backscatter are located right above regions with maximum bottom slope, with an error of two pixel sizes (30 m).

The quality of the images is not good enough for inverse modelling (calculation of the bottom slope from the radar image).

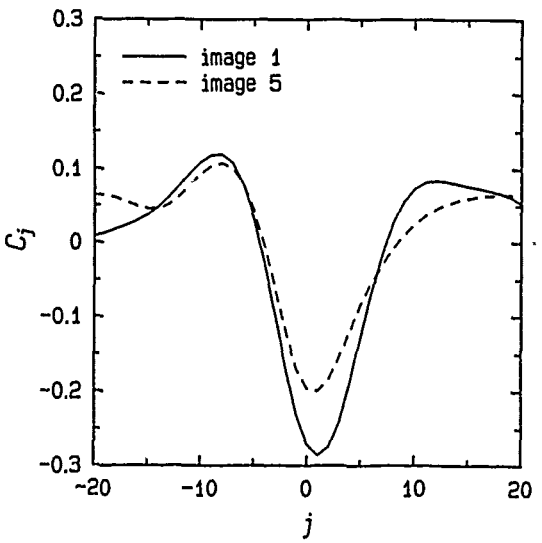


Figure 3. Correlation coefficient C_j between radar images and $\nabla_{per}(d)/d^2$.

5. CONCLUSIONS

From the results of the experiment described in this paper the following conclusions are drawn :

- The measured current velocities agree with the continuity equations.
- The extremes in radar backscatter are located right above regions with maximum bottom slope for X-band at winds exceeding 7 m/s. This is in perfect agreement with predictions by simple relaxation models.

ACKNOWLEDGEMENTS

This work has been financially supported by the Netherlands Remote Sensing Board (BCRS) as part of the Netherlands National Remote Sensing Program (NRSP).

REFERENCES

1. Alpers, W. and Hennings, I., "A theory of the imaging mechanism of underwater bottom topography by real and synthetic aperture radar", J. Geophys. Res., 89 C, pp 10529-10546, 1984.

2. Davies, A.M., "On formulating two-dimensional vertically integrated hydrodynamic numerical models with an enhanced representation of bed stress", J. Geophys. Res., 93 C, pp 1241-1263, 1988.

3. Van Gastel, K., "Imaging by X-band radar of subsurface features : a nonlinear phenomenon", J. Geophys. Res., 92 C, pp 11857-11865, 1987a.

4. Van Gastel, K., "Velocity profiles of tidal currents over sand waves", Neth. J. Sea Research, 21, pp 159-170, 1987b.

5. Hennings,I., "Abbildung von submariner Bodentopographie auf Luft- und Satellitbildern im Mikrowellenbereich und im sichtbaren Bereich des elektromagnetischen Spectrums", Ph. D. Thesis, University of Bremen, Bremen, FRG, 1988.

6. Holliday, D., St-Cyr, G. and Woods, N.E., "A radar ocean imaging model for small to moderate incidence angles", Int. J. Remote Sensing, 7, pp 1809-1834, 1986.

7. De Loor, G.P. and Brunsveld van Hulten, H.W., "Microwave measurements over the North Sea", Bound. Lay. Met. 13, pp 113-131, 1978.

8. De Loor, G.P., "The observation of tidal patterns, currents and bathymetry with SLAR imagery of the sea", IEEE J. Ocean. Engin., OE-6, pp 124-129, 1981.

9. Shuchman, R.A., Lyzenga, D.R. and Meadows, G.A., "Synthetic aperture radar imaging of ocean-bottom topography via tidal-current interactions : theory and observations", Int. J. Remote Sensing, 6, pp 1179-1200, 1985.

10. Vogelzang, J., "The mapping of bottom topography with imaging radar : a comparison of the hydrodynamic modulation in some existing models", Accepted for publication in Int. J. Remote Sensing, 1989.

11. Vogelzang, J., Wensink, G.J., de Loor, G.P., Peters, H.C. and Pouwels, H., "Sea bottom topography with X-band SLAR : evaluation of existing models", Proc. IGARSS '88, Edinburgh, Scotland, pp 1303-1307, 1988.

Removal of 180° ambiguity in SAR images of ocean waves

W.Rosenthal[1], F. Ziemer[1], K.Raney[2] and P.Vachon[2]

[1] GKSS Research Centre, West Germany
[2] Canadien Centre for Remote Sensing, Canada

Abstract

Time series of wave pattern images provide the opportunity for the three dimensional spectral analysis. In the present feasibility study a three dimensional spectral analysis was applied to the seven single looks of a SAR. Each single image covers an area of the water surface at a fixed time period (the integration time). The images are separated by a well defined constant time step. This is the necessary condition for the FFT application in the time domain.
In the past this three dimensional analysis technique was used successfully for the processing of wave image measurements from shipborne navigation radar.
The presented results demonstrate the successful remove of the directional ambiguity inherent in traditional two dimensional wave image spectra.

Key words: SAR-images, Ocean Waves and directional ambiguity.

1. Introduction

The 180° directional ambiguity is a problem, that is not restricted to radar, but occurs in any sea surface imaging system. In [1] it was demonstrated, that for the use of images sampled at two successive time steps with the step length τ, the remove of ambiguity could be achieved for $0 < k < \pi^2/g\tau^2$.
In [2] an alternative way to resolve this ambiguity problem was presented. This method was used for the present feasibility study. The concept of the methode rests on a three dimensional description of ocean waves provided by a time series of wave pattern images.
The Fourier presentation of a three dimensional sample is defined over three dimensions as well. Two dimensions are defined by the two wave number components and the third by the frequency. The property of symmetry is inherent in the three dimensional Fourier space as well. But the frequency information provides the tool to separate wave energy with positive frequencies from those with negative. Wave energy with positive frequency gives by its wavenumber vector the unique direction of propagation.
This technique is successfully used in routine in the application with shipborne navigation radar (see: [2] - [5]).

For the analysis of measurements from airborne radar sensors is, because of the lack of resolution in time, the two dimensional Fourier analysis the routine application. For processing the SAR (Synthetic Aperture Radar) imaging normally the full bandwidth in the Doppler shifted radar backscatter is used for enlarging the azimuthal resolution. The theoretical limit for the resolution is given by half of the physical length of the SAR antenna.
For a still sufficient resolution of the wave energy in the neighbourhood of the energy peak of an ocean wave field, the pixel length may be enlarged. Thus by a reduction of the spacial resolution (by splitting the Doppler bandwidth) the possibility to separate a time serie of single look images may be provided. If the time step τ is hold constant the neccessary premise for the application of the three dimensional analysis is given.
The proposed technique was applied to a couple of measurements from the LEWEX'87 data set. This data set presented in [5] provides the unique possibility to compare most of the today used instruments and numerical models for the detection of ocean wave directionality.

2. The used technique

The three dimensional grey level matrix $g(\vec{\Gamma})$ is related to the wave field by a transfer mechanism that is not discussed in this paper. This time series of wave pattern images may be written as a discrete Fourier series:

$$(1)\quad g(\vec{\Gamma}) = \sum_{j=0}^{N-1} G(\vec{\Omega}) e^{i(\vec{\Omega}\cdot\vec{\Gamma})} + c.c. \quad ,$$

where the space - time vector is defined $\vec{\Gamma}=(x,y,t)$ and the corresponding vector in the Fourier domain by: $\vec{\Omega}=(k_x,k_y,\omega_n)$.
The Fourier coefficients $G(\vec{\Omega})$ on the right side of (1) are calculated by a three dimensional FFT:

$$(2)\quad G(\vec{\Omega})=1/C \sum_{j=0}^{N-1} g(\vec{\Gamma})\, e^{-i(\vec{\Omega}\cdot\vec{\Gamma})}$$

The discrete values of the wavenumber - frequency vector are preset by:

$$(3)\quad \begin{aligned} k_x &= i\Delta k_x \,(i=-I/2,..,0,..,I/2-1)\\ k_y &= j\Delta k_y \,(j=-J/2,..,0,..,J/2-1)\\ \omega_n &= n\Delta\omega \,(n=-N/2,..,0,..,N/2-1), \end{aligned}$$

where $\Delta k_x=2\pi/L_x$, $\Delta k_y=2\pi/L_y$ and $\Delta\omega_n=2\pi/T$. L_x and L_y are the total length of the observed scene in x- and y- direction. The total observation time T presets the spacing in frequencies.
The normalised image power spectrum is defined by:

$$(4)\quad P^{(3)}(\vec{\Omega})=|G(\vec{\Omega})|^2$$

The superscript in (4) gives the dimension of the used Fourier space. From (4) the known two dimensional presentation may be deduced by the integration along all frequencies:

$$(5)\quad P^{(2)}(\vec{k})=\int_{-\infty}^{0} P_-^{(3)}(\vec{\Omega})\,d\omega+\int_{0}^{+\infty} P_+^{(3)}(\vec{\Omega})\,d\omega .$$

On the left side of (5) the symmetric spectrum is given, as it comes out the two dimensional FFT used for the analysis of one single time step. On the right side of (5) the integration is given decomposed for negative and positive frequencies.
For the further analysis we consider only positive frequencies for the integration:

$$(6)\quad P_+^{(2)}(\vec{k})=\int_{0}^{+\infty} P_+^{(3)}(\vec{\Omega})\,d\omega .$$

This technique provides unambiguous results beginning at the frequency $\omega=1/2\Delta\omega$ up to the Nyquist limit.

3. An unambiguous two dimensional spectrum derived from SAR measurements

From a SAR measurement of a wave field penetrating through the ice edge east of New Foundland (Canada) a time series of seven single look grey level scenes was provided. The technique for the scene separation is given in [6]. The time step was choosen to be τ=0.76sec. Thus the total observation time is T=4.56sec.

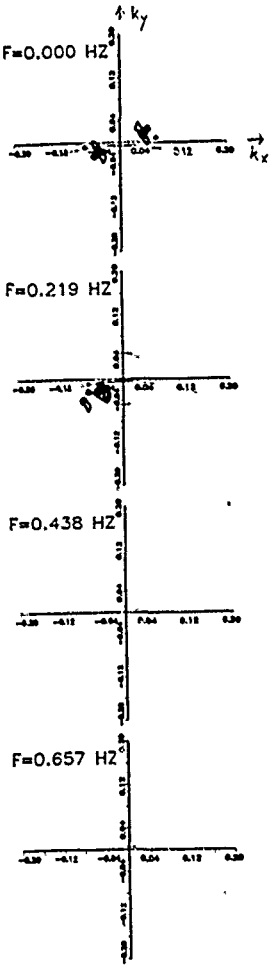


Figure 1.

Isolines of a SAR image spectrum in four wave number planes: at the zero frequency and at the three positive frequencies. The measurement was taken by CCRS.

In this and in all following figures the image power is normalised to 1. Ten isolines are given with a 0.1 distance.

This time series was transformated using a three dimensional FFT. The result, as defined by (4) is given as isolines of normalized image power in figure 1. This figure shows four wavenumber axes. The first plane is the plane with the zero frequency and the following planes are these with positive frequencies.
In the plane with the zero frequency the directional ambiguity is still inherent. This is due to the fact that each plane containes the energy from the intervall:

$$(7)\quad [f_n-1/2\Delta f, f_n+1/2\Delta f]$$

where: $f_n=\omega_n/(2\pi)$. The quality of symmetry at f_n=0 is caused by the accumulation of positive as well as negative frequencies within the same wavenumber bins into this plane. This effect is demonstrated by figure 1.

In the next plane, with the centre frequency of f_1=0.22Hz the energies from 0.11Hz to 0.33Hz are accumulated. This coveres the main part of the ocean wave signal that was detected by this measurement.

Figure 2.
Isolines of a SAR image spectrum in the frequency direction plane. The image spectrum is calculated from the sum over all four planes in figure 1.
The ambiguous part is due to the symmetric energy at f=0.

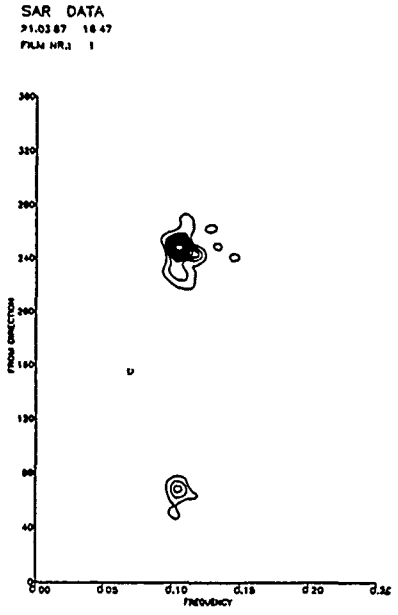
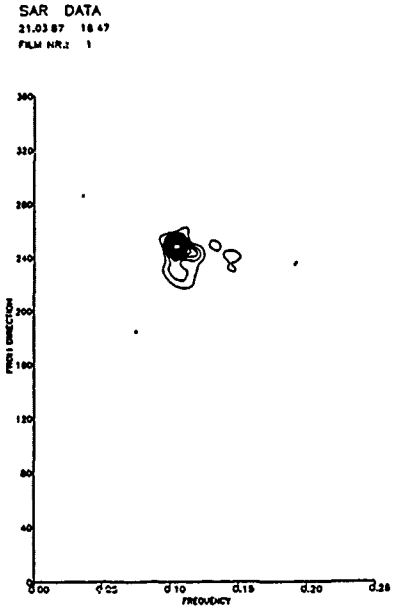


Figure 3.
Isolines of an asymmetric SAR image spectrum in the frequency direction plane. The image spectrum is calculated from the sum over the three positive frequencies in figure 1.
The ambiguous part is suppressed by leaving off the symmetric energy at f=0.



In the next two planes no more energy is to be seen. That means, as ten isolines are given, that the image power here is less then 10% below the maximum.
Figure 2. shows the integrated spectrum using (6) transformated to the frequency - direction space by:

$$(8)\quad F^{(2)}(f,\theta) = (k\ dk/df)\ F_{+}^{(2)}(\vec{k}).$$

The figure 3. gives the resulting spectrum if the zero frequency is left off. This presents unambiguous results for those energies with a frequency higher than 0.11Hz up to the Nyquist limit.

4. The detection of an opposing sea state by a SAR measurement

In figure 4. the result of a measurement gained by the navigation radar onboard the 'Tydeman' is given. For this measurement (during a total observation time of T=66.5sec) 32 timesteps were taken. Therefore the lowest frequency with unambiguous results lies at f=0.0075Hz. This is well apart from the signal caused by the waves. Thus the total resulting spectrum is free of any ambiguity.
This opposing wave system within the same frequency bin was detected by a couple of other instruments and model hindcasts of the LEWEX'87 experiment.

Figure 4.
Isolines of an asymmetric radar image spectrum in the frequency direction plane. The measurement was taken with the navigation radar onboard the vessel "Tydeman" during LEWEX'87.
The image spectrum is calculated from the sum over sixteen planes with positive frequencies.

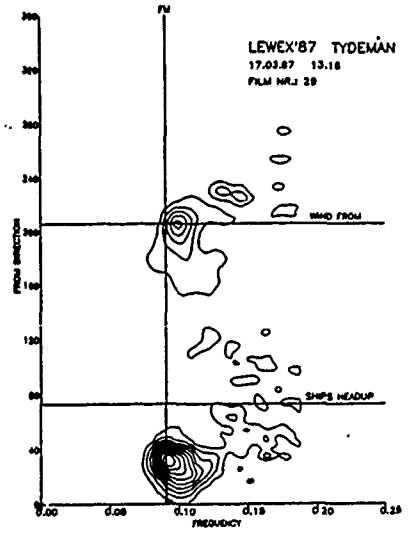


Figure 5.
Isolines of a SAR image spectrum in the frequency direction plane. The image spectrum is calculated from the sum over all four frequencies.
The ambiguous part is due to the symmetric energy at f=0.

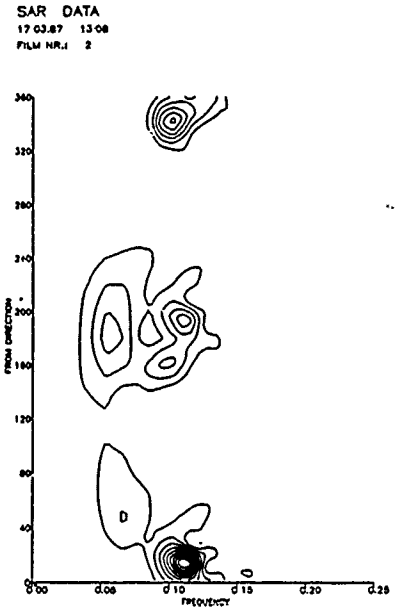
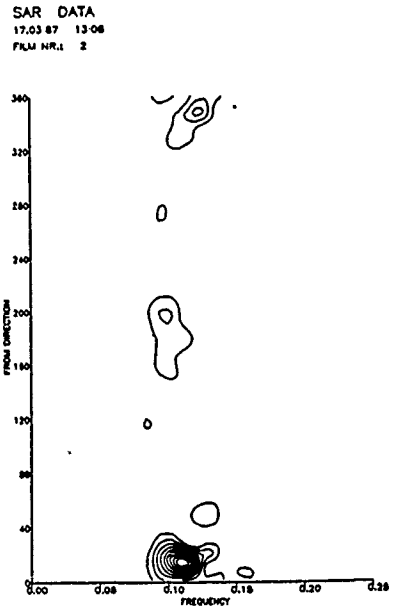


Figure 6.
Isolines of an asymmetric SAR image spectrum in the frequency direction plane. The image spectrum is calculated from the sum over the three positive frequencies.
The ambiguous part is suppressed by leaving off the symmetric energy at f=0.



At the same time as the shipborne measurement was taken an overflight with a SAR was conducted by CCRS at the location of the 'Tydeman'. The result of this measurement is given in figure 5., this corresponds to figure 2. where the energy with zero frequency is summed up as well.
Figure 6. shows the result if the zero frequency is left of. That means there is no ambiguity caused by the Fourier algorithme inherent in this presentation. The opposing wave system is clearly visible in the asymmetric SAR-spectrum. It is not detectable from a symmetric SAR-spectrum that was the only possible SAR product in the anaylsis technique used so far.

5. Conclusions

This feasibility study demonstrated, that by the use of the powerfull three dimensional FFT technique asymmetric wave image spectra may be obtained from airborne SAR measurements. The shown examples make evident that due to the short scene observation time the results in the low frequency bins remain symmetric.
The fundamental progress by using an advanced technique like the three dimensional analysis is evident and encouraging to continue with SAR derived time series of a scene development.

1. ATANASSOV,V., W. ROSENTHAL and F.ZIEMER Removal of Ambiguity of Two Dimensional Power Spectra Obtained by Processing Ship Radar Images of Ocean Waves. J.Geophys.Res., Vol. 90, pp. 1061-1067.

2. YOUNG,I.R., W. ROSENTHAL and F. ZIEMER (1985), A Three Dimensional Analysis of Marine Radar Images for the Determination of Ocean Wave Directionality and Surface Currents. J.Geophys.Res., Vol. 90, pp. 1049-1059.

3. ZIEMER, F. and W. ROSENTHAL (1987), On the Transfer Function of a Shipborne Radar for Imaging Ocean waves. Proceedings of IGARSS'87 Symposium, Ann Arbor, May 1987 pp. 1559-1564.

4. ZIEMER,F., E.STOCKDREHER and H.GÜNTHER (1988), Measured Transfer Functions for Shipmotions in Natural Seaways. Proceedings of OCEAN'88 Conference, Baltimore, Nov. 88. pp. 1212-1217.

5. BEAL, R. C., (editor), (1989), Directional Ocean Wave Spectra. The Johns Hopkins APL Technical Digest, Volume 11, Number 2 (in preparation)

6. VACHON, P.W. and R.K. RANEY (1989), Resolution of ocean wave propagation in single-pass airborne SAR imagery. Proceedings of IGARSS'89 Symposium, Vancouver, July 1989; This issue.

STATISTICAL DESCRIPTIONS OF KEEL-RELATED FEATURES IN THE UNDER-ICE DRAFT DISTRIBUTION

J. Key and A.S. McLaren

Cooperative Institute for Research in Environmental Sciences
University of Colorado, Boulder
Boulder, CO 80309-0449 USA

ABSTRACT

A precise description of the spatial characteristics of the underside of sea ice is important for studying ice dynamics, ice/atmosphere interactions, and acoustic propagation. Previous analyses have examined either thickness distributions, keel spacings, or spectral analysis and have generally been applied to a single ice regime. This study intercompares results from these methods as applied to two different Arctic ice regimes: the central Canada Basin and the North Pole area. Data were collected from narrow-beam sonar carried by a US nuclear submarine. Periodicities in the data are revealed through two-point probability functions of ice draft, as well as keel spacing distributions and spectral analysis. Relationships between ice draft, variability, periodicities, keel drafts, keel spacings, and lead widths are found.

Keywords: Sonar, sea ice, spectral analysis.

INTRODUCTION

To aid in the understanding of ice dynamics and ocean/ice/atmosphere interactions, a precise description of the under-ice draft distribution is required. Analyses to date have generally focused on a single ice regime and have examined either thickness distributions, keel spacings, or power spectra. A summary of these studies is given in Key and McLaren (1989). This study intercompares results from all of these these methods as applied to ice in two different areas of the western Arctic Ocean: the thin ice of the central Canada Basin where large open water areas are observed, and the relatively thicker ice of North Pole area. The purpose of this paper is to examine how these different statistical descriptions change with different ice regimes, and in what way they are related.

DATA

Under-ice draft data were obtained from the Arctic Submarine Labororatory, U.S. Naval Ocean Systems Center, San Diego. In August 1970 the nuclear submarine USS QUEENFISH recorded the under-ice profile using a narrow-beam, high-frequency acoustic profiler with a footprint diameter of 2.68 m. Data were interpolated to 1.45m intervals. The accuracy of the acoustic profiler is ±15 cm at best; the precision of the interpolation routine is estimated to be ±1.0 cm. Gaps occur in the data, varying in length from 4 to 90 km but account for less than 10% of the study transect. These areas were not used in subsequent analyses. The track is divided into regions based on similarity of mean draft and variability, each of approximately 150 km (Figure 1). The entire track length is also divided into 5 km subsections with points averaged over 10 m. Means and standard deviations of these subsections are shown in Figure 2.

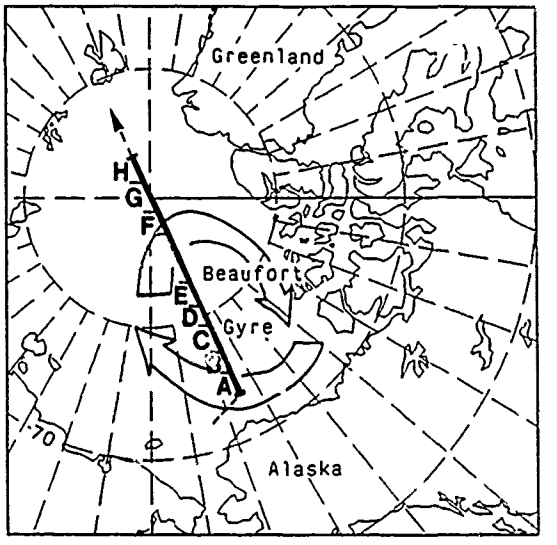


Figure 1. Track of the USS QUEENFISH in early August, 1970 across the Canada Basin and the North Pole. The study area encompasses approximately 1500 km divided into eight regions (A-H). Also shown is the Beaufort Gyre.

METHODS

Analyses include an examination of joint probability distributions of ice draft at various separation distances, numbers, sizes, and spacings of independent keels and leads; and summaries of power spectra by the distribution of spectral exponents and statistically significant periodicities. All statistics are determined for each of the eight regions. Spectral analysis is also applied to each 5 km section and summarized for the regions.

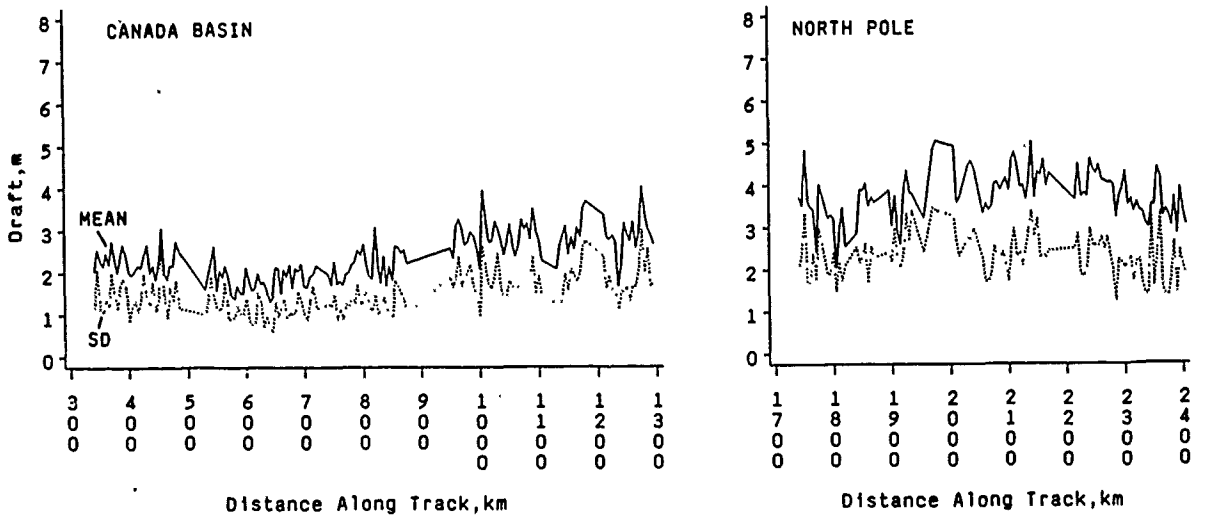


Figure 2. Mean and standard deviation of ice draft for 5 km sections across the central Canada Basin and the North Pole area.

ICE DRAFT STATISTICS

The two-point joint probability density functions (pdf) of ice draft are examined for various separation distances (Colony, 1987) as a method of summarizing the under-ice draft distribution. The estimated two-point joint pdf is

$$f'(h_i,h_j;\delta) = \frac{n_{ij}(\delta)}{N(\delta)\Delta h_i \Delta h_j}, \quad \Delta h_i = h_i - h_{i-1}, \quad \Delta h_j = h_j - h_{j-1}$$

where $f'(h_i,h_j;\delta)$ is the estimated probability of ice draft, h, at two points separated by distance δ being in thickness categories i and j, $n_{ij}(\delta)$ is the number of points at this separation distance which fall into these two categories, and $N(\delta)$ is the total number of observations at this separation distance. The two-point pdfs are determined for each region, over separation distances ranging from 1.5 m to 300 m in increments of 4.5 m. The probabilities along the diagonal of each joint probability matrix are summed for each separation distance as a measure of the relative frequency of occurrence of ice draft at two points falling in the same thickness category. A plot of these summed probabilities by separation distance is shown in Figure 3 for a region in the Canada Basin. The persistence in ice draft is obvious from the large summed probabilities for very small separation distances. It is of interest, however, that the values generally decrease with increasing separation distances, and then increase at a few of the larger separation distances, particulary with separation distances of 130-180 m. This would seem to indicate periodicities in the data, a point which will be discussed later.

KEEL- AND LEAD-RELATED FEATURES

In this section we examine distributions of independent keel and lead spacings, keel drafts, and lead widths. While spectral analysis also provides information on spacings between leads and keels, here we do not assume a periodic nature. The definition of independent keels by Wadhams and Horne (1980) is used here. Leads are defined as open water or recently refrozen areas of no more than one meter draft, with widths of at least 10 m. The original data based on point spacings of

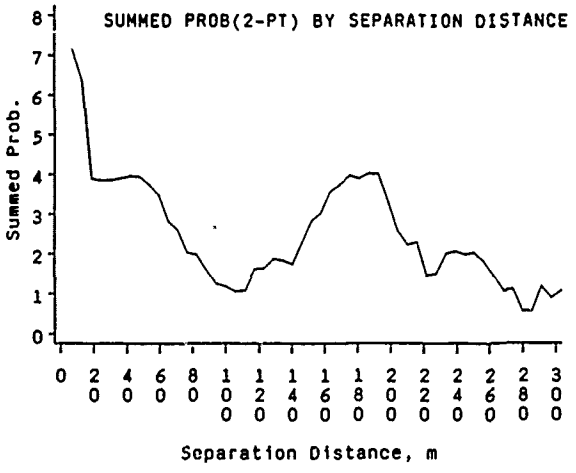


Figure 3. Summed two-point probabilities of ice at various separation distances belong to the same draft category. Probabilities in the two-point joint probability matrix are summed along the diagonal.

1.45 m are used for these analyses.

Two mathematical models of spacing distributions are compared to observed spacing distributions of keels and leads: the negative exponential (Hibler et al., 1972) and the lognormal (Wadhams and Davy, 1986). These two models are tested in each of the regions for independent keels with minimum drafts of 3.5, 5, 7, 9, and 11 m. A summary of the results is shown in Figure 4 for the same region as in Figure 3. Keel spacings are grouped into 20 m bins for keels with drafts exceeding 3.5 m. The theoretical negative exponential and lognormal distributions are also shown. A chi-square goodness-of-fit test is performed under the hypothesis that the observed and theoretical frequency distributions are not significantly different. In all eight regions the observed frequencies are found to be significantly different (0.05 level) from the negative exponential distribution. However, the observed distributions for keels exceeding 5, 7, 9, and 11 m draft are not found to be

significantly different from the lognormal distribution except in the central Canada Basin where only those distributions for keels exceeding 3.5 m draft follow the lognormal model. The more uniform distributions observed in that area are assumed to be a function of the relatively mild ice conditions encountered by QUEENFISH (McLaren, 1987, 1989). Based on this data set, keel spacing distributions cannot be modeled by either probability density function in the central Canada Basin when the threshold value for keel depth is large; e.g. 5 or 9 m for this thin ice area.

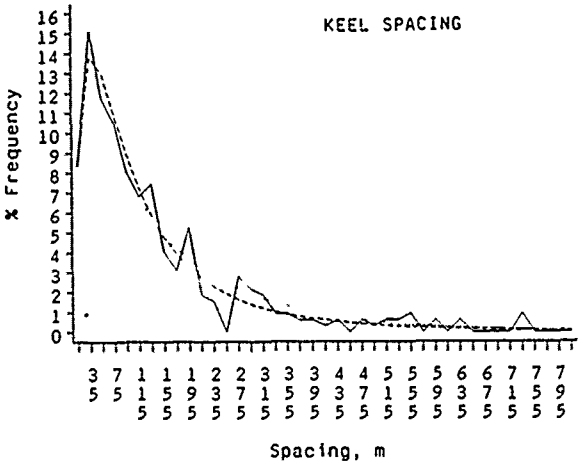


Figure 4. Relative frequency of keel spacings (solid line) in region E with negative exponential (dotted line) and lognormal (dashed line) models also shown. Spacings are grouped into 20 m bins.

SPECTRAL ANALYSIS

Spectral analysis techniques are used to identify cyclical patterns, or periodicities, in data. Spectral analysis theory is well developed; the method applied in this study follows Mitchell et al. (1966). The i^{th} spectral density estimate is a measure of the total variance in the original series that is contributed by wavelengths near the i^{th} harmonic of the fundamental wavelength. The hypothetical population continuum is assumed to be that of "red noise". To inhibit aliasing in the analysis, frequencies above the Nyquist frequency were removed with a low-pass, 1-2-1 filter.

Spectral analysis is applied to each 150 km region and to each 5 km subsection within the regions. An example of the power spectrum for the same region presented in Figures 3 and 4 is given in Figure 5 where significant harmonics are indicated by their wavelength in meters. Figure 6 shows the relative frequency of significant periodicities computed for each 5 km subsection within this same region. Histograms for other regions (not shown) indicate that although spacings along much of the track vary somewhat randomly from one subsection to the next, some patterns are apparent. Periodicities in the range of 30-130 m occur most frequently. The shorter of these may correspond to the "blisters" observed by Wadhams (1988) in sidescan sonar data which had wavelengths of 28-63 m. Also observed is an increase in the frequency wavelengths longer than 1000 m in those areas with greater mean draft and variability, primarily in the northern portion of the track. Similarly, the less variable, thin ice areas of the Canada Basin show an increase in shorter wavelengths with few greater than 1000 m.

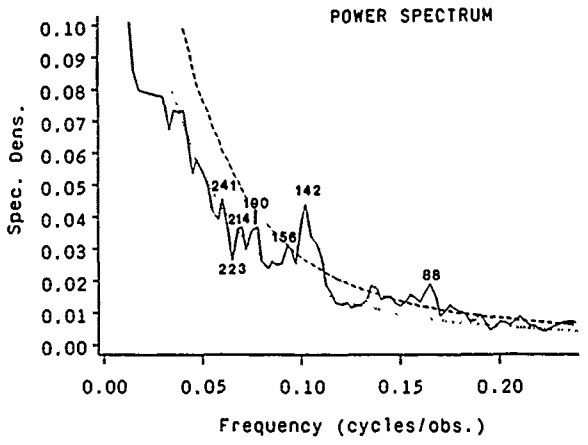


Figure 5. Power spectrum for region E with the hypothetical population continuum and the 95% upper confidence level. Significant peaks are indicated by their wavelengths in meters.

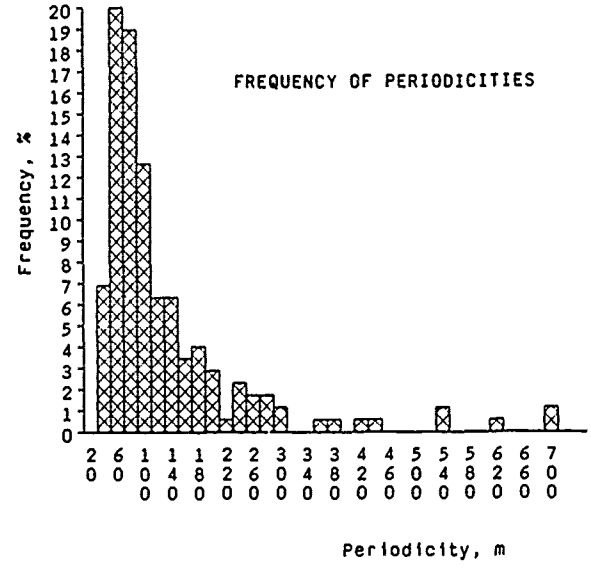


Figure 6. Relative frequency histogram of significant periodicities extracted from power spectra of 5 km subsections within region E.

DISCUSSION

There will undoubtedly be some redundancy in the broad spectrum of variables presented. Benefits in terms of processing as well as interpretability are gained by reducing this set of features to a set which includes only those containing the greatest amount of descriptive information for the analysis of draft distributions. For example, if keel spacings are periodic, we would expect to see an increase in the joint probability distribution at the separation distance corresponding to the wavelength of the periodicity, a peak in the power spectrum, and an increase in the observed frequency of keel and lead spacings also at this point. A comparison of Figures 3, 4, 5, and 6 gives an example of this situation where the two-point probability indicates a periodicity at 140-190 m, the keel spacing distribution shows a peak in the 170-190 m

bin, the power spectrum shows peaks covering the approximate range of 142-214 m, and the summary of significant peaks in the spectra has a peak in the 170-190 m bin (differences of up to 10 m in these figures can be attributed in part to differences in binning procedures). Also note that the low summed probability in Figure 3 near 220 m can also be identified by the other methods.

An examination of the keel spacing distributions and summaries of significant peaks in the power spectra of 5 km subsections reveals further similarities. In particular, there are a larger number of observed keel spacings in the 40-70 m range than expected given a lognormal distribution, even though there is overall agreement between them. This consistent deviation generally agrees with the findings of the spectral analysis where the most frequently occurring periodicities are also in this range.

Table 1 provides summary statistics for each of the eight regions. Keel-related statistics are given by minimum draft. Power spectrum measures were computed for each 5 km subsection within the region then averaged. Variance in the spectra is given for the entire spectrum and for that portion covering wavelengths of 20-150 m. An examination of this table and linear correlations (not shown) indicate that a relatively high degree of redundancy exists in the measures. In particular, most measures are highly correlated with the mean and standard deviation of ice draft.

Table 1
Regional ice draft, keel, and lead statistics.

VARIABLE	REGION A	B	C	D	E	F	G	H
DRAFT (mean) (stand. dev.)	2.3 0.3	1.8 0.3	2.2 0.4	2.9 0.4	2.9 0.5	3.5 0.6	4.1 0.4	3.5 0.4
TOTAL VARIANCE 20-5000m	1.9 1.1	1.0 0.6	1.4 0.6	2.9 1.2	3.0 1.6	4.6 2.4	5.0 2.0	3.7 2.1
TOTAL VARIANCE 20-150m	0.5 0.2	0.3 0.2	0.4 0.2	0.8 0.4	0.8 0.3	1.3 0.8	1.4 0.7	1.0 0.4
KEEL SPACING (3.5m)	201.3 10.0	364.3 28.1	211.6 10.9	136.9 4.6	133.3 5.8	93.3 4.0	82.9 1.8	100.7 2.4
KEEL DRAFT (3.5m)	5.5 0.1	5.0 0.1	5.2 0.1	5.9 0.1	6.0 0.1	6.5 0.1	6.8 0.1	6.1 0.1
# KEELS/KM(3.5m)	4.9	2.8	4.7	7.3	7.5	10.6	12.1	9.9
KEEL SPACING (5m)	455.0 33.5	1042.9 133.7	528.8 40.2	239.4 9.9	233.8 15.4	158.0 10.1	129.7 4.2	187.5 7.6
KEEL DRAFT (5m)	7.2 0.1	6.7 0.2	6.7 0.1	7.3 0.1	7.5 0.1	8.2 0.2	8.4 0.1	7.8 0.1
# KEELS/KM(5m)	2.2	1.0	1.9	4.2	4.3	6.3	7.7	5.3
KEEL SPACING (7m)	1127.5 129.5	3944.9 887.7	1654.0 204.0	560.0 37.8	518.3 55.5	279.3 26.2	221.7 11.7	359.8 23.4
KEEL DRAFT (7m)	9.1 0.2	9.2 0.3	8.7 0.2	9.1 0.1	9.4 0.2	9.9 0.2	10.1 0.1	9.7 0.2
# KEELS/KM(7m)	0.9	0.3	0.6	1.8	1.9	3.6	4.5	2.8
KEEL SPACING (9m)	2554.1 426.4	8671.8 2686.4	5360.5 963.3	1342.5 123.3	1188.5 177.1	528.8 85.2	403.8 29.5	716.4 76.2
KEEL DRAFT (9m)	10.7 0.2	11.2 0.5	10.5 0.3	10.9 0.2	11.2 0.3	11.7 0.3	11.9 0.1	11.5 0.2
# KEELS/KM(9m)	0.4	0.1	0.2	0.7	0.8	1.9	2.5	1.4
KEEL SPACING (11m)	7607.0 1773.7	14596.2 4915.8	20228.8 7560.0	3778.5 600.5	3053.9 960.2	967.4 194.2	708.1 65.8	1555.5 235.4
KEEL DRAFT (11m)	12.7 0.4	11.6 0.3	12.6 0.4	13.0 0.2	13.7 0.4	13.3 0.4	13.3 0.2	13.5 0.4
# KEELS/KM(11m)	0.1	0.1	0.1	0.3	0.3	1.0	1.4	0.6
LEAD SPACING	238.1 14.5	213.4 15.1	226.5 17.9	416.3 31.5	316.9 31.2	421.6 64.6	937.0 97.6	722.0 72.5
LEAD WIDTH	28.9 2.4	40.8 2.5	32.9 2.5	30.9 2.2	31.0 2.9	32.4 6.6	36.4 3.7	40.0 3.6
# LEADS/KM	3.7	3.9	3.9	2.2	2.9	2.2	1.1	1.3

CONCLUSIONS

Ice draft along along the track of the USS QUEENFISH through the Canada Basin and across the North Pole has been described in terms of two-point probability functions, spectral analysis of periodicities, and distributions of keel and lead spacings, drafts or widths. Each of these measures of the under-ice draft distribution were found to vary from region to region. In regions where periodicities exist - in particular regions B, C, E, F, and H - the wavelength of the periodicity is revealed by all of these measures, although each method also provides other unique information. Summary statistics derived from the keel/lead distributions and power spectra can be used to discriminate between ice regimes. The absence of detailed buoy and satellite data for 1970 and requisite oceanographic and ice mechanics information precludes detailed analysis of possible environmental forcings which produced the observed results.

These results imply that analyses of periodicities, independent keels, and simple draft distribution statistics collectively provide a relatively complete description of the under-ice topography and may be useful for discrimination between ice regimes, which in turn will aid in the investigation of environmental forcings. However, correlations between many of these measures indicate that some duplication of information exists and that not all methods are required for an adequate description the under-ice draft distribution.

ACKNOWLEDGEMENTS. This work was supported by the Office of Naval Research University Research Initiative Program contract N00014-86-K-0695.

REFERENCES

1. Hibler, W.D. III and LeSchack, L.A., "Power spectrum analysis of undersea and surface sea-ice profiles", J. of Glaciol., Vol. 11, pp345-356, 1972.
2. Key, J.R., and A.S. McLaren, "Periodicities and keel spacing in the under-ice draft of the Canada Basin recorded by the USS QUEENFISH, August 1970", Cold Regions Sci. Tech., Vol. 16, pp1-10, 1989.
3. McLaren, A.S., "The under-ice thickness distribution of the Arctic Basin as recorded in 1958 and 1970: a comparison", in: POAC '87 Proceedings, 1987
4. McLaren, A.S., "The under-ice thickness distribution of the Arctic Basin as recorded by USS NAUTILUS in 1958 and USS QUEENFISH in 1970", J. Geophys. Res., Vol. 94, No. C4, pp4971-4983, 1989.
5. Mitchell, J.M. Jr., B. Dzerdzeevskii, H. Flohn, W.L. Hofmeyr, H.H. Lamb, K.N. Rao, and C.C. Wallen, "Climatic change", Technical Note No. 79, WMO-No.195.TP.100, World Meteorological Organization, 1966.
6. Rothrock, D.A. and A.S. Thorndike, "Geometric properties of the underside of sea ice", J. Geophys. Res., Vol. 85, No. C7, pp3955-3963, 1980.
7. Wadhams, P., "The underside of Arctic sea ice imaged by sidescan sonar", Letters to Nature, Vol. 333, pp161-164, 1988.
8. Wadhams, P. and T. Davy, "On the spacing and draft distributions for pressure ridge keels", J. Geophys. Res., Vol. 91, No. C9, pp10697-10708, 1986.
9. Wadhams, P. and R.J. Horne, "An analysis of ice profiles obtained by submarine sonar in the Beaufort Sea", J. Glaciol., Vol. 25, No. 93, pp401-424, 1980.

ROLE OF ICE PROPERTIES IN WAVE-ICE INTERACTION DURING LIMEX '1987

W. D. Winsor[1], J. I. Clark[1], B. M. Eid[2], and C. M. Morton[2]

1. Centre for Cold Oceans Resources Engineering, St. John's
2. MacLaren Plansearch Limited, Halifax

ABSTRACT

The Labrador Ice Margin Experiment (LIMEX) was initiated as a pilot program in March, 1987. An interdisciplinary group of different research organizations gathered data from the Marginal Ice Zone (MIZ) and in the open waters adjacent to the ice edge. This paper presents results of the 1987 pilot program.

The Ocean Data Gathering Program (ODGP) spectral ocean wave model was used to hindcast wind and wave fields for the duration of the experiment. In addition, the C-CORE/MUN ice motion package (6 degrees of freedom) was placed on several ice floes to measure the wave induced ice motion at different locations along the MIZ. Analysis of the motion package records furnished time series plots of wave parameters as well as wave spectra within the ice cover. Interpretation of the aerial photographs provided measurements for floe size distribution, floe density and ice concentration.

Keywords: wave ice interaction, sea ice cover, wave attenuation, wave patterns in SAR imagery.

INTRODUCTION

The Labrador Ice Margin EXperiment (LIMEX) is directed to advance the knowledge of the natural processes operating during the formation, existence and decay of the sea ice cover off the east coast Canada. The first field program for LIMEX was carried out in March 1987. It was conducted as a pilot project to gather field observations and field experience of the sea ice cover at the ice transition to ocean, "the marginal ice zone", of the northwest Atlantic. This is a complex and dynamic zone is one of considerable economic importance for renewable and non-renewable resources in Canada.

The LIMEX '87 pilot project was encompasses careful documentation of observations, evaluation of measurement techniques and the formulation of hypotheses. The chief interest of the authors was to observe and quantify wave ice interaction. It is postulated that the ocean waves incident at the ice edge are crucial in the genesis of the ice conditions encountered at the ice margin. On the other hand, the ice cover plays an important role in determining the nature of the ocean waves generated at the ice edge and the character of waves propagated within the ice cover.

The winter ice cover at lower latitudes in the northwest Atlantic is an extremely variable phenomenon. This was illustrated dramatically through the LIMEX '87 project time period. The ice cover which initially stretched 400 km east from the Newfoundland coast. During the 9 days from March 9 to 23, however, this ice cover was compressed by persistent easterly winds into a narrow strip adjacent to the coastline through due to (McNutt *et al.*, 1988). This stable weather condition provided a rare opportunity to observe the effects of wave action on an ice cover with a well defined edge.

The marginal ice zone is a difficult place to conduct oceanographic and ice cover measurements. Wave buoy moorings cannot be set because of the high risk of having equipment swept away by the moving ice. Often the ice is too thin and motions too severe to set and recover instrumentation from the ice floes.

The ice motion instrumentation employed during LIMEX '87 was designed to measure the wave induced motions of bergy bits and growlers. The equipment was not ideally suited to collect the extended time series desired for wave investigation.

The LIMEX program included a major remote sensing component. This focussed on the evaluation of a prototype synthetic aperture radar (SAR), the sensor unit which is being developed for the RADARSAT platform. The wide coverage of the SAR imagery provides a exceptional opportunity for fitting local observations into a regional picture.

MATERIAL PROPERTIES OF FLOE ICE COVER

The incident ocean waves furnish the driving force while the ice cover offers resistance to the transfer of this wave energy through the ice cover. When a radiating wave field encounters a material boundary, a fraction of the wave energy passes (is transmitted) into the ice cover and a fraction is reflected. The ratio of energy transmitted and reflected depends on the equivalent of the index of refraction (optics) of wave propagation in an ice cover. There is little known about the physical properties which dictate wave propagation and wave reflection in discrete floe ice cover.

Wadhams *et al.* (1986) suggest floe size distribution to be an important property controlling wave attenuation in ice cover. Their study was conducted in a pack ice regime composed of large floes (~100 m). The LIMEX'87 work was conducted in ice cover consisting of much smaller mean floe size, (~5 m). Here other ice cover features appeared to have an important effect on the wave ice interaction. LIMEX '87 observations of wave ice interaction suggest that collisions at floe boundaries and within the ice debris filling the interstitial area between discrete floes have an important influence on the ice cover's response to wave propagation.

Lateral movement of the ice floes during convergence produces in two significant changes in the nature of the ice cover. (1) Any pulverized ice is squeezed and forced down to underlie the discrete floes and (2) the floes are rotated and jostled so that the interstitial area between the floes is minimized. Hence the floes interlock and offer increasing resistance to further lateral movement and to bending motion from wave action. The compaction of the pack is a time varying property and depends on prior mechanical activity, i.e. the mechanical history for the ice cover. The increase of ice cover strength (resistance) caused by convergence is commonly know as "ice pressure". It offers considerable resistance to vessel movement and can even beset capable ice breaking vessels. The "pressure" condition will persist as a feature of the pack ice condition after the lateral load which created it has passed. Further exposure to wind and current loads, new wave action or a combination of both will act to disperse the ice and relieve the ice pressure.

During LIMEX '87 the available measurement techniques were not specifically designed to quantify the material properties which could resist or enhance wave propagation of the ice cover. More basic research needs to be done to yield fundamental results for wave action in sea ice cover. The LIMEX '89 field program will spend additional time on this topic. Any advances in the identification of "pressured ice" especially remote sensing identification by the ice reconnaissance services will produce significant economic benefits to companies routing vessel traffic through ice covered waters.

FIELD OBSERVATIONS

Ice Motion Measurement

Ice floe motion measurement was accomplished with the C-CORE/MUN six-degree of freedom ice motion package (Lever *et al.*, 1986). This unit was originally designed to measure wave induced iceberg and growler motions and was employed without modification. The sensor cluster uses three orthogonally mounted accelerometers to detect displacement. The two instruments used in 1987 had different methods of measuring pitch and roll; package B used a gyro and package C had a set of tiltmeters. The angular measurements are required to transform the local instrument axis to a fixed reference frame. There is still some question about the validity of tiltmeter measured rotations when these sensors are faced with a combination of rotation and acceleration.

An investigation of the directional wave spectra and a more complete comparison of ice motion with the incident ocean wave regime is given a fuller treatment to a separate paper by Eid *et al.* (1989).

Incident Wave Field

Wave measurement at the ice edge was carried out with a Delft disposable wave buoy, on March 22 and 23, 1987 (Thomas, 1988). The wave field was hindcast using the meteorological data and the Ocean Data Gathering Program (ODGP) model for all the observation periods.

Remote Sensing

The C Band SAR imagery collected by the Canadian Centre for Remote Sensing's Convair 580 aircraft provided extensive coverage of the marginal ice zone through LIMEX '87. Coverage on March 23 and 26, 1987 was of particular interest to the wave ice interaction program. The SAR imagery covered from 3 to 3.5 degrees of latitude and full coverage from the Newfoundland coast to the ice edge to the east. The SAR imagery provided a unique opportunity to view the wave ice interaction process on a regional scale.

Floesize Distribution

Aerial photographs were taken from a helicopter at 600 m to access the granular nature of the ice. A 70 mm Hasselblad camera with a 80 mm focal length lens was used with Kodak Aerographic Plus-X black and white film. Ice coverage photography consisted of a series of 60 to 100 square frames of 400 m on the side except for one day when the ceiling was restricted to 300 m and the photograph frames were 200 m on the side.

Floe sizes were taken from 73 frames for March 23, 25 and 26, 1987. Each photograph was sampled by measuring floe diameters for every floe occurring on a set of lines that was superimposed on the photograph. The floe dimensions were obtained using a vector digitizing table. This sampling technique yields a floe size distribution that is normalized by floe area. Figure 1. illustrates a typical floe size distribution. Table I. summarizes the results for some locations of interest to the wave ice interaction investigation.

Case I - Dispersed Ice Floes

The first deployment of the ice motion package took place on March 22, 1987. The instrument was placed on a small floe, (~5 m), in a strip of loosely packed floes, < 10 m. All the floes were surrounded by water; adjacent floes were not colliding. Figure 2. shows the non-directional wave field estimated from the ice motion data collected. Simultaneous measurements by the John Hopkins University waverider in adjacent open water (Thomas, 1988) gives the wave height entering the pack at 2.3 m and a dominant period of 9 seconds. The "CSS Baffin" was operating in limited visibility when the measurements were conducted. There was no collaborative helicopter photography of SAR imagery on March 22, 1987.

Case II - Definite Transition

March 23, 1987 proved to be a significant observation period for wave ice interaction. It was unfortunate that the ice motion package failed to record data during that deployment. The ancillary observations are valuable for the study and are reported.

The swell in the ice cover was evident visually and has been reported by McNutt *et al.* (1988). The

SAR imagery of the ice cover of the "CSS Baffin's" operations area showed a marked transition about 4 km into the ice cover. The transition was from a homogeneous, wave patterned texture seaward, to a floe patterned texture for the ice cover landward to the coast. The decreasing wave pattern observed in the ice cover with homogeneous radar return suggested the wave propagation into the ice cover was significantly attenuated by the time the waves had travelled to the transition zone. The floe size distributions for the transect across the ice belt showed a marked reduction of floe size immediately adjacent to the ice edge. Waverider measurements on March 23, 1987 showed the incident wave field having a 2.5 m significant wave height.

Case III - Wave Attenuation

A seven hour period was dedicated to the ice motion measurement on March 25, 1987. Sensor package C was placed on a small ice floe, (~3 m), adjacent to the ice edge. Its motion was recorded for 5.5 hours. Package B was deployed three time at three different locations, nominally at 1, 2, and 4 km into the ice pack. Figures 3. (a), (b), and (c) give the non-directional spectral densities for each deployment. For comparison Figure 4., shows the waves attenuating and the shorter period waves disappearing with distance away from the ice edge. The energy attenuation is summarized as: -3.9 dB with dominant period of 8.7 seconds at 1 km, -8.3 dB and period 9.5 seconds at 2 km and -11.4 dB and period 18.1 seconds at 4 km.

The helicopter aerial photography showed the floe size distribution to be relatively uniform over the ice cover attenuating the wave field, (see Table I). The floes were relatively closely packed. The marked attenuation of the waves over a relatively short distances indicated the ice cover energy absorption mechanisms were effective through the observation period.

Case IV - Wave Penetration

Ice motion measurements, helicopter aerial photography and SAR coverage were obtained for March 26, 1987. The motions measured at the two stations, one is given in Figure 5. The non-directional spectra showed a long period swell, 15 seconds, propagating in the ice cover with little attenuation, -1.7 dB, between the stations. This observation was collaborated by visual observation of the swell running in the ice cover and the SAR imagery collected for the same time period. The SAR imagery showed a uniform wave pattern for the whole ice cover along the east coast of the Avalon. The apparent wavelength for the swell was estimated at 300 m.

The more complete penetration by the ocean waves through the ice cover during this period of observation might be attributed to two factors. (1) The longer period wave incident at the ice edge, 15 seconds as compared to 8 to 9 seconds on March 25. (2) The ice cover appeared to have more separation between the floes then the ice cover sampled on March 25. Techniques need to be developed further to characterize the granular nature of pack ice cover.

DISCUSSION

The observations of the LIMEX '87 project showed that the wave ice interaction is a complex phenomenon and the marginal ice zone cover is a complicated material. The discrete floe composition of sea ice cover is analogous to a granular material (Erlingsson, 1988). It seems to behave like a two dimensional granular material with particle sizes large compared to thickness. The third dimension is the ice thickness which is also an important factor in assessing ice cover strength. The average strength of the particular ice cover dictates how the incident wave field crosses the ice edge and propagates within the ice cover.

Wind and current act on the air/ice and ice/water boundaries causing the floes to converge (compact) and disperse; this depends on the direction of the applied loads and boundary restraints on the ice cover. Ice convergence strengthens the ice cover and increases its resistance to the radiation of the wave action. The SAR imagery of March 23, 1987 showed an ice field with considerable lateral strength. The major portion of the wave energy appeared to be attenuated in a narrow band of ice adjacent to the ice margin.

Subsequent motion measurements and SAR imagery, on March 26, 1987, showed the wave field propagating further into the ice cover. Results indicated that the ice cover strength changed with time. The associated surface measurements suggested the decrease in strength was caused by a combination of reduced wind stress and the gradual wave action on the floes. The accumulating volume of crushed ice between the floes suggested the relative movement between floes caused continual failure at the floe perimeter. Jordaan and Timco (1988) have investigated the role of crushed ice during the ice-crushing process and have shown that this viscous ice material consumes the largest part of the failure energy involved.

The crushed ice appears to perform a somewhat different function in wave ice interaction. The crushed ice flows into the interstitial area between the floes lowering the flexural strength of the granular ice cover. The weakened cover permits the wave field to radiate deeper into the ice cover where the process continues. While wave energy attenuation can be approximated by linear potential theory (Wadhams *et al.*, 1986, 1988), the attenuation coefficient is dependent on wavelength of the incident wave field (Eid *et al.*, 1989) and on the material properties of the ice cover involved. Therefore the attenuation coefficient can be expected to vary rapidly as the ice cover converges and disperses, and as the ice crushed and ground by wave bending and floe to floe collisions. Although thermodynamic effect of re-freezing was not encountered through the LIMEX '87 period it must be assumed to have a significant effect on the behaviour of the winter marginal ice zone.

SUMMARY

The results collected during the field experiment do not conclusively confirm the relative importance of floe size distribution in the attenuation of ocean waves in the marginal ice zone. The authors speculate that the wave penetration is a major factor in dictating the floe size distribution. Further work is planned on the mechanical behaviour of ice floes in a wave environment. It is clear that the ice cover is effective in filtering the wind waves. The spectra of the ice motion time series indicates the ice cover encountered during LIMEX '87 behaved like a low-pass filter with a roll off at about 0.2 Hz.

Other features like floe to floe separation (compactness), frequency of floe collisions, floe to

floe interlock and fraction of ground ice appear to have a more direct effect on wave ice interaction. The interactions are a function of the spatial distribution of the ice cover. It is not yet clear how to assign quantities that will be effectively correlated with the mechanical relationship (model) for wave ice interaction. The helicopter aerial photographs and the SAR imagery are two data sets that might be applied to extract this sort of information. Both techniques offer limited temporal coverage. It is planned to investigate these concepts with image capture and analysis techniques to explore ways to quantify the granular nature of sea ice cover.

Much more complete data collection will be required before it will be possible to apply these results to help estimate of risk of ice loads at offshore structures. Future results may prove that offshore platform may not need to be designed to the break continuous ice cover. The combination of wave and ice loads may represent a lower environmental load than either wave or ice loads acting independently. The objective of LIMEX '89 is to collect a more extensive data set under a different set of conditions, add to the data base of ice conditions and aid the engineering of safe and economic systems to recover both renewable and non-renewable resources from Canada's ice covered oceans.

ACKNOWLEDGEMENTS

This study resulted from a combination of Unsolicited Proposal funded by Supply and Services Canada (DSS UP Funds), a NSERC CRD Grant in conjunction with Mobil Oil Canada Ltd., and contributions from DFO (MEDS and BIO), EMR (CCRS) and PERD. The contribution from these sources are acknowledge with gratitude. The authors are also grateful to Dr. Ron Wilson of MEDS, the Scientific Authority for this study, and Mr. Jim Ransom, of Mobil Oil Canada Ltd, the industry liaison officer of the NSERC CRD Grant, for this support and encouragement.

REFERENCES

Eid, B. M., Morton, C.M., Winsor, W.D. and Cardone, V.J., "Wave-ice interaction study during LIMEX '87", Proc. of the 2nd Inter. Workshop on Waves, Vancouver, 11 p., April 1989.

Erlingsson, Bjorn, "Two-dimensional deformation patterns in sea ice", J. of Glaciology, Vol. 34, No. 118, pp. 301-308, 1988.

Jordaan, I.J., and Timco, G.W., "Dynamics of the ice-crushing process", J. of Glaciology, Vol. 34, No. 118, pp. 318-326, 1988.

Lever, J.H., Klein, K., and Diemand, D., "Design, development and deployment of instrumentation packages to monitor the motion of glacial ice masses at sea", Ice Technology, T.K.S. Murthy, J.J. Connor, C.A. Brebbia, eds. Proc. of the 1st Inter. Conf., Cambridge, MA., pp. 225-236, June 1986.

McNutt, L., Argus, S., Carsey, F., Holt, B., Crawford, J., Tang, C., Gray, A.L. and Livingstone, C., "The Labrador Ice Margin Experiment, March 1987 - A pilot experiment in anticipation of RADARSAT and ERS 1 data", Trans., Amer. Geophysical Union, Vol. 69, No. 23, pp. 634-635, 643, 1988.

Thomas III, W.L., "Labrador sea wave and ice measurements in support of the March Labrador Sea Ice Margin Experiment (LIMEX)", David Taylor Research Centre, Rprt. No. DTRC/SHD-1212-05, 29 p., 1988.

Wadhams P., Squire, V., Ewing, J.A. and Pascal, R.W. "The effect of the marginal ice zone on the directional wave spectrum of the ocean", J. of Physical Oceanography, Vol 16, pp. 358-376, 1986.

Wadhams P., Squire, V., Goodman, D., Cowan, A. and Moore, S. "The attenuation rates of the ocean waves in the marginal ice zone", J. Geophysical Res., Vol. 93, No. C6, pp. 6799-6818, 1988.

TABLE I

SUMMARY - FLOE SIZE DISTRIBUTIONS

location	mean floe size (m)	std. dev. (m)
March 23, 1987		
ice edge	4.3	2.1
mid-point	4.8	3.2
near coast	9.9	7.3
March 25, 1987		
ice edge	9.8	6.7
2km from edge	7.3	5.3
4km form edge	10.9	7.3
March 26, 1987		
across the pack		
ice edge	5.4	3.4
mid-point	7.7	4.7
at coast	18.0	11.1
along coast		
north end	6.7	3.6
at ship	7.0	3.7
south end	7.1	3.7

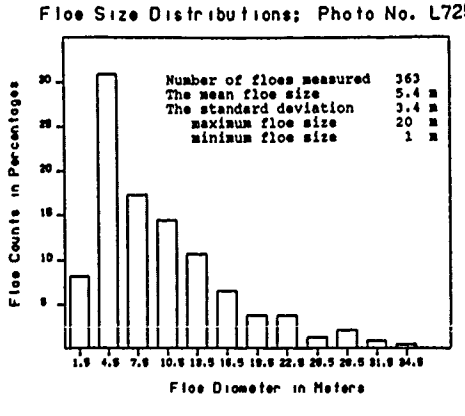


Figure 1. A typical floe size distribution for the ice cover encountered in LIMEX '87.

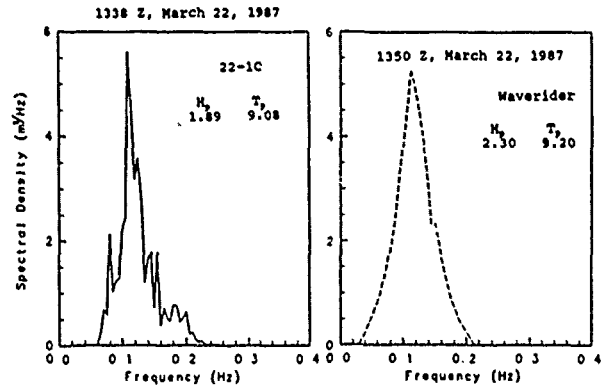


Figure 2. The non-directional spectra density for ice floe motion and waverider on March 22.

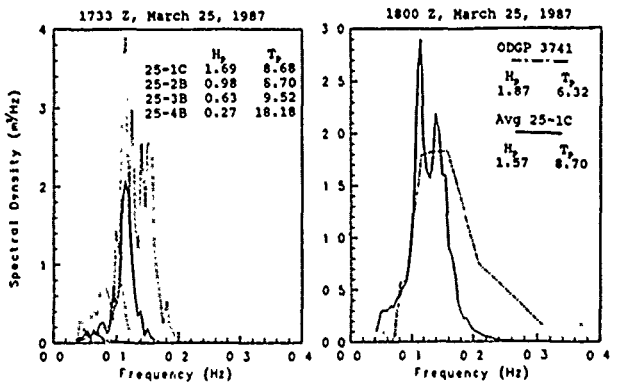


Figure 4. Comparing the spectra of Figure 3 with the ice motion at the ice edge and concurrent ODGP wave hindcast.

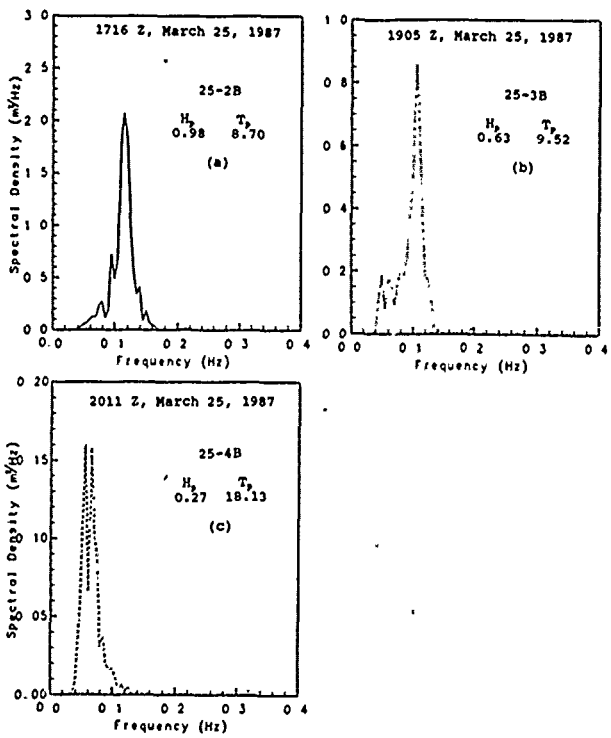


Figure 3. The wave spectra densities showing attenuation with distance, (a) 1 km, (b) 2 km, and (c) 4 km.

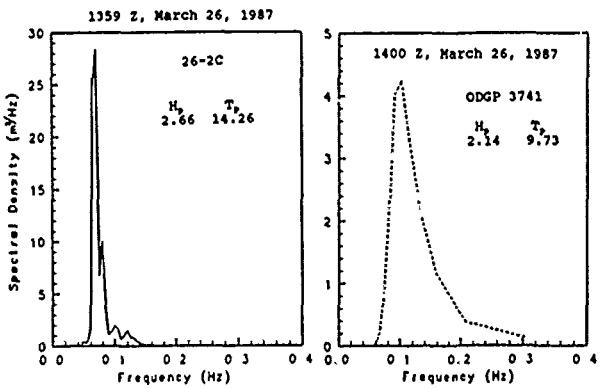


Figure 5. The ice motion and wave hindcast on March 26 with the long period swell in the ice cover.

IDENTIFICATION AND VOLUME ESTIMATION OF ICEBERGS BY REMOTE SENSING IN THE BARENTS SEA

Sylvi Vefsnmo[1], Stig Magnar Løvås[1], Sveinung Løset[1], Terje Næss[2]

[1] Norwegian Hydrotechnical Laboratory, SINTEF
7034 Trondheim-NTH,
Telephone: 47-7-592300
Telefax: 47-7-943345
[2] Fjellanger Widerøe A/S

ABSTRACT

Offshore activities like petroleum exploration, fishing and hunting in the Barents Sea are heavily affected by. the ice conditions. In the planning stage of activities the major challenge is to make precise estimates of impact forces by ice colliding with a structure/vessel. Surveys in the area have revealed such a high population of icebergs that those represent the design criteria for offshore structures/vessels rather than the harmless first year ice and more seldom encountered multi-year ice.

To estimate the probabilities of collision and calculation of forces caused by impact with ice, one needs to know the intrusion of icebergs and ice floes. Research surveys in the ice covered waters are expensive and hence for practical reasons give limited coverage in time. Satellite imagery has the best time coverage but the disadvantages are the necessity of cloud-free atmosphere as well as limited resolution and interpretation.

To improve time and aerial coverage, remote sensing might be a valuable tool in this environmental mapping. This paper describes how Landsat-TM imagery and additional information are used to identify and map icebergs in the Barents Sea during spring 1988. When the icebergs are identified, the next step is to estimate interesting size parameters for icebergs, especially the volume.

Volume estimates of icebergs can be performed in several ways with varying results in regard to the consumption and accuracy. Three methods are used: Photogrammetric analysis, simplified visual analysis using stereoscope and a formula involving a shape factor based on iceberg classification. All these methods are primarily used in combination with stereo aerial photography. The paper describes these methods and evaluates the results.

INTRODUCTION

Oil resources and discoveries in conventional offshore source areas are becoming fewer and smaller. Hence the industry is turning its attention towards more remote areas, often with a harsh environment.

At present, much interest is focussed on the Barents Sea, where exploration activity has recently increased, and might accelerate with additional licensing in the future.

It is costly to design and use offshore exploration drilling platforms which can withstand the impact of icebergs. Hence the oil industry is inclined to apply less costly concepts involving ice contingency planning and ice management with possible evasive action. In view of the significant cost of interrupting drilling operations by disconnecting and re-entering the well, some method is required to determine whether or not a particular iceberg represents a significant threat to the platform.

To estimate the probabilities of collision and calculation of forces caused by impact with ice, one needs to know the intrusion of icebergs and ice floes. The driving force of the Barents Sea dynamics is the global radiation imbalance, with a surplus of incoming heat in the equatorial regions and a deficit of outgoing radiation in the polar regions (Carstens et al., 1988). The ocean polar front is located further north than the better known atmospheric ones, cutting right across the Barents Sea. Research surveys in the ice covered waters are expensive and hence give limited coverage in time. Satellite imagery has the best time coverage but the disadvantages are the necessity of cloud-free atmosphere as well as limited resolution and interpretation.

The NOAA AVHRR system, the Landsat system and the Spot system are all in the visual, near-IR and/or thermal-IR spectral region and they therefore require sunlight (visual, near-IR) and a reasonably cloud-free atmosphere. Light haze can be tolerated, but denser fog or clouds will make imagery impossible.

The most detailed (and most expensive) images today are Spot images, covering 60 x 60 km^2 with geometrical resolution either 10 m or 20 m, and a repeat cycle of 26 days. Landsat images give a better areal coverage (full scene: 185 x 185 km^2) and have a shorter repeat cycle (16 days). Due to overlap, the effective repeat cycle is even less. Landsat images are also less expensive. The best resolution of the Landsat - TM images is 30 m.

IDENTIFICATION OF ICEBERGS IN LANDSAT IMAGES

In several Landsat images we could clearly see numerous ellipse-shaped traces as a result of small ice floes moving past grounded icebergs. These ellipse-shaped traces are caused by tidal current and give valuable additional information in order to detect icebergs in Landsat images. The discovery of these traces increases the chance of using Landsat images to

extend the knowledge of the changing populations of icebergs in the shallower parts of the Barents Sea.

A Landsat TM image from June 1st 1988 centered on 75°50'N, 22°50'E (southwest of Hopen) has been analyzed and is shown in Figs. 1 and 2. The image covers an area of 90 x 90 km. This paper presents results from the iceberg identification analysis. The position of the icebergs is estimated based on the known positions of the island Hopen in the adjacent northern image. Totally 50 icebergs were identified and they are plotted in Fig. 3. Due to the bathymetry in the area and the ellipse-shaped traces, most of the 50 icebergs are grounded. The depths in the area vary from 35 - 100 m.

Fig. 1. LANDSAT quarter scene southwest of Hopen

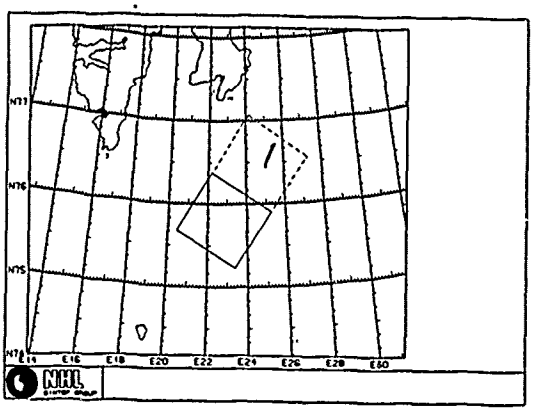


Fig. 2. Positions of the LANDSAT TM image

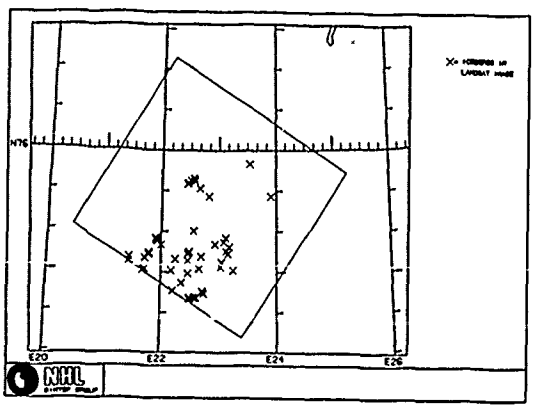


Fig. 3. Positions of identified icebergs in the LANDSAT TM image June 1st 1988

In 1988 Norwegian Hydrotechnical Laboratory (NHL) and Norwegian Polar Research Institute (NP) conducted the Ice Data Acquisition Program, Vessel Deployment (IDAP '88). The intention of the expedition was to obtain data on iceberg and sea ice movement through the deployment of Argos buoys on selected icebergs and sea ice floes. The ice conditions in the Barents Sea between 74-77°N and 19-28°E were investigated on a 10 days survey with R/V "LANCE" late March 1988. About 200 icebergs were detected on this survey. The highest concentration of icebergs was located at the eastern slope of Spitsbergenbanken. About 50 to 60 of these icebergs were grounded.

During this survey 10 Argos buoys were deployed on icebergs. In the aftermath the icebergs have been tracked. On June 1st 1988 two of these buoys were within the area covered by the Landsat image. These two icebergs are also identified in the image.

During a survey with K/V "ANDENES" in late May 1988, a total of 56 icebergs were observed from the vessel. Fig. 4 shows the icebergs on this survey close to the area covered by the Landsat image. A comparison between the observed icebergs in Fig. 4 and the detected icebergs in Fig. 3 shows that a majority of the icebergs are still in the area.

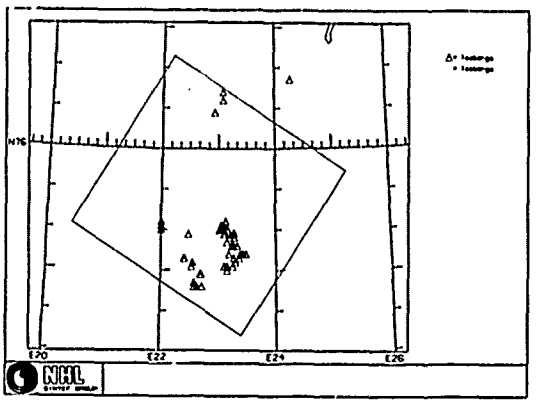


Fig. 4. Icebergs observed from K/V "ANDENES" in late May 1988

VOLUME ESTIMATION

In the following 3 methods for volume estimation of icebergs will be described:

* stereo photogrammetric analysis
* simplified stereoscopic analysis
* shape factor method

Stereo photogrammetric analysis

Between the 22nd and 27th of March 1988 a number of icebergs/bergy bits were photographed from the air along predetermined tracks in the Barents Sea, southeast and east of Spitsbergen. All significant icebergs/bergy bits discovered were stereo-photographed by a precision mapping camera from 3000 feet altitude with the sun elevation between 12° and 15°.

During IDAP'88 some 135 icebergs/bergy bits were photographed, and 102 of them have been analyzed using the overlapping photographs. These icebergs have been fitted into an analytic stereoplotter to create a stereographic model of each iceberg. This model is then used to draw a topographic map of each iceberg based on regular photogrammetric principles.

The stereo-operator constructs the iceberg maps with 1 m interval between contour lines and digitizes the contour lines into a data base at the same time. The resulting digital terrain model is then used to calculate different parameters of each iceberg. These parameters are:

* maximum height
* maximum waterline length
* maximum waterline width
* waterline plane area
* above waterline volume

The volume calculations are performed using the plane area of each height contour. Nominal accuracy of height measurements in stereomodels with 60% overlap is 0.015% of the aircraft altitude, i.e. ± 0.14 m at 3000 feet. Overall accuracy of the volume calculations of the icebergs is about ± 4.6%. This method provides data with high accuracy, but is also somewhat time-consuming (about 2.5 - 3 hours per iceberg).

Simplified stereoscopic analysis

The equipment needed for using this technique is a small stereoscope and a ruler. With this equipment the maximum and mean length are measured as well as width (perpendicular to maximum length) and the sun shadow length on the level ice. The maximum values are rather easy to measure, while the mean size parameters become a subjective estimate.

When using α as the known sun elevation, the height parameters are calculated as:

$$H = S \cdot \tan \alpha + H_0, \qquad H_0 \text{ - sea ice freeboard}, \quad S \text{ - sun shadow length}$$

These calculations and scale transformations are performed with computer software. The mean values are used to estimate area and volume, as follows:

$$V = A \cdot H_{mean}, \quad A = L_{mean} \cdot W_{mean}$$

During the ICEBASE (Sea Ice Investigations in the Barents Sea) programme both stereo photogrammetric analysis and simplified stereoscopic analysis were performed for 9 icebergs (Næss et al., 1988). The objective of this intercomparison was to test the difference in accuracy. The deviations (per cent) for the various parameters measured are tabulated in Table 1.

Table 1. The deviations (per cent) in the parameter estimates between stereo photogrammetric analysis and simplified stereoscopic analysis for 9 icebergs.

	Volume	Area	L_{max}	H_{max}
Deviation (%)	1.3	2.7	4.2	3.2
	1.0	2.7	1.5	13.3
	0.28	0.61	0.20	0
	2.2	1.1	0.50	0.54
	12.0	16.0	0.35	2.2
	15.8	20.0	3.4	3.9
	20.7	33.0	21.7	5.3
	3.2	23 0	2.7	0.92
	1.9	15.>	4.9	0
Mean deviation (%)	6.5	12.8	4.4	3.3
Standard deviation	7.6	11.6	6.7	4.2

The largest deviations are found for the area measurements. The reason for this is mainly the discrepancies in drawing the line between the iceberg and the surrounding ice. This line is sometimes difficult to define. These discrepancies will have minor effect on the volume estimates. The deviations are smallest for tabular icebergs since the mean parameters are easier to estimate for this category of icebergs. This is a very fast method for estimating the size parameters (less than 10 minutes per iceberg).

NHL has developed software to calculate maximum length, width, perimeter and area based on digitized coordinates. A stereoscope is used to estimate maximum and mean sun shadow, and to outline the iceberg water level area on a transparency. Then the outlined area is digitized, and the digitized coordi-nates and the shadow lengths are input to the software. The output of the software are more accurate estimates of all parameters except the height parameters. Especially for highly irregularly shaped water level areas, the increase of accuracy for the area and volume estimates is significant. This increase in accuracy requires more labour time per iceberg than the above described method (about 20-30 minutes).

Shape factor method

During the ICEBASE programme and IDAP'88 a total of 206 icebergs were analyzed. The size parameters of these icebergs were estimated by using stereo photogrammetric analysis and simplified stereoscopic analysis (= 70 % by stereo photogrammetric analysis). Based on this analysis a shape factor (f) is calculated and defined as: $f = V/H_{max}\, L_{max}\, W_{max}$. The purpose was to estimate the reliability of the volumes estimated by using this shape factor method. The frequency of occurrence distribution for the shape factor is shown in Fig. 5. The mean value of the shape factor is 0.35 and the standard deviation is about 0.12 (or 34 %). As an attempt to increase the accuracy, the frequency of occurrence for the shape factor is determined based on classification of the icebergs.

The icebergs were classified into four categories: pinnacled, tabular, sloping and weathered. Table 2 shows the mean values and the standard deviations of the shape factor for all these iceberg categories.

Table 2. Statistical data for the shape factor (f).

Iceberg category	Mean value of shape factor	Standard deviation	Standard deviation (%)
Pinnacled	0.30	0.08	25.2
Tabular	0.53	0.11	21.2
Sloping	0.34	0.09	25.1
Weathered	0.36	0.09	25.0
All	0.35	0.12	33.7

The figures in Table 2 clearly show an increase in accuracy, but still the deviations are quite high. The mean values of the shape factor for the pinnacled, sloping and weathered icebergs are about the same. When analysing these 3 categories as one, called non-tabular, the mean value of the shape factor becomes 0.32, while the standard deviation increases slightly (26.1 %). The frequency of occurrence distributions are shown in Fig. 6 and Fig. 7.

The analyses of the shape also searched for any connection between the shape factor and the size of the parameters. Scatter plots were made of the shape factor versus the following parameters:

* height (H)
* length (L)
* width (W)
* height divided by length (H/L)
* height divided by width (H/W)
* width divided by length (W/L)
* height divided by both width and length (H/W·L)

An example of these scatter plots is shown in Fig. 8.

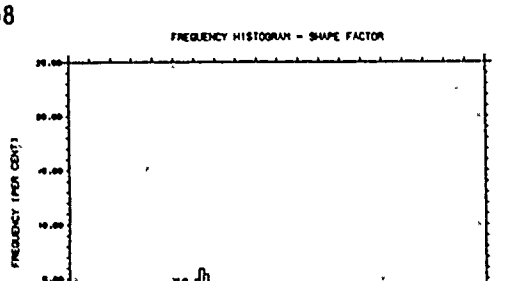


Fig. 5. Frequency of occurrence of the shape factor for all icebergs

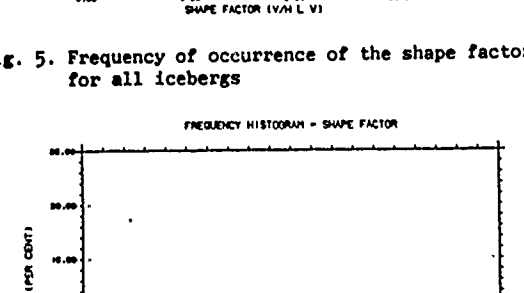


Fig. 6. Frequency of occurrence of the shape factor for tabular icebergs

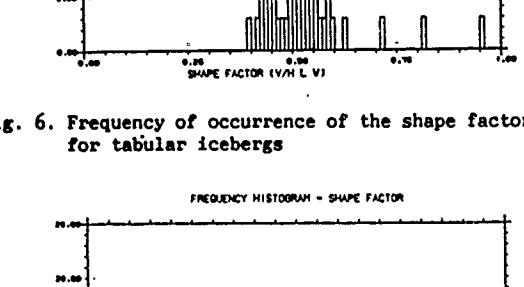


Fig. 7. Frequency of occurrence of the shape factor for non-tabular icebergs

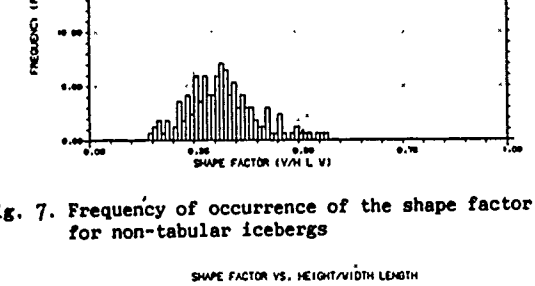


Fig. 8. Scatter plot of the shape factor (f) versus height/(width·length) for pinnacled icebergs

Based on these scatter plots, there seems to be no unique connection between the shape factor of an iceberg and a combination of the size parameters.

CONCLUSIONS

Recent research justifies the assumption that the icebergs rather than the sea ice represent the design load in the Barents Sea. The indentation pressure is a function of contact area between iceberg - structure (Sanderson, 1986) and the shape of the iceberg does also affect the sharing between overturning moment and translation forces put on the structure, e.g. if the iceberg hits the caisson instead of a column.

As shown by the standard deviation of the shape factors (Table 2), there is a significant decrease in accuracy when using the shape factor method instead of the two other methods. The experiences so far clearly show that the information provided by the 3-dimensional model and the stereo photogrammetric analysis is of high value and justifies its cost.

If a customer wants a fast estimate of a limited set of size parameters, the method of using simplified stereoscopic analysis might very well be adequate. Because of the poor accuracy, the method of using a shape factor is limited for research purposes. But the method is a helpful tool to get a low cost and fast estimate of the volume of an iceberg.

When performing ice reconnaissance in the vicinity of drilling platforms it is important to sort out possible hazardous icebergs from the more harmless ones (icebergs which can be handled with ice management techniques). By using the shape factor (f) and a load factor (γ) the volumes are estimated as: $V = \gamma \cdot f \cdot H_{max} \cdot L_{max} \cdot W_{max}$. When using i.e. $\gamma = 1.6$ one will get maximum volume estimates, which should be adequate to separate harmless icebergs from those who need better volume estimates. When using a load factor less than 1.0 (i.e. 0.8) one will get a minimum estimate of the volumes, and if this minimum volume estimate is bigger than the design criteria for the ice management technique or the drilling platform itself one will have to prepare for evacuation or taking into account other countermeasures like active ice management.

ACKNOWLEDGEMENT

IDAP'88 (Ice Data Acquisition and Analysis Project) was funded by the Operators North of 62°N (OKN), and the authors gratefully acknowledge for permission to release information from this project. The research programme ICEBASE (Sea Ice Investigations in the Barents Sea) was funded by Mobil Exploration Norway Inc., BP Petroleum Development (Norway) Ltd and Esso Norge a.s.

REFERENCES

Carstens, T., Løset, S. and Løvås, S.M., 1988, "Ice Drift Modelling for the Barents Sea", Polartech'88, Trondheim

Darell, L., Jensen, H. and Løvås, S.M., 1988, "IDAP'88 Vessel Deployment, Volume 3: Cruise Report K/V "ANDENES", NHL report STF60 F 88091

Hisdal, V. and Berge, T., 1987, "Solhøydediagrammer for området fra 70° til 82°N", Norsk Polarinstitutt.

Løset, S. et al., 1989, "IDAP'88 Vessel Deployment, Volume 2: Field Observations and Analysis", NHL F88093

Næss, T., Løvås, S.M. and Johansen, I.M., 1988, "Sea Ice Investigations in the Barents Sea, Stereo Aerial Photography. ICEBASE Report No 3", NHL STF60 F 87078

OBSERVATIONS OF SEA ICE DRIFT OFF NEWFOUNDLAND USING SATELLITE IMAGERY AND ICE BEACONS

I. K. Peterson and S. J. Prinsenberg

Physical and Chemical Sciences Branch,
Department of Fisheries and Oceans,
Bedford Institute of Oceanography,
P.O. Box 1006,
Dartmouth, N.S., B2Y 4A2

Daily sea ice velocity fields were obtained by manually tracking ice features in 80 pairs of sequential cloud-free visible NOAA/AVHRR satellite images for the region off Labrador and Eastern Newfoundland from 1984-1986. Although cloud-free images were found for winds from all directions, the images were least likely to be cloud-free when winds were onshore. During the winters of 1985 to 1987, ice velocities were also obtained using a total of 28 Argos satellite-tracked ice beacons, deployed off the Labrador coast by helicopter. The imagery-derived velocity fields were in good agreement with these data. With light to moderate winds, the influence of the Labrador Current could clearly be seen at the shelf break. Strong topographic steering was evident around the numerous banks and saddles on the shelf. Daily geostrophic wind, which was calculated from FNOC surface pressure data on a 380 km grid, accounts for about half the variance in daily ice velocity, for both the imagery and beacon data.

KEY WORDS: Newfoundland shelf, sea ice drift, AVHRR imagery, ice beacons

INTRODUCTION

The understanding of sea ice motion over the continental shelf off Labrador and Newfoundland is important for the hydrocarbon exploration and development taking place off the Canadian east coast, since sea ice exerts forces on offshore platforms, interferes with surface shipping support and affects iceberg-towing operations. The study of ice drift may also provide information on ocean circulation patterns for this region which could be useful for the prediction of iceberg trajectories and the drift of fish eggs and larvae.

The Labrador/Newfoundland shelf (Figure 1) consists of a series of banks and saddles, with a mean depth of about 200 m off Labrador, about 300 m over the Northeast Newfoundland shelf, and less than 200 m over the Grand Banks. The main core of the Labrador Current is located over the continental slope.

In this paper, observations of ice drift over the continental shelf, obtained from sequential satellite images for 1984-1986 are described. These are compared with data obtained from Argos satellite-tracked ice beacons deployed on ice floes by helicopter in 1985-87, and with gridded geostrophic wind data obtained from the Atmospheric Environment Service (Swail, 1985).

METHODS

Pairs of relatively cloud-free visible NOAA/AVHRR images (each at about 1600-1800 Z) approximately 1 day apart were navigated by means of ephemeris data, to an accuracy of 1-2 km. An example of a pair of images from March 1985 is shown in Figure 2. Ice tracking was done manually by flickering between the two images. In general, floes are smaller than the resolution of the imagery (1 km) across, so that diffuse patches of floes are tracked. Where possible however, individual ice floes larger than 1 km were tracked, using the centre of the floe. In fact, these large floes are usually aggregates of smaller floes

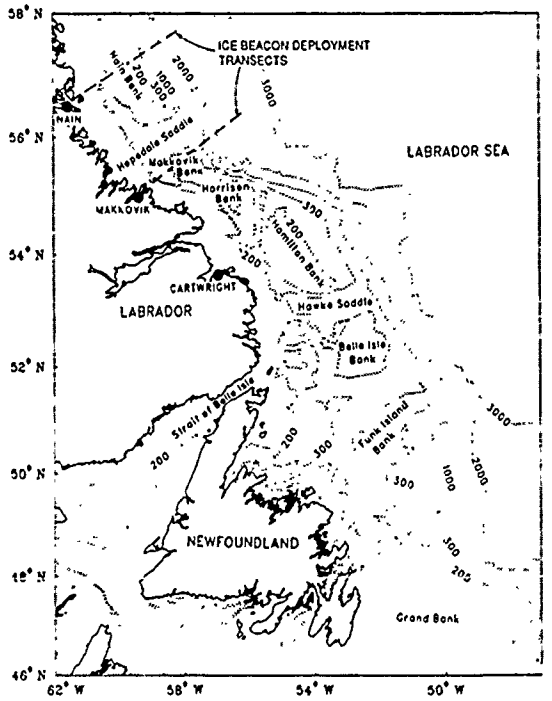


Figure 1. The bathymetry off Labrador and eastern Newfoundland (depths are in meters).

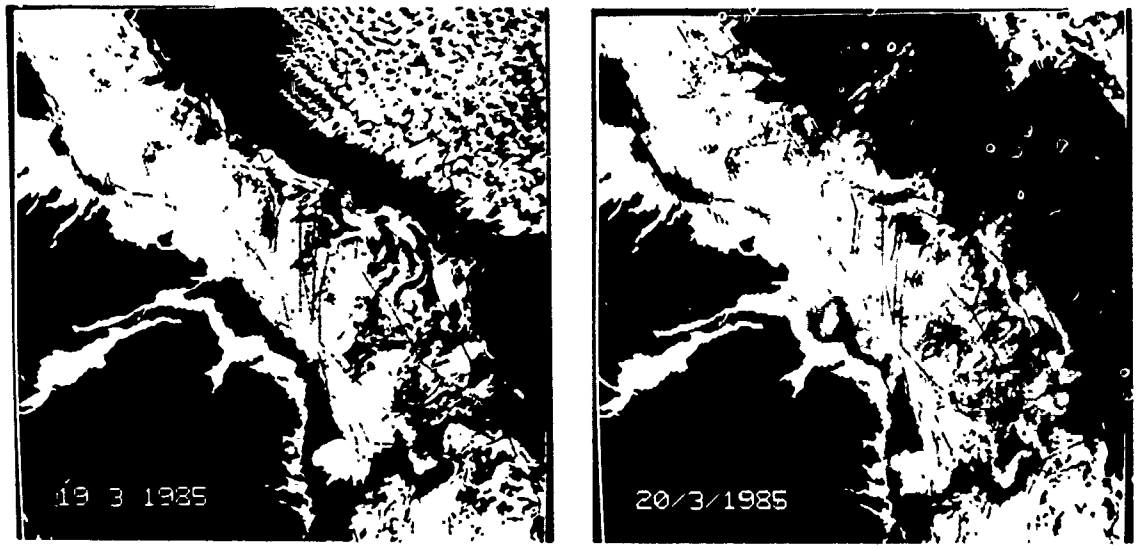


Figure 2. Visible AVHRR images off southern Labrador between 52 and 57°N, for March 19, 1985 (left), and March 20, 1985 (right).

which have frozen together or been compressed together against the shore. Obviously, the error associated with tracking will be higher for patches of floes than for individual floes. The error of the latter is probably 1-2 km.

In total, 80 velocity fields were obtained for 1984-86: 30 in 1984, 32 in 1985, and 18 in 1986. In addition, 12 velocity fields have been generated for 1987, but are not included in this analysis.

Ice velocities were also obtained for 1985-87 from satellite-tracked beacons deployed on ice floes off Labrador by helicopter (Peterson and Symonds, 1988). For 1985 and 1986, eleven beacons were deployed (half in January, and half in February), while for 1987, only one deployment of 6 beacons was carried out in January. For each deployment, three beacons were deployed off Nain and three off Makkovik (Figure 1). Some of the beacons were equipped with anemometers. About 7-12 positions per day were computed by Service Argos from the Doppler shift in beacon transmissions, with an accuracy of about 200 m. Daily ice velocities were calculated by linearly interpolating beacon positions to a fixed time of the day, for a total of about 1200 buoy-days of data.

EXAMPLES OF ICE VELOCITY FIELDS

Three examples of daily ice velocity fields obtained from satellite imagery for March 1985 are shown in Figures 3, 4 and 5 (narrow arrowheads). The ice velocities obtained from ice beacons over the same time interval (wide arrowheads) are in good agreement with those obtained from imagery. The corresponding geostrophic wind field is shown in the inset in the upper right hand corner, covering the same area as the large map.

For the first example (Figure 3), strong persistent alongshore winds for several days caused the pack-ice to move uniformly parallel to the coast on March 6-7. The other two velocity fields correspond to relatively weak wind conditions. The one for March 10-11 (Figure 4), follows several days of weak southwesterly winds. In contrast to the

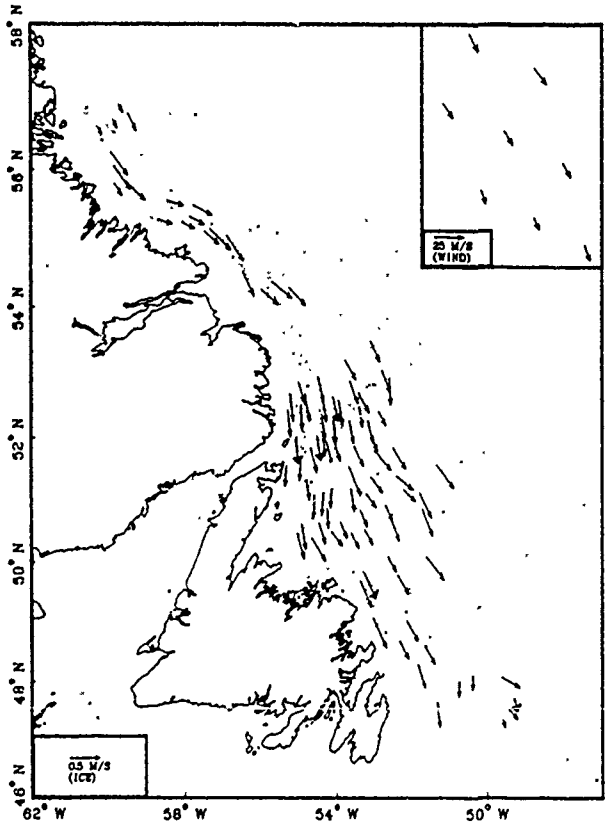


Figure 3. Daily ice velocities for March 6-7, 1985 from satellite imagery (narrow arrowheads) and ice beacons (wide arrowheads).

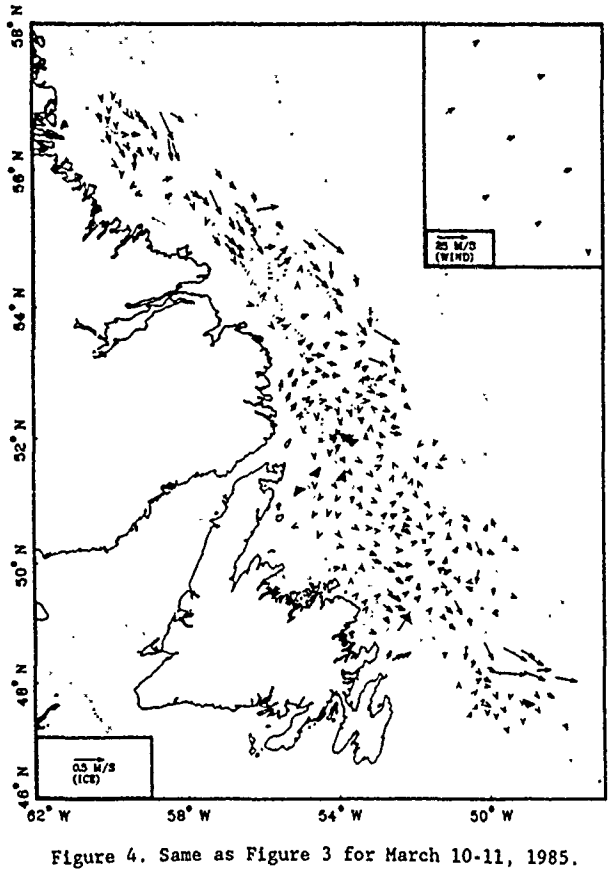


Figure 4. Same as Figure 3 for March 10-11, 1985.

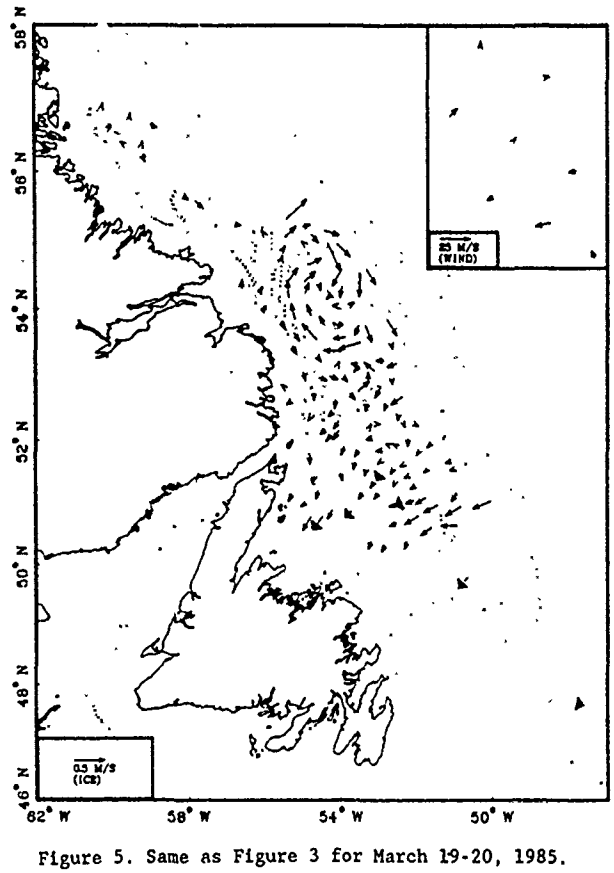


Figure 5. Same as Figure 3 for March 19-20, 1985.

previous figure, it shows a weak, less organized flow field. Yet the effect of the Labrador Current over the shelf is still visible. The flow pattern in the south is similar to the results of a barotropic model for this region (Greenberg and Petrie, 1988). The velocity field for March 19-20, 1985 (Figure 5) follows a period of longshore winds, and was derived from the images in Figure 2. The ice edge changed from being straight along the 1500 m contour on March 19, to a wavy pattern with a wavelength of about 80 km on March 20. A strong clockwise flow of ice is visible around Hamilton Bank, with a weaker counterclockwise flow around Hawke Saddle to the south of it. This may have resulted from ocean currents spinning down after several days of longshore winds.

WEAK WIND CASES

Figures 4 and 5 demonstrate the variability in the flow patterns associated with weak wind conditions, while at the same time they show the importance of ocean currents and bottom topography. A summary of many of the features seen in the other velocity fields corresponding to weak wind conditions is as follows.

The effect of the main core of the Labrador Current is clearly visible over the shelf break, where ice speeds are high, and the flow closely follows the bathymetry, centred over the 1000 m isobath. Although the ice edge usually coincides with the shelf break, in April 1984 the ice extended further offshore, and counterclockwise eddies about 30 km in diameter could be seen offshore of the Labrador Current (Peterson, 1987).

Over Nain Bank, and inshore of the bank along the marginal trough, the flow tends to be southward, but is deflected offshore at Hopedale Saddle. The flow over Makkovik and Harrison Banks and Cartwright Saddle is usually very weak. Along the southern edge of Hamilton Bank, part of the flow from the Labrador Current is often deflected inshore, where it sometimes continues around the western edge of the bank, as shown in Figure 5. Inshore of Hamilton Bank, there is usually a weak southward flow.

The flow along the northern edge of Belle Isle Bank tends to be offshore, with the flow shown in Figure 5 being an exception. A cyclonic eddy can often be observed in the southeast corner of Hawke Saddle. South of this, the currents are usually strongest over the 300 m contour on the western side of the channel separating Belle Isle Bank from the mainland.

There are many other instances of topographic steering east of Newfoundland, in particular a northeastward flow north of Funk Island Bank (Figure 4), and an eastward flow along the northern edge of both Notre Dame Bank and Grand Bank.

WIND EFFECTS

Because the use of visible AVHRR imagery depends on cloud-free images, there is likely to be a bias towards certain wind conditions which are most often associated with cloud-free conditions. In Table 1,

Table 1. Frequency and mean wind speed by quadrant (the direction the wind is from) (a) for all days, and (b) for days with cloud-free image pairs only.

Quadrant	Total Number	Total Frequency(%)	Cloud-free Images Number	Cloud-free Images Frequency(%)	Mean Wind Speed (m/s) Total	Mean Wind Speed (m/s) Cloud-free only
SW	73	27	15	21	8.4	7.1
SE	37	14	3	8	10.2	11.1
NE	71	26	18	25	11.8	8.9
NW	93	34	41	44	10.1	8.7

the fr quency of cloud-free images is shown as a function of wind direction, for the winters of 1984-86, The wind direction is for the grid point at 52° 48 N, 54° 42 W, and the image pair was considered cloud-free if any ice features within half a grid unit could be tracked. The overall frequency of winds in the four quadrants is also shown for this grid point, for the sum of the periods in each year between the first and last image pair.

For the three periods in 1984-86, the overall winds were in the southeasterly (direction from) quadrant only 14% of the time (37 days out of 274), of which the image pairs were cloud-free only 8% of the time (3 days out of 37). On the other hand, the overall winds were in the northwesterly quadrant 34% of the time, of which the image pairs were cloud-free 44% of the time. The other two quadrants were intermediate.

The mean wind speed in each quadrant is also shown in Table 1, along with the mean wind speed for the cloud-free image pairs. Overall, the mean wind speed for the cloud-free image pairs is 8.6 m/s, compared with 10.1 m/s for the total period. With imagery data, there is also a bias toward low ice speeds, since it is easier to track slow-moving than fast-moving ice features in imagery.

Ice beacon data, on the other hand, are also biased. The drift tracks, and hence the data distribution, are largely determined by wind conditions, so that there is likely to be a bias toward onshore winds near the coast, and offshore winds far from the coast. For gridded data, there is expected to be a bias toward slow-moving beacons, simply because they will stay in the region longer.

The effect of wind on ice drift in different grid squares was evaluated for the image data (1984-86) and the beacon data (1985-86) separately, using complex regression analysis. A regression model of the form

$$U_i = |A|\, e^{i\theta}\, U_a + U_c + \epsilon$$

was used, following Thorndike and Colony (198), where U_i is the ice velocity, U_a is the geostrophic wind velocity, $|A|$ is the scaling factor, θ is the turning angle, and ϵ is the error. The offset, U_c, which is the mean ice drift minus the effect of the mean daily wind, is termed the mean current.

The results of the regression were similar for the image data and beacon data, in spite of the wind biases mentioned above. The regression model accounted for about half the variance in ice velocity, in agreement with other studies of coastal ice drift (Thorndike and Colony, 1982; Johannessen et al, 1983).

CONCLUSIONS

Ice drift data obtained from satellite imagery is in good agreement with that from satellite-tracked ice beacons. The velocity fields derived from the imagery are of high spatial resolution over a large area, and therefore complement the beacon data, which provide continuous time series of ice drift.

Because of the high spatial resolution, ice tracking using satellite imagery also represents a useful oceanographic tool for studying details of the surface flow pattern in this region, and for verifying numerical models of ice drift or ocean circulation. Ice tracking is done over a period in the year when little other oceanographic information is available. Details of the ocean surface flow pattern may have implications for the drift of fish eggs and larvae. For instance, northern cod spawns along the Labrador shelf break during March to May, (Fitzpatrick and Miller, 1979) and the onshore flow observed at Hawke Saddle may represent an important pathway for the drift of fish eggs and larvae closer to shore, where they later develop.

ACKNOWLEDGEMENTS

We thank Val Swail of the Canadian Climate Centre for providing the geostrophic wind data, and Linda Payzant for assisting with the image analysis. This work was supported in part by the Canadian Federal Panel on Energy Research and Development (PERD).

REFERENCES

Fitzpatrick, C. and R.J.Miller,"Review of spawning times and locations for some commercial finfish on the Newfoundland and Labrador coasts", Fish. Mar. Serv. Tech. Rep. 905:v + 14 p, 1979.

Greenberg, D.A. and B.D. Petrie, "The mean barotropic circulation on the Newfoundland shelf and slope", J. Geophys. Res 93:15541-15550, 1988.

Johannessen, O.M., J.A. Johannessen, J. Morison, B.A.Farrelly, E.A.S.Svendsen, "Oceanographic conditions in the Marginal Ice Zone north of Svalbard in early fall 1979 with an emphasis on mesoscale processes", J. Geophys. Res. 88:2755-2769, 1983.

Peterson, I. "A snapshot of the Labrador Current inferred from ice floe movement in NOAA satellite imagery", Atmosphere-Ocean 25:402-415, 1987.

Peterson, I.K. and G. Symonds, "Ice floe trajectories off Labrador and eastern Newfoundland: 1985-1987", Can. Tech. Rep. Hydrogr. Ocean Sci. No. 104:vi+101p, 1988.

Swail, V.R. "Geostrophic wind climatology of Canadian marine areas", Canadian Climate Centre Report No. 85-9, 1985.

Thorndike, A.S. and R. Colony, "Sea ice motion in response to geostrophic winds", J. Geophys. Res. 87:5845-5852, 1982.

EFFECTS OF THE SEA ICE PATTERNS DURING THE AUGUST REVERSAL OF THE BEAUFORT SEA GYRE ON ATMOSPHERIC FORCING

Ellsworth LeDrew and Douglas Johnson,
Earth-Observations Laboratory,
Department of Geography,
University of Waterloo,
and the Institute for Space and Terrestrial Science,
and
Jim Maslanik
CIRES,
University of Colorado,
Boulder, Colorado
Phone: (519) 885-1211 X2783
FAX: (518) 888-4632

Typically the Beaufort Sea Gyre has an anticyclonic circulation. Recent studies of monthly data have illustrated a cyclonic reversal during August for most years of record. Examination of satellite imagery reveals that the sea ice diverges during this reversal. Consequently there is probably an increase in the heat flux from the ocean to the atmosphere. This anomalous heat flux may have a feedback effect on the stabilization of the low pressure cell that drives the reversed circulation.

We are examining Defense Meteorological Satellite Program visible and IR imagery as well as SMMR and recent SSM/I microwave imagery to study this problem. Specifically, for individual days during the reversal period we are mapping the area and pattern of open water and ice concentration evident on the imagery. Using these data in conjunction with National Meteorological Centre gridded atmospheric data we are calculating the surface turbulent heat flux from the water and ice. These values provide the lower boundary condition for a diagnostic model of atmospheric dynamics from which we can assess the role of the sea ice divergence on synoptic evolution.

In this paper we report on the linkage between the sea ice concentration as derived from satellite imagery and the spatial patterns of the surface heat flux for several cases. Through sensitivity analysis we evaluate the spatial resolution of ice pattern and concentration required to force atmospheric processes at the synoptic scale. We also evaluate the effect of the difference in the US and Canadian concentration algorithms on the atmospheric forcing.

SOME OBSERVATIONS OF THE SNOWCOVER ON SEA ICE IN THE GULF OF BOTHNIA

G. B. Crocker
Centre for Cold Ocean Resources Engineering
Memorial University of Newfoundland
St. John's, Newfoundland, A1B 3X5
(709) 737-8358, FAX: (709) 737-4706

Dirk Werle
ÆRDE Environmental Research
43 Parkhill Road
Halifax, Nova Scotia, B3P 1R4
(902) 477-2266

Abstract

During March of 1988, detailed measurements of the physical properties and depth distributions of the snowcover on the sea ice in the Gulf of Bothnia were made as part of BEPERS-88 (The Bothnian Experiment in Preparation for ERS-1). The measurements included profiles of the density, crystal structure, salinity, temperature, and brine volume (all at 1.2 cm depth intervals), and a total of 1084 snow depths. The profiles revealed that the near-surface layers of the snowpack contain a large liquid fraction, even when the ambient temperature is well below 0°C. The volume fraction of this highly saline brine, which is drawn up from the ice surface in the early stages of ice and snowcover development, varies temporally in response to brine drainage processes and to temperature (following the phase relation for the salt mixture). Most importantly, it was found that with the snow temperatures and salinities observed in this region, significant quantities of liquid are present even in mid-winter. The implications of this liquid fraction for the dielectric and backscattering properties of the medium (at radar frequencies) and for the interpretation of SAR imagery, are discussed. The depth measurements are divided into 5 groups representing 5 different ice surface roughnesses. The depth distributions show that ice surface roughness strongly influences the snow depth variations, with the general trend being toward decreasing snow depth variability with decreasing ice surface roughness, until the ice is very smooth when snow depth variability again increases. This phenomenon is attributed to changes in the aerodynamics of the deposition process, and the implications of the observed distributions for estimating ice surface elevations or ice surface roughness from laser profilometry (which records changes in snow surface elevation) are assessed.

MULTISENSOR REMOTE SENSING OF THIN ICE

I.G. Rubinstein, Earth Observation Laboratory, ISTS and CRESS
R.O. Ramseier, Ice Research, AES and ISTS
Petrie Building, York University, 4700 Keele Street
North York, Ontario, Canada M3J 1P3
Tel: (416) 739-4685
Fax: (416) 739-4687
E-mail: Envoy 100, York.M

The problem of detecting new ice (less than 0.10 m thick) from spaceborne sensors will be addressed in this presentation.

The similarities in signatures of new-ice pixels and ice-free ocean pixels, for the whole range of frequencies (from visual to cm wavelengths), lead to difficulties that limit the usefulness of the sensors' information. In addition, the algorithms developed for ice information retrieval are usually set up for global use, ie without taking into account regional differences.

Our previous work has shown that, in order to fully utilize remote-sensed information, a degree of flexibility in data interpretation as to be built-in into the model. This flexibility should include both regional differences and multi-sensor data use.

Our present algorithm introduces two new modules: one that allows for regional differences and another that includes multi-sensor information. Once a region of interest is identified from the global algorithm output, and if the presence of new ice is suspected, the user can selectively "turn-on" either one or both modules. In the proximity of an ice edge or in the marginal ice zones, the multi-sensor module will allow more accurate ice edge location. Data from sensors of high spatial resolution, such as OLS or NOAA, will be utilized for this purpose.

Once these modules are fully tested, they will be incorporated into a global algorithm that "flags" specific regions for the more detailed processing.

We will present the results of testing these modules on data from SSM/I (Special Scanning Microwave / Imager), and OLS or NOAA.

ESTIMATION OF THE THICKNESS OF UNDEFORMED FIRST YEAR ICE USING RADAR BACKSCATTER

E.D. Soulis, *W.C. Lennox*, and *J.F. Sykes*,

Department of Civil Engineering, University of Waterloo, Waterloo, Ontario N2L 3G1, Canada

Tel:(519)885-1211 Telex:069-55259 Fax:(519)888-6197

Abstract– Areal fractions of undeformed young and first year ice are important for the determination of the spatial variation of ice strength and mass for the modelling of ice motion.

This paper presents a basis for treating these types of ice as a single class. A series of relationships are derived that link the thickness of undeformed ice to radar image grey-level

The approach is based on a physical optical model for the radar return of undeformed ice. In this model most of the of the variation in return at a given angle of incidence can be explained by changes in the Fresnel reflection coefficient which is a function of the dielectric properties of ice. The complex dielectric constant is related to salinity and surface temperature which in turn depends on the age and thickness of the ice. The combination of these factors is a hyperbolic relationship that is well approximated by a power-law expression that relates ice thickness to image grey-level.

The results correspond well with a 13.3 GHz data set from the Canadian Centre for Remote Sensing and with an analysis of 10.0 GHz SLAR imagery of the Beaufort Sea

Keywords–radar imagery, backscatter, ice thickness

Introduction

Estimates of the spatial variation of ice thickness in a region are important data for the modelling of ice motion Thickness data is used in two ways. First, all motion models, with the exception of those for free drift, require some characterization of the strength of the ice pack. Various strength models have been proposed (Hibler, 1980; Coon, 1980) and all of them require the estimation of certain global parameters as well as local estimates of ice thickness. Secondly, all motion models, including those for free drift, require estimation of the local mass of the ice pack. Because ice thickness is not directly measurable from radar backscatter, acquiring these data from radar imagery is a significant challenge.

Fortunately, thickness estimates are only required at a spatial resolution that is determined by the element size of the numerical model being used in a region. This area can vary from 1 km^2 to 10,000 km^2 depending on the size of the region under study. Ice mass and ice strength are estimated at each node from the frequency distribution of the ice thickness in the neighbourhood of the node. Normally, this histogram is constructed using concentrations of various ice categories and typical thickness ranges for each category. Nodal ice mass is calculated using the mean of the thickness distribution and nodal ice strength is calculated in various ways that use, in effect, up to the first three moments of the undeformed young and first-year ice portion of the distribution (Soulis et al, 1987).

Nevertheless, meeting these requirements requires considerable effort. Producing an accurate thickness histogram requires identification of the various sub-classes of undeformed young ice (grey, grey-white) and undeformed first-year ice (thin, medium, and thick) which is a non-trivial image segmentation problem because the signatures of these sub-classes have significant overlap. Texture measures are some help (Holmes et al, 1984) but multi-channel imagery appears to be required (Livingstone et al, 1987).

This paper provides a basis for treating the above classes as a single class by presenting a relationship between radar backscatter and ice thickness that is sufficiently accurate on average to permit the transformation of backscatter statistics to thickness statistics. Attempts have been made to use imaging sensors to measure ice thickness (Page and Ramseier, 1975) but the results have not been highly successful because in general there appears to be poor correlation between ice surface characteristics and ice thickness. There is, however, some evidence that such a relationship does exist for cold undeformed young and first-year ice. Table 1 gives the backscatter and thickness ranges for these ice classes from a Canadian Centre for Remote Sensing data set (Livingstone et al, 1987). Backscatter appears to decrease monotonically with ice thickness.

Table 1: Winter ice signatures at 45° incidence angle (from Livingstone et al, 1987)

Ice Type	Backscatter (σ^0_{HH} dB) mean	standard deviation	Thickness (cm)
Young ice			
Grey	-12.85	1.61	10-15
Grey-white	-14.92	0.93	15-30
First-year smooth	-21.49	1.81	30-200

Radar returns from first-year ice

From energy balance considerations, the radar return from first-year ice is given by

$$\sigma^0(\theta) = \sigma^0_{ss}(\theta) + T_s^2(\theta)\left[\sigma^0_{sv}(\theta') + \frac{1}{L^2(\theta')}\{\sigma^0_{is}(\theta') + T_i(\theta')\sigma^0_{iv}(\theta'')\}\right] \quad (1)$$

where σ^0 is the backscatter coefficient, σ^0_{ss} is σ^0 for the snow surface, σ^0_{sv} is σ^0 for the snow volume, σ^0_{is} is σ^0 for the ice surface, σ^0_{iv} is σ^0 for the ice volume, T_s is the power transmission coefficient of the air/snow boundary, T_i is the power transmission coefficient of the snow/ice boundary, L is the one way loss through the snow layer, θ' is the angle of refraction in the snow, and θ'' is the angle of refraction in the ice. Kim (1984) shows that for first-year ice, the last term including $\sigma^0_{iv}(\theta'')$ becomes negligible at frequencies in the 10 GHz range and usually can be omitted

For dry snow, which is the most likely condition when the ice is cold, the real part of the dielectric constant for snow is only 1.6 or 1.7 (Hallikainen and Ulaby, 1981). Therefore, there is very little refraction of the radar beam as it passes through the snow cover. That is, θ' is very close to θ.

Thus, using representative values for the snow volume scattering coefficient, equation (1) simplifies to

$$\sigma^0(\theta) = \sigma^0_0 + \sigma^0_1\sigma_{is}(\theta) \quad (2)$$

where σ^0_0 and σ^0_1 are functions of θ. Therefore a surface-scattering model should explain most of the variation in $\sigma^0(\theta)$.

Among the many surface-scattering theories, the Kirchoff or physical-optical formulation is one of the most widely used. This theory is applicable to surfaces with gentle undulations whose horizontal dimensions are large compared with the incident wavelength. Thus a surface can be modelled as a collection of facets at which plane reflection is occurring. The return power can then be described statistically based on the probability density function and the autocovariance function of the reflecting surface This is still a formidable task and a variety of analytical approximations have been derived.

The following physical-optical model was proposed for first-year ice by Kim (1984) and has been verified experimentally by Xue et al (1986) using measurements of artificial ice at the Cold Regions Research and Engineering Laboratory outdoor tank. Surface scattering from first-year ice is computed using a physical-optics model with an exponential correlation function (Kim et al, 1985). Most of the radar return is contained in the incoherent scattering component and therefore, $\sigma^0_{is}(\theta)$ can be determined by

$$\sigma^0_{is}(\theta) = 2k^2|R_p(\theta)|^2\cos^2\theta\exp\{-(2k\sigma\cos\theta)^2\}\cdot \sum_{n=1}^{\infty}\frac{(4k^2\sigma^2\cos^2\theta)^n}{n}\frac{k^2(n/L)}{(4k^2\sin^2\theta + n^2/L^2)^{3/2}} \quad (3)$$

where k is the radar wave number, $R_p(\theta)$ is the complex Fresnel reflection coefficient for polarization p at incidence angle θ, σ is the rms height of the surface and L is the correlation length of the surface. This model applies to surfaces with an average radius of curvature greater than the radar wavelength and an rms slope of the surface of less than 0 25. Most undeformed first-year ice meets these conditions (Kim, 1984).

Substituting equation (3) in equation (1) gives

$$\sigma^0(\theta) = \sigma^0_0 + \sigma^0_1 C|R(\theta)|^2 \quad (4)$$

where C is a function of θ

The incoherent scattering component varies considerably with angle of incidence which accounts for the change in brightness across a radar image. To give a uniform appearance either sensitivity time control is used during image acquisition or some post-processing technique is used to correct for the angular variation in σ^0_{is}. This results in

$$\sigma^0 = \sigma^0_2 + \sigma^0_3|R(\theta_r)|^2 \quad (5)$$

where θ_r is some reference angle and σ^0_2 and σ^0_3 are constants for the image.

Finally, the digital grey levels in an image are normally a linear transformation of returned power. Therefore, for undeformed first-year ice

$$s = s_0 + s_1|R(\theta_r)|^2 \quad (6)$$

where s is image grey-level and s_0 and s_1 are constants that depend upon the signal processing that has occurred.

Thus the grey-levels observed in an well-calibrated image will vary with the magnitude of the Fresnel reflection coefficient of the ice surface.

The Fresnel reflection coefficient

The Fresnel reflection coefficient $R(\theta_r)$ is the ratio of the field intensity of a reflected plane wave to the field intensity of a plane wave incident at an interface between two media at angle θ. Maxwell's equations for propogation of an electromagnetic wave in a lossy homogeneous medium are

$$\nabla \times E = -\mu\frac{\partial H}{\partial t}$$
$$\nabla \times H = -\sigma E + \epsilon'\frac{\partial E}{\partial t}$$

where E is the electric field vector, H is the magnetic field vector, μ is material permeability, ϵ' is material permittivity and σ is material conductivity.

Requiring that the tangential and normal boundary conditions match at the interface and evaluating $|R(\theta_r)|^2$ gives

$$|R(\theta_r)|^2 = \frac{[\cos\theta_r - \sqrt{r}\cos(\theta/2)]^2 + r\sin^2(\theta/2)}{[\cos\theta_r + \sqrt{r}\cos(\theta/2)]^2 + r\sin^2(\theta/2)} \quad (7)$$
$$\theta = \arctan\left[\frac{\epsilon''}{\epsilon' - \sin^2\theta_r}\right]$$
$$r = \{[\epsilon' - \sin^2\theta_r]^2 + [\epsilon'']\}^{\frac{1}{2}}$$

where θ_r is the angle of incidence of a wave in medium 1 on the surface of medium 2 measured from a normal to the surface, ϵ'' is the imaginary part of the dielectric constant given by σ/ω and ω is the frequency of the incidence wave. By substituting equation (7) into equation (6), the variation in the grey-level for undeformed ice is determined by two calibration coefficients s_0 and s_1 and the complex dielectric constant for sea ice.

The relationship between dielectric constant and thickness

The complex dielectric constant, ϵ, determines the electromagnetic properties of a material. The real part, ϵ', is related to the amplitude of a wave propogating through the medium and the imaginary part, ϵ'', is related to the attenuation of the wave. Neither component depends directly on ice thickness. Rather, both are functions of ice salinity, S, and ice temperature, T, both of which can be related to thickness. Vant et al (1978) used AIDJEX data to relate ϵ, S, and T empirically and found ϵ well approximated at 4 GHz by

$$\epsilon' = 3.05 + 7.20V_{br} \quad (8)$$

$$\epsilon'' = .024 + 3.29V_{br} \quad (9)$$

where V_{br} is the volume of brine in the ice in parts per thou-

sand. The volume of brine can be calculated by (Frankenstein and Garner, 1967)

$$V_{br} = S(\frac{-49.185}{T} + 0.532) \qquad -27.9°C \leq T \leq -0.5°C \quad (10)$$

The values of the dielectric constant calculated from equations (8) and (9) may be adjusted for other frequencies using (Kim, 1984)

$$\begin{aligned} \epsilon'_f &= (.995 - .00154f)\epsilon'_{4GHz} \qquad 5GHz \leq f \leq 18Ghz \quad (11) \\ \epsilon''_f &= (.914 - .00546f)\epsilon''_{4GHz} \end{aligned}$$

Since both ice thickness and salinity vary in time, it is possible to establish a relationship between the two. An approximation for the change in ice thickness (Cammaert and Muggeridge, 1988) is Stefan's equation in the form

$$h = C\sum_{i=1}^{t}(T_m - T_a)^{1/2} \quad (12)$$

where h is ice thickness in meters, T_m is the freezing point of sea ice ($-1.8°$ C) and T_a is mean daily air temperature. Replacing T_a with an average value for the winter months of the Arctic of about $-20°$ C gives

$$h = h_0 t^{(1/2)} \quad (13)$$

where h_0 is a constant equal to 10.6 cm and t is time in days. Nakawo and Sinha (1981) found that surface salinities are about 12 parts per thousand (ppt) after initial freezing and drop about 0.5 ppt per month thereafter. Therefore

$$S = S_0 - kt \quad (14)$$

where S is salinity in ppt, S_0 is initial salinity, k is 0.5 ppt/month, and t is time in months. Combining these two equations gives

$$S = S_0 - k(h/h_0)^2 \quad (15)$$

The relationship between temperature and thickness involves some assumptions about the temperature gradients through the ice and snow and at the air/snow interface. The heat transfer mechanisms at the snow surface are primarily convection and radiation whereas the transfer through each material is by conduction. The boundary condition at the ice/water interface is the melting point of ice, T_m, and at the top of the air/snow boundary layer is T_a, the ambient air temperature.

The temperature gradient across the air/snow interface does not appear to be large. Maykut and Untersteiner (1971), using a rigorous radiation model, found that predicted snow surface temperatures were within 2°C of measured air temperatures. Nakawo and Sinha (1981) were able to model field measurements of snow and ice temperature profiles by assuming the surface temperature of snow and air temperature were the same. Therefore, by equating heat flux through the snow layer with heat flux through the ice,

$$\frac{k_s}{h_s}(T_a - T_s) = \frac{k_i}{h_i}(T_s - T_m) \quad (16)$$

where k_s and k_i are the thermal conductivities of snow and ice, h_s and h_i are the thicknesses of snow and ice, and T_s is the surface temperature of the ice.

Solving equation (16) for T_s gives

$$T_s = \frac{k_s\frac{h_i}{h_s}T_a + k_iT_m}{k_i + k_s\frac{h_i}{h_s}} \quad (17)$$

Following Nakawo and Sinha (1981), k_s is .0006 cal cm^{-1} s^{-1} deg^{-1}, k_i is .005 cal cm^{-1} s^{-1} deg^{-1}, and T_m is $-1.8°$C, which gives

$$T_s = .545T_a + 123.4 \quad (18)$$

where T_s and T_a are in degrees Kelvin.

Comiso (1983) conducted a linear fit between for T_s to the predictions of the Maykut and Untersteiner (1971) model (Table 2) for each of the four seasons. The parameters are close to the values predicted by equation (18) which confirms equations (18) and (17) as approximations for ice surface temperature.

A power law approximation

The relationship between returned radar power and ice thickness is obtained by combining equations (7), (8), (9), (10), (15) and (17). Results for SLAR imagery of the Beaufort Sea and the CCRS signature data (Table 1) are plotted in Figures 1 and 2. In both cases the exact curves are well approximated by

$$|R(\theta_r)|^2 = r_0 + r_1h^\beta \quad (19)$$

where r_0, r_1 and β are constants.

Substituting equation (19) into equation (6) gives

$$s = s_0 + s_1h^\beta \quad (20)$$

Table 2: Linear regression of ice surface temperature on snow surface temperature for the Maykut-Untersteiner model (from Comiso, 1983)

Season	Linear fit parameters		Months used in fit
	a	b	
Spring	0.524	126.656	Dec.,Jan.,Feb.,Mar.
Summer	0.533	126.735	Mar.,Apr.,May, Jun.
Fall	0.513	132.306	Jun.,Jul.,Aug.,Sep.
Winter	0.669	93.214	Sep.,Oct.,Nov.,Dec.

$T_s = aT_a + b$

Results and conclusions

Equation (20) is a parametric relationship between s and h. The values of s_0, s_1, and β will vary from image to image but they can be readily determined from a minimum of signature information. Figure 3 shows the best fit of equation (20) to the first-year signatures for SLAR imagery of the southern Beaufort Sea used by Soulis et al (1987). The data points used in the calibration correspond to the mean thickness and backscatter for the two undeformed first-year ice types identifiable in the imagery. The grey-level value at the thickness that separates the classes was taken as the grey-level at which either class was equally likely. Similarly, Figure 4 shows the best fit to the CCRS data set (Table 1).

Both curves have the shape predicted in Figures 1 and 2. This supports the use of equation (20) as a basis for interpreting the thickness of undeformed young and first-year ice on radar imagery.

Acknowledgements

Funding for this study was provided by the National Science and Engineering Research Council (NSERC) post graduate scholarship awarded the senior author and the NSERC strategic grant entitled: *The Image Analysis of Sea Ice.*

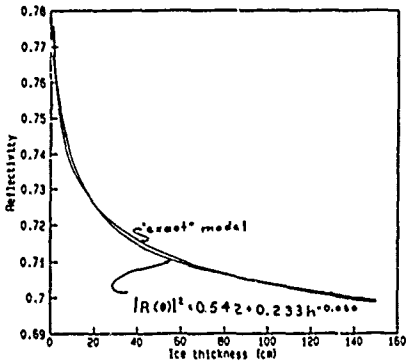


Figure 1: Backscatter versus ice thickness for $\theta = 82.1°$ and $f = 10.0$ GHz

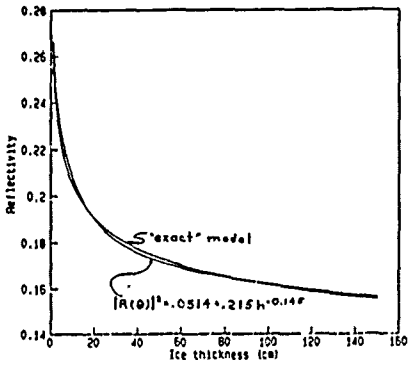


Figure 2: Backscatter versus ice thickness for $\theta = 45.0°$ and $f = 13.3$ GHz

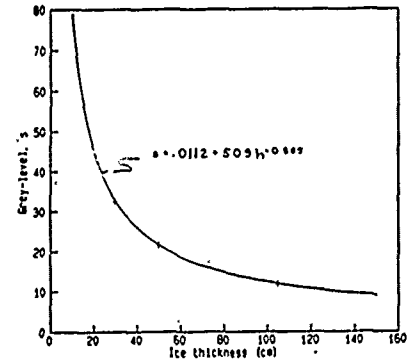


Figure 3: Grey-level versus ice thickness for the Beaufort Sea SLAR imagery

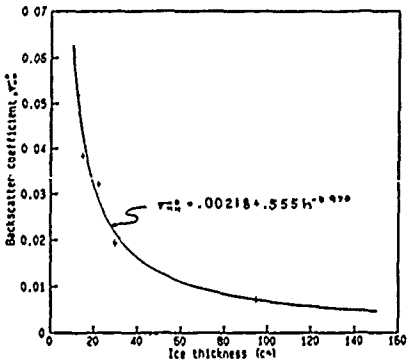


Figure 4: Backscatter versus ice thickness for the CCRS data set

References

[1] Cammaert, A.B. and Muggeridge, D.B., Ice Interaction With Offshore Structures, Van Norton-Rheinhold, New York, 1988.

[2] Comiso, J.C., Sea ice effective microwave emissivities from satellite passive microwave and infrared observations, *Journal of Geophysical Research*, Vol. 88, No. C12, 1983.

[3] Coon, M.D., Maykut, G.A., Pritchard, R.S., Rothrock, D.A. and Thorndike, A.S., Modelling the pack ice as an elastic-plastic material, *AIDJEX Bulletin* 24, 1-105, 1974.

[4] Frankenstein, G. and Garner, R., Equation for determining the brine volume of sea ice from −0.5° to −22.9°C, *Journal of Glaciology*, Vol. 6(48), 1967.

[5] Hallikainen, M. and Ulaby, F.T., Dielectric and scattering behaviour of snow at microwave frequencies, *Proceedings of the IGARSS Symposium*, Zürich, Switzerland, 1986.

[6] Hibler, W.D., Modeling a variable thickness sea ice cover, *Monthly Weather Review*, Vol. 108, No. 12, 1980.

[7] Holmes, Q.A., Nuesch, D.R. and Schuchman, R.A., Textural analysis and real-time classification of sea-ice using digital SAR data, *IEEE Transactions on Geoscience and Remote Sensing*, Vol. Ge-22, No. 2, 1984.

[8] Kim, Y.S., Theoretical and Experimental Study of Radar Backscatter From Sea Ice, Ph.D. Thesis, University of Kansas, 1984.

[9] Kim, Y.S., Moore, R.K., Onstott, R.G. and Gogineni, S., Towards identification of optimum radar parameters for sea-ice monitoring, *Journal of Glaciology*, Vol. 31, No. 109, 1985.

[10] Livingstone, C.E., Singh, K.P. and Gray, A.L., Seasonal and regional variations of active/passive microwave signatures of sea ice, *IEEE Transactions on Geoscience and Remote Sensing*, Vol. GE-2[illegible]

[11] Maykut, G.A. and Untersteiner, N. [illegible] a time-dependent thermodynamic model [illegible] *of Geophysical Research*, Vol. 76, No.

[12] Nakawo, M. and Sinha, N.K., Growth rate and [illegible] file of first-year sea ice in the high arctic, *Journal of Glaciology*, Vol. 27, No. 96, 1981.

[13] Page, D.F. and Ramseier, R.O., Application of radar techniques to ice and snow studies, *Journal of Glaciology*, Vol. 15, 1975.

[14] Soulis, E.D., Sykes, J.F. and Venkatesh, S., Variation of pack ice strength using radar backscatter, *Proceedings of the Eleventh Canadian Symposium on Remote Sensing*, University of Waterloo, 1987.

[15] Ulaby, F.T., Moore, R.K. and Fung, A.K., Microwave Remote Sensing Active and Passive, Vol. II., Addison-Wesley Publishing Company, Reading, Massachusetts, 1982.

[16] Vant, M.R., Ramseier, R.O. and Makios, V., The complex-dielectric constant of sea ice at frequencies in the range 0.1-40 GHz, *Journal of Applied Physics*, Vol. 43, No. 3, 1978.

[17] Xue, D., Moore, R.K., Williams, T.H.L. and Onstott, R., Microwave backscatter coefficients of artificially grown first-year sea ice. *Proceedings of the International Geoscience and Remote Sensing Society '86 Symposium*, Zürich, Switzerland, 1986.

A 3-DIMENSIONAL MODELLING OF FOREST CANOPIES FOR HIGH RESOLUTION IMAGERY

Richard Fournier and John R. Miller,
York University,
Centre for Research in Experimental Space Science,
North York, Ontario, M3J 1P3

Abstract

Spatial variation in the reflected radiance, referred to as texture, which is observed by an airborne sensor over a forest canopy contains valuable information. This information relates to the canopy's physical morphological parameters such as tree species, height, biomass, spacing and canopy shape. Human interpreters have been highly successful in providing estimates of such physical parameters from textural information of air photos. However, the development of textural analysis of digital imagery in remote sensing applied to forestry is currently at a relatively primitive stage. Future application of high resolution remote sensing images to forest canopies may be dependent on the development of models and/or algorithms to assist in extraction of canopy parameters and spectral scene classification.

This research consists of a comprehensive study of the influencing parameters of a forestry scene at high spatial resolution. Utilizing a geometric/optical approach, a canopy radiance model is formulated in order to explore the relationship between image radiometric texture and canopy physical parameters. The geometric/optical model assumes a hierarchy of primary and secondary influencing factors. The primary geometrical parameters of influence are: the tree shape, the pixel dimension, the sun and the viewing direction in addition to optical characteristics, such as the amount of collimated and diffuse sky illumination, the reflectivity and transmissivity factors, and the projected and mutual shading. Secondary parameters of influence modelled include the light penetration in the canopy, a more refined evaluation of the diffuse sky illumination, and the radiance from surrounding pixels in the scene. A modelled image can then be generated from the reconstituted scene by taking these factors of influence into account. The textural model features which best produce the variance of the image radiance patterns have been investigated through the examination of airborne imagery compared to the modelled texture.

Keywords: canopy model, canopy parameters, forestry, geometric/optical model, radiance patterns

The availability of high resolution images gives rise to new perspectives in the field of interpretation of remote sensing images. This work focuses on the influence of forest canopy parameters, and their modelling, on the assessment of the parameters' importance in the resulting radiance patterns of remote sensor imagery. The investigation of the canopy scene features are highly relevant to improving the retrieval of information in collected images. The pixel size assumed in the following research is typically 0.5 m, but may vary by about a factor of three. Such spatial resolution comes about because airborne sensors provide imagery at that resolution, and the extraction of important canopy parameters is expected to be optimized at that scale (maximum spatial coverage with maximum information retrieval). The understanding of canopy parameters' influence remains at a primitive stage. The complexity of forest canopies in their architecture and their various environments makes the task arduous. In fact, the shapes and the spectral responses within a forest scene vary drastically, at this resolution.

We propose to approach the problem from a theoretical and deterministic perspective. This approach consists of building a 3-dimensional model, which contains the necessary geometrical and optical assumptions to reproduce a scene in its principal characteristics. The geometrical reproduction

of the scene allows the quantification of the model's variables as close as possible to their physical representation. A good geometric representation will determine the effectiveness of the model implementation. The radiometric interactions are applied individually in hierarchical order so that each contribution to the resulting radiance pattern can be distinguished. The development of the geometrical/optical model involves the elaboration of its basic structure at two levels. First, the geometric representation of the modelled scene relies on certain considerations:

- each individual tree (refered to as scene object) or species will have a distinctive mathematical expression to represent its geometrical shape. The variables embedded in the algebric formulation will directly express the important physical canopy parameters, such as the tree's height, base diameter and overall shape,
- the placement of each individual represented object builds the scene,
- the forest canopy is represented as a grid covering the modelled trees. The distance between each point of the grid is directly correlated with the pixel dimension. Therefore, the basic model element becomes the grid's polygons which can be directly related to the image's pixels, and
- the vectorization of the polygon normal, sun and point of view which represent the geometrical directions of importance.

Secondly, the optical aspects of the model development are dependent on the above geometrical considerations. A list of primary factors of influence would thus contain the:

- specification of the illumination characteristics (direct and diffuse) for each modelled pixels,
- evaluation of the shadowed pixels,
- spectral response of each modelled object considered in terms of the selected wavelength and the relevant vectors' direction.

The secondary factors of influence result in more subtleties in the variance of the radiance textural pattern. The decision to choose these factors is based on consideration of literature, coupled with logical inference. Thus it was decided to model the effects of the:

- light penetration in the canopy,
- viewed portion of the sky to correct the received diffuse irradiance at each modelled pixel,
- radiance reflected from other points of the scene.

Due to practical constraints, certain assumptions are made to reduce the complexity of the processes implemented. These assumptions are primarily made for the modelling of the secondary parameters of influence.

Tree Shape

The first modelling consideration involve the geometrical representation of the basic component of a forest, the tree. A flexible mathematical expression is required to deal with the wide variety of tree shapes present in a forest canopy. The generation of geometrical trees relies initially on an equation which calculates height values in terms of the input basic geometric tree parameters. A modified version of the simple equation suggested by Horn (1971) is convenient for the representation of tree shapes. It basically expresses the transverse section of the shape in the vertical and in one horizontal plane which can be visualized as the cross-section of a tree. Adopting cylindrical 3-dimensional cartesian coordinates, where z represents the vertical axis, the four parameters are related by the equation:

$$z = z_{ref} + [c^a - (bx)^a]^{1/a} \qquad (1)$$

Here the variable x can be replaced by y because shape symmetry in the horizontal plane is assumed at this point. Each of the constants a, b and c expresses a different physical parameter. The parameter a defines the convexity of the line representing the tree shape in a vertical plane such a straight line, a convex or a concave curve can be generated as a takes on a value equal to, greater than, or less than unity, respectively. The acceptable values for a range from almost 0 to a value near 10, at which point the resulting curve takes the shape of a square function. Parameter b expresses the ratio of tree height to width (from the center to the maximum vertical extension). The third parameter c, is the tree height from the maximum horizontal extension to its summit. The maximum horizontal extension of the tree, y_{max} or x_{max}, is expressed as the simple ratio of c/b. In order to represent the height of the branching structure above the ground, it is convenient to consider a reference height (z_{ref}) from the ground to the point where the tree is maximally extended horizontally. Accordingly, the variable z

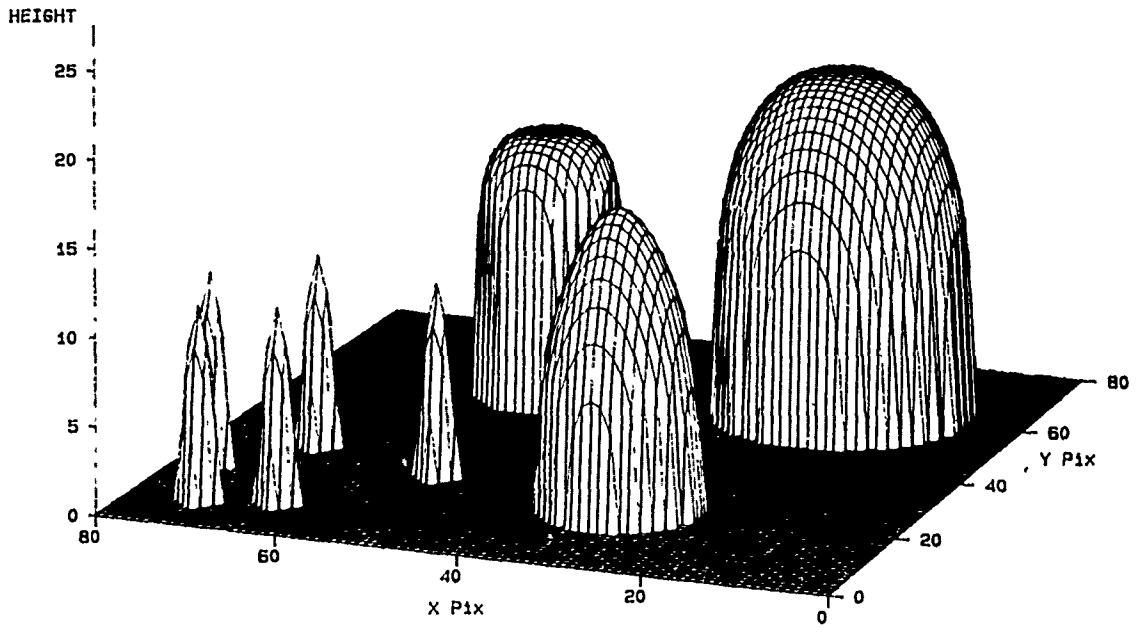


Figure 1; Example of a 3-dimensional surface representation of a modelled scene which include a variety of tree geometries.

represents the tree's vertical height with respect to ground level.

When rotated about the vertical axis at a given angular interval, the curve generated by the Equation 1 - which is the tree profile in two dimensions (horizontal and vertical) - generates a cylindrically-symetrical, 3-dimensional object. Furthermore, this equation provides a flexible tree geometry representation with only four significant parameters. As shown in Figure 1, the object cross-section generated by the Equation 1 is integrated into a 3-dimensional surface. The resultant scene therefore appears as a grid, or mesh containing the trees which reveal their shape at the tips of their branches. The application of this mathematical expression requires the generation of the typical and extreme values of the parameters a, b, c and z_{ref} for different species and environments.

Scene Geometry

Each of the mathematically-defined trees is represented on a 3-dimensional axis system. The horizontal sampling points are regularly spaced following a grid where the mesh is directly associated with the selected spatial resolution. A line segment joins each point to its adjacent x and y value to form polygons with diverse orientations. These polygons can be considered as the modelled pixel's boundary. The shadowed pixels can be evaluated at that stage, since all the necessary geometrical considerations are present. By utilizing the direction of the sun's vector, two distinct types of shadowing can be specified:

1. Shadowing on a generated object due to sun location and polygon orientation,
2. Shadowing due to a *mutual masking* of polygons in the scene.

The final scene radiance is characterized by contrasting shadowed zones. Since more variability in the radiance pattern is usually observed between shadowed and non-shadowed areas of an imagery, three subroutines representing secondary parameters of influence have been introduced. As a result, more continuous radiance values are generated at the boundary between the shadowed and non-shadowed polygons.

Interaction Radiation/Canopy

Both tree and scene background signatures are associated with the modelled elements composing the scene, therefore, they are assigned a particular reflectivity coefficient (or albedo) at a selected wavelength. Lambertian reflective properties are often attributed to forest canopy and leaves/needles;

The final radiance from the modelled element is a consequence of three distinct processes. The most important process is the reflected insolation (L_I), which includes the direct and diffuse components of the incoming irradiance. The two other processes are the radiances issued from the transmitted flux within trees (L_T), and the reflected flux from surrounding surfaces (L_S). The total radiance at an element (L_{Elem}) is therefore the simple algebric sum of these three components:

$$L_{Elem} = L_I + L_T + L_S \qquad (2)$$

The implementation of this canopy model is dependent on the cumulative effect of various geometrical/optical factors. The model initially simulates the rough radiance pattern, expected from a geometrical/optical point of view, using basic parameters. These *primary canopy parameters of influence* are expressed by the tree geometry, the scene geometry, and the radiation/canopy interaction with respect to the spectral irradiance of a collimated source. Secondary parameters are also considered, allowing for a more accurate explanation of the variability in the observed radiance patterns. Among these factors are: 1) the radiation penetration in the canopy, 2) an estimate of the diffuse radiation influences, and 3) the inter-pixel radiance processes. The distinction between primary and secondary parameters of influence is based upon their relative importance in relation to the modelling implementation and to the interaction between the incoming radiation and the forest canopy.

Correlation Model/Sensor Images

From the inspection of the modelled and sensor images, the overall radiometric patterns appear to correspond reasonably well. The general shape of the bright and shadowed zones match the patterns suggested by the implementation of the primary factor of influence, i.e. the *croissant* shape of the brighter area and a shadowed area starting roughly from the top of the tree, expanding in a V shape in the direction of the sun rays and creating a large dark area on the background, dependent on the tree surface masking of the collimated irradiance. Some secondary radiometric features also corroborate the effects of the secondary parameters of influence modelled. The light penetration definitely influences the observed radiometric values at the extremity of the shadowed areas and at the interface between the tree bright and dark areas. The contribution from the surrounding areas appears significant, as predicted by the model. This contribution of radiance makes the shadowed area on the tree and on the background near the tree slightly lighter than other dark areas farther from the tree's influence. This feature can almost certainly be attributed to the contribution of neighbouring pixels, since the sensor's radiance patterns show the same tendencies as the modelled radiances on these areas. The effect of the adjustment of the portion of sky viewed at each pixel is not clearly apparent in the sensor image.

The modelling of the tree geometry as a simple object, as it is done in the actual model, has its limits in the reproduction of the radiance pattern of a real sensor image. The results from the simulation suggest that efforts should be made in improving the representation of the tree architecture before futher attempts at improving the algorithms describing the parameters of influence.

The geometrical/optical model, as it is now implemented, provides an excellent opportunity to investigate the factors of influence and to simulate a wide variety of forest canopy configurations. Improvements are presently being investigated which address the present model weaknesses. The model modularity will permit modification in both the number and description of parameters of influence. The model could be adapted to other applications, like terrain or urban modelling.

References

Horn, H. S., "The adaptive geometry of trees", Monograph in Population Biology - 3, Princeton Univ. Press, Princeton, New Jersey, 1971.

PREPROCESSING OF AIRBORNE MULTISPECTRAL SCANNER DATA FOR FOREST DAMAGE ASSESSMENT IN HILLY TERRAIN

F. Frey, A. Bischof, P. Meyer, K.I. Itten and K. Staenz

Remote Sensing Laboratories
Department of Geography
University of Zurich-Irchel
CH-8057 Zurich, Switzerland

ABSTRACT

The usefulness of airborne scanner data flown without inertial platform navigational information is severely affected because of the uncontrolled geometry. Also radiometric corrections require a stable geometry, therefore special attention had to be given to careful specific preprocessing steps. A precise approach was necessary because of the heterogeneity of the forest stands in our testsite as well as the hilliness of the terrain.

Using ground control points, a simulation of the flight path could be calculated. A digital elevation model then allowed for geometric correction of the scanner data. Simultaneously, aspect and slope information from the digital terrain model was used to radiometrically correct terrain induced illumination variations.

These preprocessing steps were considered imperative prior to a multispectral analysis of the data for forest damage assessment.

Keywords: Airborne MSS data, preprocessing, forest damage, geometric corrections, radiometric corrections, digital elevation model.

1. INTRODUCTION

Forest damage is a visual effect of man induced and natural factors affecting our environment The study of this forest damage and its dynamics over time may answer questions on causes of the whole problem. A first step is certainly an inventory and classification of the damages. This in most european countries is achieved through the interpretation of false color infrared airphotos. A second problem is the continous surveillance of the development and dynamics of the damages, which necessitates large investments. Attempts to automize some of the procedures have to be discussed.

Parallel to the first operational forest damage assessments with airphoto interpretation, tests of utilizing airborne multispectral scanners were conducted. These should provide a basis for later applications of spaceborne techniques, using for instance Landsat or SPOT type data for the monitoring task. One cannot say, that those early experiments were particularly successful (Kadro et al., 1985). In large homogeneous monoculture forest stands, in flat terrain only, useful results were achieved. But hilly terrain, asking for geometric corrections and associated radiometric control was another case. Moreso, in a typical Swiss forest environment, the small scale variability within the forest caused difficulties in assessing the treetype' specific damages. Because in Switzerland forest damages in prealpine and alpine areas are of prime interest, a study was carried out to try to precisely correct airborne scanner data for geometric and radiometric effects before the assessment of the damage itself.

Of course, much literature can be found on various aspects of preprocessing and processing of forestry airborne multispectral scanner data. But if corrections have been performed at all, very rarely geometric and radiometric corrections were applied simultaneously. A specific problem arose in our tests, because no navigational control data from an INS was available from the flights. Therefore a methodology had to be developed for such specific cases.

2. PROCEDURE

On 29 August 1985 our testsites in northwestern Switzerland were flown with a tandem configuration of 2 Bendix M2S Multispectral Scanner Systems by DLR (the German Aerospace Research Establishment). One testsite in a flat valley bottom served for forest stand and damage classification, whereas our second testsite covered hilly terrain, and was meant for the assessment of geometric and radiometric correction procedures. The 2.17 mrad system allowed for 6.3 to 6.9 m ground resolution cells. Two scanners were used in parallel to simulate with the additional bands the spectral characteristics of the Landsat Thematic Mapper. As a consequence 2 differing optics produced the data, therefore different correction factors were necessary for each system. Tests with the data of the second scanner which incorporated Thematic Mapper Simulation channels 5 and 7 revealed, that due to transmission errors, those bands could not be used.

As a first systematic step, the image data was corrected for sensor related radiometric effects. Every 60th scanline showed erraneous

information and had to be replaced by averaging values from the neighbouring lines. Pixel numbers in different bands were not matching due to shifts. They had to be individually adjusted.

Based on 1 : 5000 scale base maps a digital terrain model was formed using digitizing equipment and the ARC/Info software package. A 5 m grid size was chosen for x and y. A height resolution of 1 m was selected because total test area height differences amounted to 250 m.

From the digital elevation model, the corresponding values for aspect and slope angle were derived, as well as a synthetic illumination calculated according to the time of the overflight.

Geometric distortions inherent in airborne multispectral scanning may prevent a useful information extraction. The three principal factors governing those distortions are:

- the scanning manner itself
- the instability of the platform
- the topography

Fig.1 shows the effect of topography and viewing angle on the sideways distortion. It may be consulted for the discussion on whether a digital elevation model should be used to correct the geometry or not. If it is, as in our case, necessary, a rubbersheet stretching approach using passpoints and interpolation techniques is clearly not feasible.

Guindon (1980) proposes as a first step a simulation of the flight path of the aircraft. The following assumptions are made:

- constant flying height
- the flight path is regarded as a straight line
- roll, pitch and yaw are in a first attempt neglected
- the true airspeed is a constant

As in other approaches ground control points are chosen which are easily discernible in the sanner image as well as in the maps. Their image coordinates (p,l), map coordinates (u,v), and elevation (H) are retained. For each passpoint pair the image based ground range (g_i) is calculated according to the following form:

$$g_i = (h_M - H_i) \times \tan\theta_{max} \times \frac{(N_O - 2p_i + 1)}{N_O - 1} \qquad (1)$$

g_i = ground range (image based)
h_M = flying height
H_i = elevation of ground control point
θ_{max} = FOV
N_O = pixels per scan line
p_i, l_i = image coordinates of ground control point

The flight path in map space should be a tangent to all circles with centers (u,v) and radii (g_i) as shown in Fig.2. Because the assumption of the flight line to be a straight line cannot strictly be maintained, we have to iteratively calculate a best possible solution. Therefore a map ground range (a_i) is calcultated from the assumed flight position for each passpoint as follows:

$$a_i = \frac{\tan\phi \times u_i - v_i + v_O - \tan\phi \times u_O}{\sqrt{\tan^2\phi + 1}} \qquad (2)$$

a_i = ground range (map based)
ϕ = heading
u_O, v_O = intercept
u_i, v_i = map coordinates of ground control point

The best possible positioning of the flight path is obtained through minimizing the RMS difference between the calculated ground ranges in map and in image space. Initial values for flying height, heading and intercept (origin of flight path) are thus corrected.

In a second step the flying speed is determined with the linear function:

$$\text{line\#} = f(d) \qquad (3)$$

d = distance along flight path (see Fig.2)

Through a backward transformation, taking into account the flight path parameters and the speed function, every output pixel (in the elevation model) is assigned to an image pixel. Subsequent geocoding is performed using a nearest neighbour algorithm. Thus the data is corrected for panoramic distortions as well as for errors due to the topography.

First accuracy tests in overlaying a digitized street network onto the image showed, that due to crosswind during the flight, a yaw correction has also to be considered. The correction algorithms were therefore modified, implementing now a constant yaw factor. The improvement is demonstrated in Fig.3.

Concurrent to the geometric correction, radiometric errors should also be reduced. Radiometric correction algorithms have to take into account the precise local incidence angle. This possibility is lost after the geometric corrections (Woodham, 1985). Teillet (1985) states three conditions for a precise radiometric correction:

- the data should be radiometrically calibrated
- an atmospheric model should be available
- models should exist for the reflective characteristics of all objects

In this test, the first two conditions had to be neglected because the information was missing. Therefore, topographically induced illumination variations only were corrected. We applied a cosine correction to the data according to aspect and slope factors from the digital terrain model. This procedure is recommended whenever atmospheric and object specific correction factors are missing.

An atmospheric model as well as path lenght corrections had been tested in a previous campaign (Staenz et al, 1987).

The combination of the knowledge of the flight path with the digital terrain model allows for the calculation of zones in shadows. We have to differenciate between three types of shadows: sun shadows, sensor shadows and object shadows. We recommend to mask out all zones of shadows for separate treatment in classification. Areas in sensor shadows of course cannot be analyzed at all.

3. RESULTS

Multispectral digital scanner data can be considered geometrically corrected when they have been transformed into a defined ground coordination system with a controlled error tolerance. Though there are many ways of describing the precision of such transformations, as for instance "mean locational errors of passpoint pairs" (see Tab.1), these measures allow only a local error assessment. To be able to get a feeling on the remaining errors, a digitized street network was overlaid onto the data. This allowed for a visual control of the fit of the corrections.

A comparison with an overall correction using just an affine transformation, showed drastically the superiority of the chosen approach.

The radiometric correction that we performed is only a first step. However it may be very difficult, if not impossible, to correct for each tree types' individual reflective characteristic.

The final classification step was up to now only tested in our second "flat terrain" testsite. It revealed, that given the ground resolution cell of approx. 3.5 m and the mixture of the tree types within a stand, only very rough separations of major stands could be achieved as shown in Fig.4.

A heavy influence of the object-specific shadow was noted, which necessitated masking and restriction of the analysis to the center strip of the image data.

The classification was performed using a hierarchical binary decision tree, which reduced the decision rules to density slicing (DS) of various synthetic variables. These were obtained in part after pretreatment of the data with "edge preserving smoothing" (EPS) and region growing (IS), (Meyer, 1988). The variables used are shown in Fig.4.

All image preprocessing and processing was performed on a DIPIX Aries-II image analysis system. Special software was developed at the Remote Sensing Laboratories.

4. CONCLUSION

The described method of simulation of the flight path necessary for the geometric correction is feasible only with rather short flight legs of approx. 3 to 4 km. It is also only recommended when the flying was done under calm atmospheric conditions.

For any application of airborne multispectral scanners in hilly terrain, where a minimum of geometric reliability has to be reached, high precision INS data as well as roll and yaw compensation elements have to be recorded simultaneously to the multispectral data gathering.

It has been shown how airborne multispectral scanner data can be geometrically and radiometrically corrected for an application in hilly terrain. In the parallel test of the applicability of such data with even higher resolution to forest damage assessment however, only rather broad forest stand classifications were achieved. This may lead to the conclusion, that only in monoculture forest stands, damage assessment through multispectral data analysis can be successful.

5. REFERENCES

1. DIPIX SYSTEMS LTD, ARIES SYSTEMS Users Manual, Part I, II, III, Aries Version 4.0, Canada, 1989.
2. Guindon, B., "Integration of MSS and SAR data of forested regions in mountainous terrain", Proceedings of the 14(th) International Symposium on Remote Sensing, Winnipeg, pp 1673-1690, 1980.
3. Jansa, J., "Rektifizierung von Multispektral-Scanneraufnahmen", Geowissenschaftliche Mitteilungen, Technical University of Vienna, No.24, Vienna, 1983.
4. Kadro, A. & Kunz, S., "Ergebnisse computergestützter Waldschadenerhebungen mit Multispektralscannerdaten", ISP-IUFRO-Symposium, Zurich, pp 179-182, 1985.
5. Meyer, P., DIPIX-Memo No.20, Dept. of Geography, University of Zurich, 1988.
6. Rose, A., "Entzerrung von Scannerbildern mit Prädiktionsansätzen", Verlag der Bayerischen Akademie der Wissenschaften, Ph.D., Reihe C, No.303, Munich, 1984.
7. Staenz, K. & Meyer, P. & Itten, K. I., "Viewing angle corrections of airborne multispectral scanner data acquired over forested surfaces", Proceedings of IGARSS'86 Symposium, Zurich, pp 671-676, 1986.
8. Teillet, P.M. & Guindon, B. & Goodenough, D.G. , "On the Slope-Aspect Correction of Multispectral Scanner Data", Canadian Journal of Remote Sensing, Vol.8, No.2, pp 84-106, 1982.
9. Teillet, P.M., "Réflexions sur la Correction Radiométrique des Images en Télédétection", Proceedings of the Fifth Congress of the "Association Québecoise de Télédétection", Chicoutimi, Québec, p.25, 1985.
10. Woodham, R.J. & Lee, T.K. , "Photometric method for radiometric correction of multispectral scanner data", Canadian Journal of Remote Sensing, Vol.11, No.2, pp 132-161, 1985.

standard deviation	x-direction y-direction	± 2.56 pixel ± 2.42
rms	x-direction y-direction	0.587 0.556
r^2	x-direction y-direction	0.9998 0.9999

Tab. 1: Statistical results of geocoding accuracy

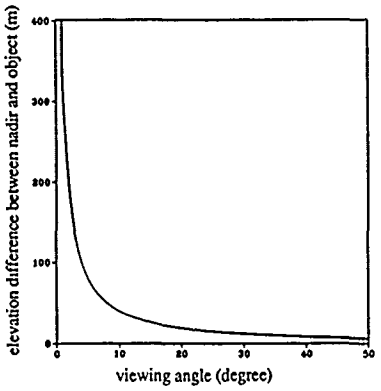


Fig.1 : Limits of 1 pixel displacement due to topography and viewing angle (pixelsize 6.5m)

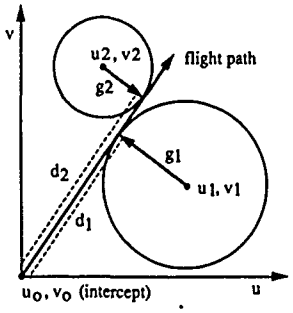


Fig.2 : Flight path in map space

Fig.3 : a) Result of geometric correction without yaw factor

b) Result of geometric correction including a yaw factor

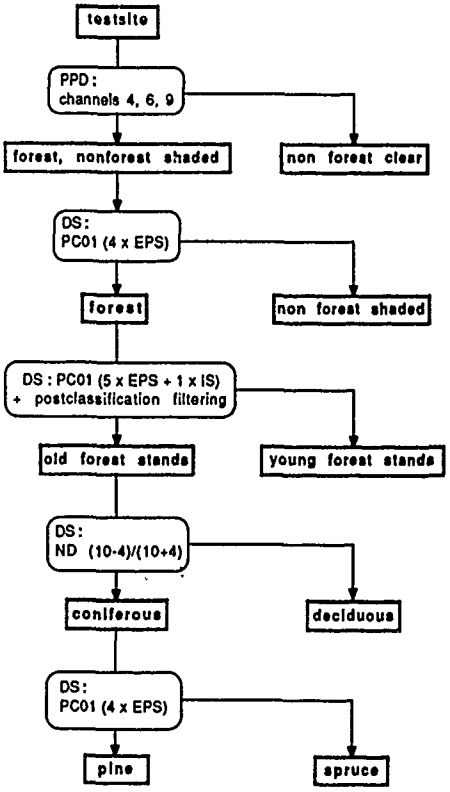


Fig.4 : Major forest stand classification through hierarchical separation

PPD	parallelepiped-classification
DS	density slicing
EPS	edge-preserving smoothing filtering
IS	image segmentation
PC01	first principal component
ND	normalized difference

CLASSIFICATION OF FOREST TYPES BY MICROWAVE SCATTEROMETER

Martti Hallikainen, Teemu Tares, Juha Hyyppä, Erkki Somersalo

Helsinki University of Technology
Laboratory of Space Technology
Otakaari 5 A, 02150 Espoo, Finland

ABSTRACT

Helicopter-borne scatterometer measurements of vario forest and surface types were conducted in the spring 1988 in Espoo, near Helsinki. The HUTSCAT (Helsinki University of Technology Scatterometer) operates at 5.4 GHz and 9.8 GHz and employs four linear polarization modes (VV, HH, VH, and HV). At each of the 8 channels, the radar can measure the backscattering properties of a target with a range resolution of one meter. The following aspects of forest remote sensing were investigated: (a) the location of backscattering sources for various tree species, (b) determination of tree height, and (c) identification of tree species. Identification of tree species was examined using both the properties of the radar return vs. range spectrum and the principal component analysis.

Key words: Radar, Forest, Tree height, Tree species identification.

1. INTRODUCTION

Ground-based scatterometer measurements of single trees have provided useful information on the effects of tree parameters to the backscttering behavior (Sieber, 1985; Zoughi et al., 1986). Ground-based measurements of single trees cannot, however, give detailed statistical information on the backscattering properties of tree species without getting extremely laborous. Airborne radar studies are needed in order to better understand the interaction of microwaves with forest canopies. The distribution of backscattering sources within the forest canopy is employed in theoretical backscatter modeling (Paris, 1986).

Helicopter-borne scatterometers can be used to collect data on both single trees and forest canopies. So far, the value of experimental results is limited by the number of channels (frequency and/or polarization) available. (Hoekman, 1987; Bernard et al., 1987; Pitts et al., 1988).

In this paper, we summarize the first results from forest measurements with an advanced helicopter-borne 8-channel ranging FFT scatterometer, HUTSCAT.

2. INSTRUMENTATION

Our new dual-frequency helicopter-borne scatterometer is a frequency-modulated continuous wave (FM-CW) radar that operates at 5.4 GHz and 9.8 GHz. The HUTSCAT employs four linear polarization modes (VV, HH, VH, and HV) at both frequencies. The 8-channel system can measure the backscattering properties of a target with a range resolution of one meter. The ranging capability for each channel simultaneously is obtained by performing the Fast Fourier Transform (FFT) in real time to the received time-domain signal.

The forest measurements described in this paper were made with the test version of the radar. The final version will be completed in May 1989. A detailed description of the scatterometer is given in another IGARSS'89 paper (Hallikainen et al., 1989).

The main difference between the test version and the final version of the radar is that the test version can collect data using only one channel at a time, whereas the final version collects data using eight channels simultaneously. In the test version, the incidence angle is fixed to 23 degrees off nadir, thus being identical to that of the ESA ERS-1 SAR to be launched in 1990.

The antenna 3 dB (two-way) beamwidth is 3.8 degrees at both 5.4 and 9.8 GHz. Since the incidence angle is 23 degrees, the antenna footprint on the ground (main axis of the ellipse) is 0.076*H and the slant range depth is 0.031*H, where H is the flight altitude. For an altitude of 50 meters (used in the experiment) the antenna footprint is 3.9 meters and the slant range depth is 1.5 meters.

3. DESCRIPTION OF EXPERIMENT

The measurements described in this paper were conducted on 29 April 1988 in Espoo, near Helsinki. Air temperature was about 6°C at 11 a.m. when the measurements were started and about 10°C (forest) to 15°C (open areas) at 2 p.m. when the measurements were completed. There was practically no wind and no clouds. No precipitation had been reported in the area since 25 April. However, snow cover had melted only one week before the measurements and the soil was partially frozen in forested areas of the test site. Due to recent snow melt run-off, the volumetric water content of the soil was high.

Description of the five target categories is given in Table I. Three forest categories were measured; mature spruce, young spruce, and a mixture of aspen and birch. Additional targets included logged land (clear-cut area) and agricultural land (plowed bare field). According to the present practice in Finland, the logged area had been plowed after harvesting. At the time of the measurement, aspen and birch were defoliated.

Table I. Characteristics of the target categories, Espoo, 29 April 1988.

Category	Tree Height (m)	Understory	Soil State	Soil Water Content (cm^3/cm^3)	Soil Temperature Below Surface (°C)				
					1cm	3cm	6cm	10cm	20cm
Mature Spruce	20 ± 5	None	Partly frozen, no snow	0.54	1.0	0.6	0.4	0.0	0.0
Young Spruce	8 ± 2	Grass, Hay	Partly frozen, no snow	0.47	1.7	1.2	1.1	1.1	1.1
Aspen and Birch (Defoliated)	14 ± 3	Moss	Partly frozen, pools	0.60	9.9	4.4	3.0	2.0	1.1
Logged Land	-	Plowed after harvesting	Thawed	0.35	-	-	-	5.0	3.8
Agricultural Land	-	Plowed field	Thawed	0.40	16.0	11.0	8.5	6.2	3.2

The backscattering properties of each target were measured at 5.4 and 9.8 GHz, employing four linear polarization modes (HH, VV, HV, VH) at both frequencies. The incidence angle was 23 degrees off nadir. The range resolution was one meter. Consequtive backscattering spectra were obtained in 0.5 meter intervals along the flight track.

4. EXPERIMENTAL RESULTS

4.1. Determination of Tree Height

Figure 1 depicts consequtive radar return spectra for mature spruce at 5.4 GHz, VV polarization. The backscatter power peaks at the minimum and maximum distances from radar correspond to backscatter from tree tops and the ground, respectively. The tree height along the flight track can be determined simply by taking the difference between the two distances and by making the necessary correction due to the 23^0 incidence angle.

Figure 2 shows a sample of the radar-derived tree heights in the test site, obtained by using the 5.4 GHz VV channel. The measured minimum tree heights are due to small open areas between the trees and the true tree height varies from 12 to 29 meters. The radar-derived tree heights were confirmed in an approximate manner by the ground truth.

In general, the like-polarization gives a clear radar return from both the tree tops and the ground, thus being suitable for tree height measurements. The cross-polarized radar return is weaker, as discussed in Section 4.2., and does not provide good capability to measure the tree height. Based on the present data set, both C-band and X-band can be used for these measurements.

4.2. Identification of Backscattering Sources in Forest Canopy

Figure 3 shows the backscattering sources for mature spruce and the mixture of aspen and birch at 5.4 GHz, VV polarization. The intensity of backscatter is coded in grey, with the lightest areas representing the highest intensity. Figure 3 indicates that the dominant backscattering source for mature spruce are the tree tops. In some cases, the attenuation of the spruce canopy is so high that no return from the ground is obtained. The results for aspen and birch indicate that most of the backscatter comes from the ground. At the time of the data taking, deciduous trees were defoliated.

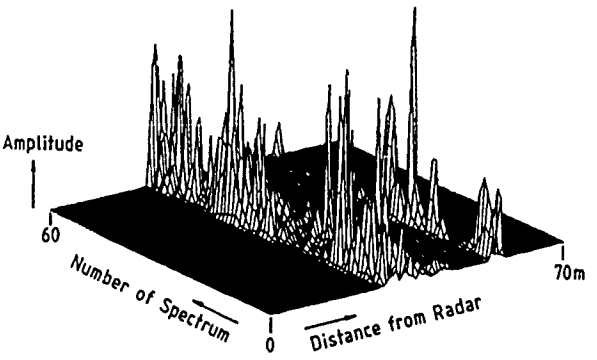


Figure 1. Consequtive radar return spectra for mature spruce along the flight track; 5.4 GHz, VV polarization.

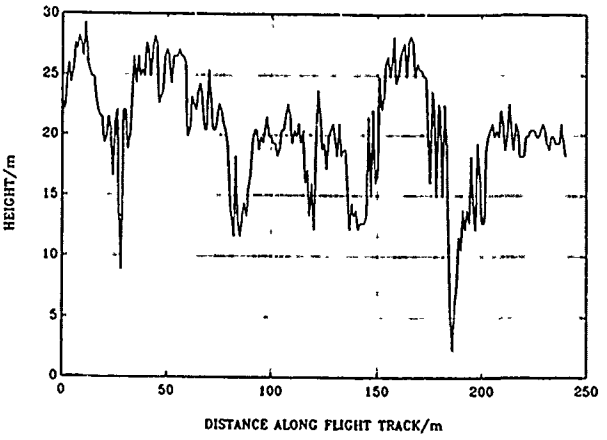


Figure 2. Radar-derived tree height along the flight track for mature spruce; 5.4 GHz, VV polarization.

Figure 4 shows the average radar return spectra for mature spruce and the mixture of aspen and birch at 5.4 and 9.8 GHz for HH, VV, and HV polarization modes. The difference in the location of the main backscattering source for the two tree species is clearly visible. The results were obtained by averaging 60 consequtive spectra.

4.3. Identification of Tree Species

The experimental backscattering coefficient for each of the forest and surface types was around -6 to -12 dB at 5.4 GHz, VV polarization. The other channels did not appear to provide a better capability to discriminate between the five categories used in the experiment. Hence, other methods were examined in order to increase the discimination capability.

The experimental results in Figure 4 indicate that the shape of the radar return vs. range (return spectrum) depends on the tree species. The feasibility of using the properties of the return spectra to discriminate between various forest types was investigated by using standard multivariate statistical methods. The following characteristics of each spectrum can be computed: (a) norm, (b) center of mass, (c) skewness (describes the

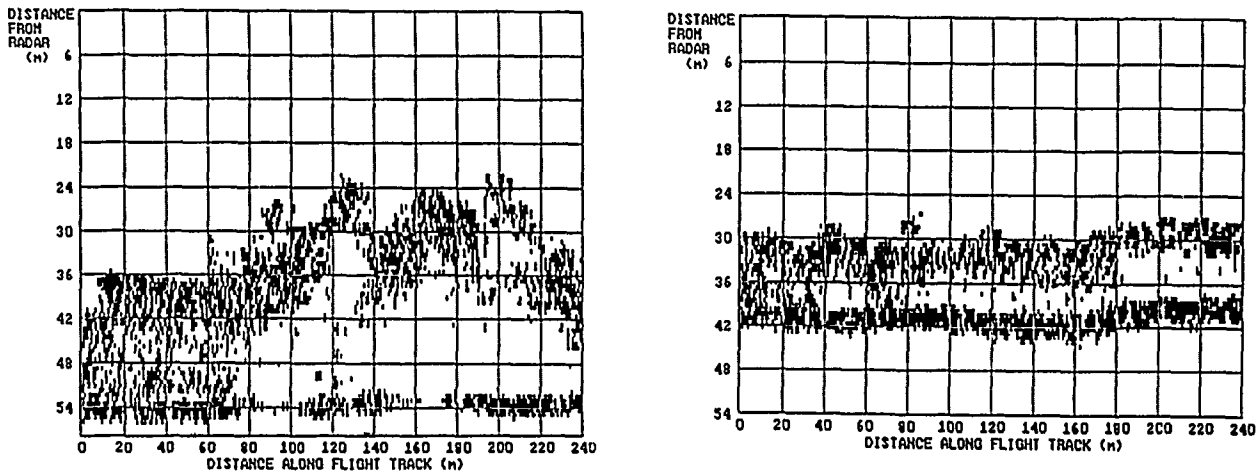


Figure 3. Backscattering source distribution for mature spruce (left) and aspen and birch (right) along the flight track; 5.4 GHz, VV polarization. Backscatter amplitude is coded by grey scale.

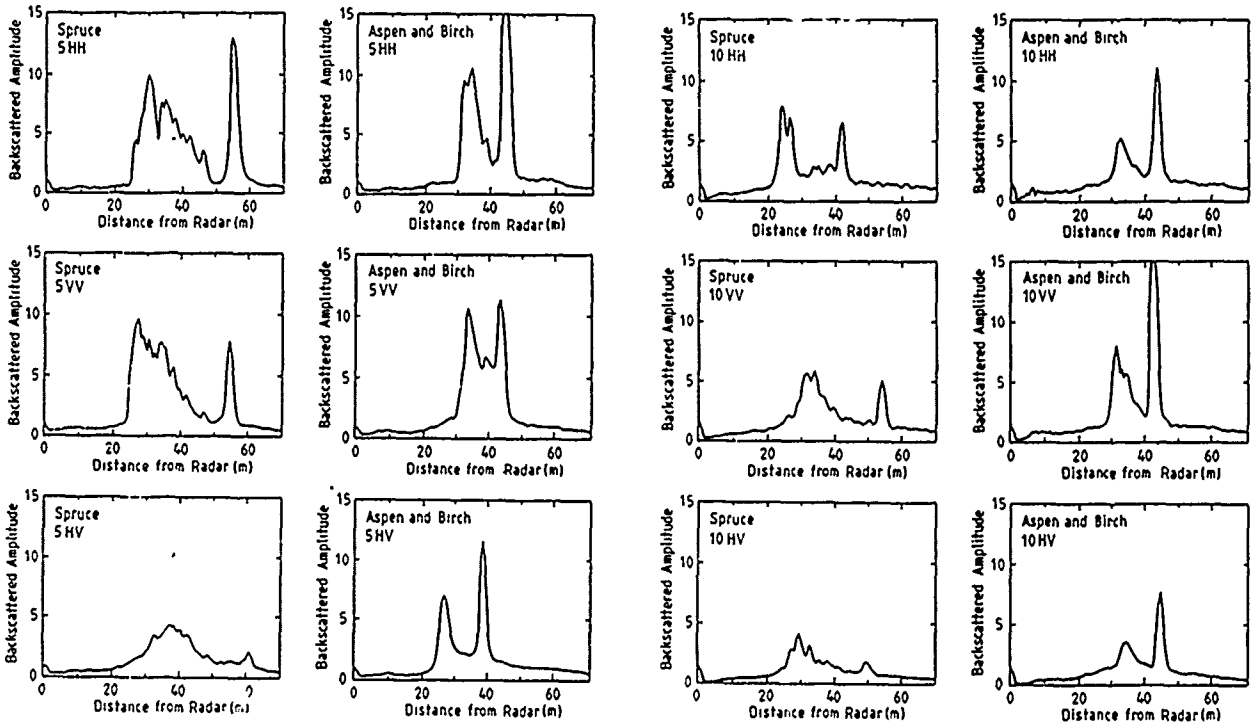


Figure 4. Average shape of radar return vs. range for mature spruce and aspen and birch at 5.4 and 9.8 GHz, for HH, VV, and HV polarization modes. 60 consequtice spectra were averaged.

symmetry of spectrum around the center of mass), and (d) spread (desribes the depth of the target).

Additionally, the principal component analysis can be applied to the experimental results. The above statistical quantities and methods are briefly reviewed in another IGARSS'89 paper (Hallikainen et al., 1989); for a more detailed treatment see e.g. (Wilks, 1963).

The application of the principal component analysis to the raw data from the FM-CW radar is not practical, since the return spectra obtained from different flight altitudes would produce several separate subgroups. Therefore, the spectra are first scaled in an appropriate way to preserve only the shape of the spectrum.

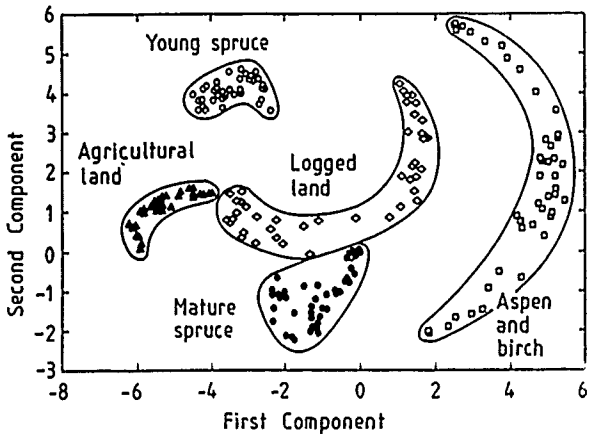


Figure 5. Tree species discrimination by the principal component analysis; 5.4 GHz, VV polarization. See text for description of data points.

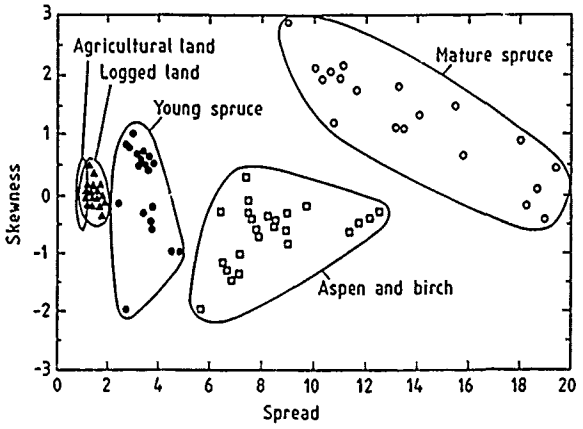


Figure 6. Spread versus skewness for averaged radar return vs. range spectra; 5.4 GHz, VV polarization. See text for description of data points.

From each target, 15 successive spectra are first averaged. Next, the scaling explained above is performed and the data matrix (each column corresponds to a single measurement) is formed with the discrete sample vectors of the scaled spectra. From each target, 45 averaged scaled spectra are included. Finally, the two first principal components are calculated and the data is projected to these directions. The scatter plots obtained for the first and the second component in Figure 5 show that the data clusters to five different groups corresponding to the original categories. The results in Figure 5 strongly suggest that identification of tree species is possible.

In Figure 5 the points corresponding the spectra of aspen and birch and of logged land are clustered in an arc-shaped configuration. This phenomenon is sometimes referred to as the "horseshoe effect" and it is known to be a mathematical artifact.

As explained above, the scaling of the spectra retains only the information that is contained in the shape of the spectra. Figure 6 shows a plot of the spread vs. skewness of each of the averaged spectra. The results in Figure 6 indicate that overlap exists only between the categories of agricultural land and logged land.

5. CONCLUSIONS

The results obtained with the new 8-channel scatterometer indicate that the capability of the radar to instantaneously measure the backscattered power vs. range is valuable for remote sensing of forests. The tree height along the flight track and the location of the backscattering sources within the forest canopy can be determined.

Using the principal component method, the radar return vs. range spectrum can be used to identify tree species. The data have to be scaled in an appropriate way to eliminate the effect of the flight altitude. Additional measurements are needed to investigate the feasibility of the method for other seasons.

6. REFERENCES

1. Bernard, R., Fr 'al, M.E., and Vidal-Madjar, D., "Nadir-looking 'borne radar and possible applications to 'estry". Remote Sensing of Environment, Vol 1, pp. 297-309, 1987.

2. Hallikainen, M.T. et al., "A helicopter-borne 8-channel FFT scatterometer", 1989 IEEE Int. Geoscience and Remote Sensing Symposium (IGARSS'89), Vancouver, 10-14 July 1989.

3. Hoekman, D.H., "Measurements of the backscatter and attenuation properties of forest stands at X-, C-, and L-band". Remote Sensing of Environment, Vol. 23, pp. 397-416, 1987.

4. Paris, J.F., "Probing thick vegetation canopies with a field microwave scatterometer". IEEE Trans. Geoscience and Remote Sensing, Vol. GE-24, pp. 886-893, 1986.

5. Pitts, D.E., Badhwar, G.D., Reyna, E., Ulaby, F.T., Brunfeldt, D.R., "The use of a helicopter mounted ranging scatterometer for estimation of extinction and backscattering properties of forest canopies, Part I: Experimental approach and calibration". IEEE Trans. Geoscience and Remote Sensing, Vol. GE-26, pp. 140-143, 1988.

6. Sieber, A.J., "Forest signatures in imaging and non-imaging microwave scatterometer data", ESA Journal, Vol. 9, pp. 431-447, 1985.

7. Wilks, S., "Mathematical Statistics", 2nd edition, New York: John Wiley and Sons, 1963.

8. Zoughi, A., Wu, L.K., and Moore, R.K., "Identification of major backscattering sources in trees and shrubs at 10 GHz", Remote Sensing of Environment, Vol. 19, pp. 269-290, 1986.

INTEGRATION OF LASER RANGEFINDER AND MULTISPECTRAL VIDEO DATA FOR FOREST MEASUREMENTS

D.H. Currie, V.L. Shaw and F.G. Bercha

The Bercha Group
Suite 250, 1220 Kensington Road N.W.
Calgary, Alberta, T2N 3P5
Phone: (403) 270-2221; Telefax (403) 270-2014; Telex: 03-827666

ABSTRACT

Attempts to improve the accuracy and timeliness of forest measurements have led to the application of new sensor technologies which collect high resolution digital data sets from an airborne configuration. Two devices which have shown particular promise are the laser rangefinder and the multispectral video camera. These instruments are capable of measuring a number of primary forest parameters, including tree height, accurately and economically. Used in an integrated sensor package, they can generate three dimensional views of the forest body and underlying terrain; however, the data sets collected are large, with a high level of complexity.

The objective of the work described herein was to develop a system for extracting forest information from simultaneously acquired coincident laser and video data. Test data sets were acquired over an area in west central Alberta and consisted of multiple low level flight lines over a variety of forest cover types. An integrated sensor package was used, consisting of a high frequency laser rangefinder and a six band multispectral video camera. Extensive ground data were compiled and supplemented with historical data from permanent sample plots maintained by the Alberta Forest Service. This information was used to develop and test a system for filtering and correlating the remotely sensed data in a semi-automatic manner. Analysis of the information derived showed a high degree of accuracy in the measurement of tree heights and stand densities, while species, crown diameters, and vigor could be estimated. Current research is aimed at the generalization of the software system to provide support for a wide variety of forest cover types as well as a direct and logical interface to forestry GIS systems.

KEYWORDS: Laser, video, forestry.

INTRODUCTION

The purpose of the work described herein was to develop an airborne data collection and analysis facility which would be useful for forest assessment, either as a stand-alone system or in combination with other airborne and ground based data acquisition systems. Previous investigations by the authors and others have demonstrated the utility of the laser rangefinder for direct measurements of tree heights in an airborne configuration. By pairing this unique sensor with a second one, capable of obtaining images of the forest canopy, and analyzing the acquired data as an integrated set, a system which could generate a three dimensional representation of the forest canopy in a digital format was predicated.

Prior to this work, the laser systems were often operated in conjunction with a video camera in order to recover the aircraft flight line in post mission processing. These two sensors are complementary in the sense that their operational envelopes overlap (flying height, speed, etc). Additionally, they offer two very different data sets since the laser measures in the vertical plane, while the video camera produces a horizontal image. Unfortunately, conventional video imagery is not easily interpreted using the analytical techniques which are required for management of the large quantities of data produced by these systems. By using a multispectral video camera, the same image data may be broken down into several discrete images representing the scene reflectance at a variety of wavelengths. These images can be captured using a video digitizer and analyzed in a manner analogous to the digital analysis of Landsat imagery.

Testing of the integrated system was performed in west-central Alberta, near the town of Drayton Valley. This area offers a variety of forest cover types in single-species and mixed stands located on rolling terrain. Operations consisted of several flights over pre-selected areas using the laser-video system, followed by a ground truth program aimed at validating the forest parameters measured by the airborne system.

LASER RANGEFINDER DATA

Reduction of the laser data is a two step process. The first requirement is to determine the range from the aircraft to the ground. This involves filtering out the ranges which are intercepted by the foliage so that the remaining range measurements indicate the ground surface.

These ground ranges are considered an intermediate result for scaling the video images.

Once the ground trace is defined, the second step is to subtract the short ranges to give the tree heights. A considerable amount of research has been devoted to determining how well the laser actually detects the top of the tree crown (Aldred 1985). The factors which come into play here are the sensitivity of the laser to the small branches at the top of deciduous trees and the difficulty in determining whether or not the laser detected the highest point on the crown. Both of these effects will cause a shorter tree height than is actually the case to be measured. Studies with the laser used in the present application have shown that sensitivity to small objects is not a problem as a fairly wide beam dispersion is used and the ranging electronics provide excellent control over the target selection. Thus, for deciduous trees with large, flat-topped crowns, the height measured will be fairly accurate. Another consideration is that the laser is not maintained in a constant vertical orientation and will therefore produce height measurements which are slightly longer than the actual tree size. The net effect has been found to be that the raw range data is within 0.5 meters of the control data when averaged over small segments of the flight line.

An important aspect of the laser cross-section is its display of the vertical foliage distribution along the profile as shown in figure 1. This output is useful for discriminating between tree species during image interpretation and can also be used to locate defoliated crowns which may be due to disease or insect attack.

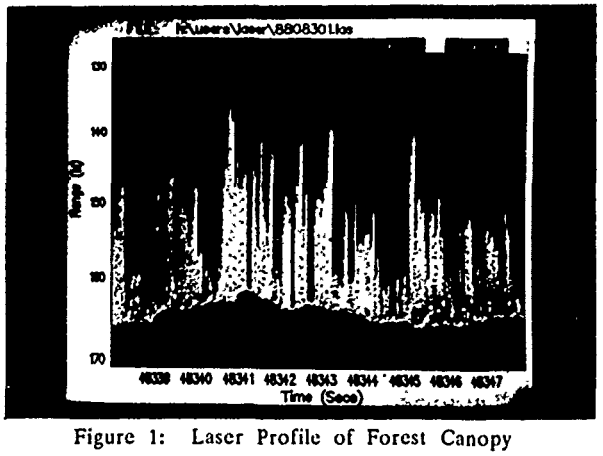


Figure 1: Laser Profile of Forest Canopy

MULTISPECTRAL VIDEO IMAGERY

The multispectral video camera used for this application was a Xybion model MSC-02. This system utilizes a charge coupled device (CCD) sensor mounted behind a rotating disc containing six narrow band spectral filters (Frost 1985). The disc rotates at a speed of sixty RPM and is synchronized with the scan frequency of the video sensor. Thus each successive field of the video imagery is obtained through a different filter. A complete set of six images is recorded ten times per second. The filter disc also acts as a shutter to reduce image blur due to aircraft motion. This sensor configuration has a number of advantages over other technologies in that the successive images have the same attributes since they are acquired by the same sensor-lens combination. The individual bands are easy to separate because they are recorded at different epochs. On the other hand, the nature of the imagery makes it nearly impossible to analyze without a digital frame capture system as the live video flickers rapidly through the six spectral bands. The serial nature of the data acquisition also implies that the images must be registered to each other to eliminate the effects of aircraft motion.

The video imagery is captured using a video "frame grabber" board interfaced to a digital image analysis system. This system utilizes high speed analog to digital converters (ADC's) to convert the incoming video signal to a digital raster. Consecutive fields are captured and registered to form a multiband image (Fig. 2). The time codes stored in the video signal are also captured and used to annotate the images for later correlation with the laser data. Each multispectral video image is cropped and then radiometrically enhanced prior to classification so as to equalize the relative brightness of each band. These procedures do not lead to any loss of data.

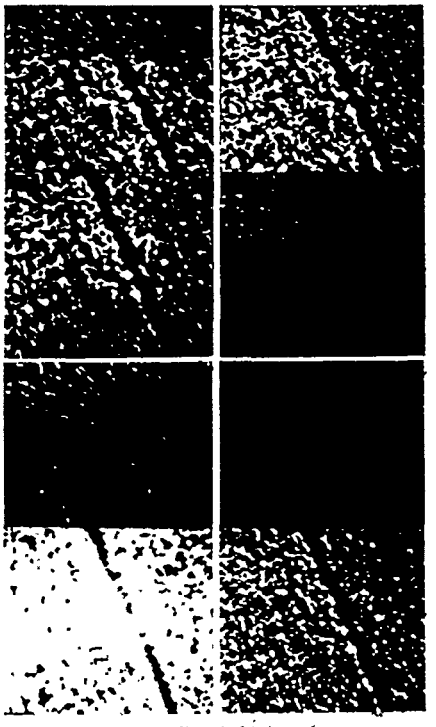


Figure 2: Multi Band Video Imagery

The goal of the analysis phase is to obtain a general classification of the image by separating tree crowns from open ground (Fig.3). In mixed stands, separation of the deciduous species from the coniferous species is sometimes possible but, due to the limited bandwidth of

the video sensor, identification of the individual species is unreliable. One of the main factors affecting the image quality is the amount of solar illumination. General classifications are first attempted using an unsupervised clustering technique. If this approach fails, a supervised approach is used. At this point the laser data can be used to advantage. Since different tree species present different vertical profiles in terms of leaf and branch distributions, the analyst can identify the various types by correlating the laser data to the video image using the synchronized time codes embedded in the two data sets. This generally leads to a successful classification.

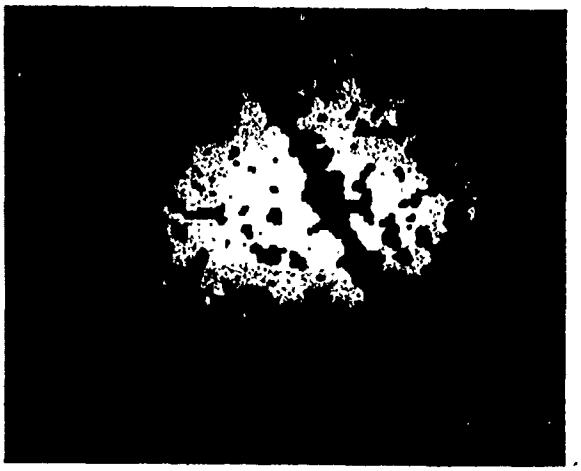


Figure 3: Thematic Overlay Showing % Crown Cover

GROUND DATA

Two methods were used to gather control data for this study. The first method was designed to establish the relationship among age, height, and diameter at breast height (dbh) of several species of trees in the study area. To do this 691 trees were sampled at random and classed by height and species group so as to obtain a significant sample of each species population and to indicate the distribution of age and dbh relative to height. For each species group, five height classes were established and thirty trees falling in each group were sampled. The data collected in this manner gave a good indication of the growth rates and relative distributions of the various tree species in the area.

The second ground data acquisition procedure was designed to specifically test the laser system accuracy. Several flightlines were accurately marked on air photos of the area at a scale of 1:30,000. At well defined sections of these lines, a "strip cruise" was performed covering an area of ten meters on each side of the line. In the area covered by the strip, each tree was measured for dbh and its species recorded. At ten meter intervals along the strip, the height of a dominant tree was measured. This data, augmented by the data from the first method, gives a clear indication of the stem density, species distribution, and canopy height at a location which was also viewed by the airborne system.

ANALYSIS AND RESULTS

The products of the airborne data interpretation were tabulated to show average stand height, crown closure, and the ratio of deciduous to coniferous species along the strip cruise areas. These parameters were compared to the measurements acquired during the ground truth data acquisition phase.

The average tree heights measured over the profile lines were within 0.5 meters of the ground values. This repeats results obtained in previous investigations. The determination of crown closure was better in single species coniferous stands than in mixed or deciduous areas, mainly because of the difficulty in differentiating the lowlying ground foliage from the deciduous crowns. Better results in these stands were achieved using the integrated data sets than with the video imagery alone. Similar problems were encountered in determining the ratio of species occurrence.

CONCLUSIONS

The processing of data obtained from an airborne laser rangefinder integrated with a multispectral video camera was described. The results of flight tests of this system were compared with data obtained by field cruising techniques. The sensor package described has the advantages of relatively low cost while providing a direct measurement of tree height and thematic inventory data which can be easily processed by semi-automatic means.

ACKNOWLEDGEMENTS

Cost sharing with the Bercha Group for the work described above by the National Research Council and the Western Diversification Program of Canada is gratefully acknowledged.

The work described was directed by Valerie Shaw. Her untimely death, on February 23, 1989, prevented her from participating in the preparation of this paper.

REFERENCES

1. Aldred, A.H.; Bonnor, G.M. 1985. Application of airborne lasers to forest surveys. Information Report PI-X-51, Can. For. Serv. Petawawa, Ont.

2. Frost, P.A. 1985. A multispectral video imaging and analysis system Xybion Elec. Systems Corp. Cedar Knolls, NJ

3. MacLean, G.A.; Krabill, W.B. 1986. Gross merchantable timber volume estimation using an airborne Lidar system. Can. J. of Rem. Sens. 12:1 pp 7-18

4. Miesner, D.E. 1985. Fundamentals of airborne video remote sensing. Proc. of X Biennial Wkshp. on Color Aerial Photogr. in the Plant Sci. Ann Arbor, Mich.

5. Nelson, R.; Krabill, W.; MacLean, G. 1984. Determining forest canopy characteristics using airborne laser data. Rem. Sens. of Env. 15:201-212

Application of Aerial Multispectral Videography and Colour/Colour IR Photography in Sugar Maple Decline Assessment

X. Yuan and J. Vlcek

Faculty of Forestry
University of Toronto
Toronto, Ontario M5S 1A1

D. King

Dept. Of Civil Technology
Ryerson Polytechnical Institute
350 Victoria Street
Toronto, Ontario M5B 2K3

and

D. McLaughlin

Ontario Ministry of Environment
Phytotoxicology Section
880 Bay Street
Toronto, Ontario M5S 1Z8

ABSTRACT

This paper presents an application of aerial multispectral videography and colour/colour IR photography in sugar maple decline assessment. Videography is a low cost alternative to multispectral scanning and photography. In this study, large scale 4-band video of several test plots was acquired using the 4-camera video system developed in the Faculty of Forestry, University of Toronto. In the video data, various spectral and textural measures for the sampled trees within and around several test plots were computed. They included principal components, band ratios, and first-order and second-order texture transforms. These measures were evaluated and compared with the results from photo interpretation and ground surveys. The most significant measures were then selected to form a numerical maple decline index. Colour and colour IR photography was acquired simultaneously with the video at 1:2,000 scale. Trees in each test plot were identified and interpreted for decline level based on: leaf chlorosis, evidence of visible branches, and general crown texture. A decline index derived from photo interpretation was also developed primarily based on leaf chlorosis characteristics. The analysis results so far have shown that the aerial photography method is subjective but internally quite accurate and that videography is an objective, quantitative and inexpensive alternative for maple decline assessment.

KEY WORDS : MULTISPECTRAL AERIAL VIDEO IMAGING, COLOUR/COLOUR IR PHOTOGRAPHY, MAPLE DECLINE, AND SPECTRAL AND TEXTURE ANALYSIS.

INTRODUCTION

Aerial colour and colour infrared photography has long been used for forest inventory purposes. Multispectral video imaging, as a newly developed remote sensing technique, has been shown to have a great potential in many forest applications using both analogue and digital image analysis methods (Vlcek et al. 1986; Yuan et al. 1987).

Sugar maple (Acer saccharum) decline has become a severe problem in northeastern American forest regions in recent years due to the interaction of several biotic and abiotic causes or stress factors, such as acid rain, soil nutrient deficiencies, increased ozone pollution, etc. Providing timely and accurate assessment of the level and extent of damage to this economically important tree species is a primary goal of remote sensing in forest resource management. Advantages of using remote sensing techniques to detect and assess maple decline are obvious considering the general decline symptoms. Basically, there are 5 stages in the development of decline: (1) early leaf discolouration followed by premature leaf fall, (2) progressive

deterioration of young twigs and branches of increasing sizes, (3) progressive dieback of buds, twigs, and branches from upper outermost parts of the crown, (4) decline of part or almost all of the leaf crown, and (5) death/recovery of affected trees. These changes are manifested in both crown spectral and textural characteristics providing opportunity for large scale imaging and individual tree assessment using spectral and textural analysis.

The main purpose of this study was to investigate and evaluate the capability of multispectral video imaging in comparison with aerial colour/colour IR photography for sugar maple decline assessment. Digital-video-based and photo-interpretation-based maple decline indices were developed for operational purpose.

DATA ACQUISITION

Aerial photography (colour and colour IR) and multispectral video of nine Ontario Ministry of Environment (MOE) test plots were acquired simultaneously on August 16, 1988 between 1130 and 1500 EST from an altitude of 610 m AGL.

Multispectral video was acquired using the 4-camera video sensor developed at the Faculty of Forestry, University of Toronto (Vlcek and King 1985). It incorporates four aligned black and white solid-state video cameras, a sequential switcher which provides frame-rate multiplexed video from all cameras that is recorded on a single VCR, and a colour encoder which provides colour/false colour imagery that is recorded on a second VCR. The cameras are RCA TC-2800 models equipped with Canon 2/3", 16 mm focal length TV lenses. Selected interference filters were: 430-470 nm, 530-570 nm, 665-675 nm and 780-820 nm. Digitization of the recorded black and white multiplexed video signal at 512 x 480 sampling rate was carried out using a Matrox PIP-1024a frame grabber capable of freezing four consecutive video frames. The resulting ground pixel sizes were : 0.48 m (H) x 0.38 m (V). Registration of individual video bands was then conducted to eliminate shifts due to aircraft translation in the 1/15 s data acquisition interval. A linear contrast enhancement was also applied to each image to fully utilize the whole dynamic 8-bit grey level range.

The photographic portion of the mission was carried out by the Ontario Centre for Remote Sensing (OCRS). The cameras utilized were Vinton 70 mm format cameras with 76 mm focal length lenses. The films employed were Kodak Aerocolour 2445 negative 70 mm film and Kodak Aerochrome 2443 IR 70 mm film. A 500nm cut-on filter was used on the IR camera.

The nine OME plots were located in representative parts of sugar maple woodlots in several regions of southern Ontario. Of the nine sites which were flown over, only seven were successfully covered by photography and four by video. The remaining sites were missed due to navigational problems. Field visits were made to each of the sites to identify plot locations on the aerial photographs and visually assess the decline status of each plot.

DATA ANALYSIS METHODOLOGY

Video Data Analysis Methodology

A number of spectral and textural measures were computed and evaluated with ground-based information and photo interpretation results. The measures which were most associated with decline symptoms were then selected for use in the decline model.

An image of windows of selected tree crowns for derivation of spectral and textural measures was first defined. Selection of window size was critical because too large a window would include some boundary shaded pixels which could increase variations of pixel values, while too small a window could not reflect the actual tree crown spectral and texture patterns. In this study, the window size was determined in accordance with the tree size. Experience revealed that an appropriate window size for a tree crown was one that included 70% - 80% of the total pixels falling on the crown.

For each window selected, spectral and texture measures were calculated as described below. For sites in which the OME test plot was not successfully covered, a representative test plot nearby was selected.

(a) Spectral Transformation. Three spectral measures were used, two principal components and one ratio component. Principal component transformation is a useful technique in multispectral image analysis to reduce data dimensionality by derivation of uncorrelated principal components formed by linear combinations of the original spectral variables. Since the first two components usually explain most of the variation, they were selected as two spectral measures. Another spectral measure was band ratios. They were selected in order to reduce radiometric variations caused by sun illumination angle and view angle variations.

Before carrying out spectral transformations, each subimage of tree crowns was standardized by subtracting the band mean and dividing by the standard deviation for scene comparison purposes. This was done to eliminate the effect caused by varied imaging formation conditions from site to site and varied image pre-processing that may have compressed, enlarged, or shifted grey level distributions.

(b) Texture Transformation. The texture representation methods used were first-order texture transformations, and the second-order co-occurrence method. One measure from the former was mean Euclidean distance (MED) and three measures from the latter were contrast (CON), entropy (ENT) and angular second moment (ASM). There are some advantages and

disadvantages in each method discussed by Irons and Pattersen (1981), Harlick (1979), Weszka et al. (1976), and, Conners and Harlow (1980).

A similar data normalization procedure to that described above was also applied before performing texture transformations. In addition, the resulting data were compressed into a 6-bit grey level range. This process did not change the actual grey level distributions but greatly reduced the size of the co-occurrence matrix from 256 x 256 to 64 x 64, and thus, greatly reduced computation costs. The intersample spacing for the co-occurrence matrix was 1 pixel and an average of 8 sampling angles were used to ensure rotationally invariant texture. All texture measures were derived from the band 2 image because it displayed the best image contrast and sharpness resulting in good visual appearance of different texture patterns.

(c) Evaluation of Spectral and Texture Measures. Evaluation of each spectral and texture variable in maple decline assessment was carried out through a statistical analysis of ground information, photo interpretation results and data collected from two of the test plots (STEED and MILLER). Since the average ground-based OME index of each plot was computed from a small number of trees that were often difficult to identify on the video images, most of the evaluation was done based on photo interpretation of identical trees sampled from both photographic and video imagery. Colour photo interpretation was carried out by interpreter 2 based on three variables: chlorosis (4 levels), visible branches (4 levels), and crown texture (4 levels). The sum of the three variables was used as a photo decline index (PDI). Correlations between this photo decline index (PDI) and the spectral and texture variables were calculated and the best variables were selected for the final maple decline index model.

(d) Formation of Maple Decline Index Model. The linear regression models obtained from the above analysis only showed the linear relationship between the selected variables and the photo decline index within a scene. For between scene comparisons, there were two possible approaches. The first was to calibrate video data for each scene so that an absolute comparison could be done. The second was to use a common reference object so that anything in the image could be compared to it, thus, the comparison would be relative. In this study, the latter approach was adopted. A healthy tree was selected from each site as the reference and a spectral and texture score were calculated. A linear model was proposed as follows.

$$VDI = K_1 + K_2 (S_t - S_r) + K_3 (T_t - T_r)$$

where: VDI : video decline index for a sampled tree
S_t : spectral score for the tree
T_t : texture score for the tree
S_r : spectral score for the reference tree
T_r : texture score for the reference tree
K_1, K_2, K_3 : constants.

The constants K_1, K_2, and K_3 were determined based on the photo decline index (PDI) so that both decline indices would have a similar dynamic range.

Aerial Photo (Colour and Colour IR) Interpretation

To develop an operational photo interpretation methodology using colour/colour IR photography, the keys shown in Table 1 were developed to assist in the interpretation process and assure consistency between interpreters.

Table 1. Keys to maple decline assessment for colour/colour IR photography.

Film type	Decline level	Colour	Colour density	Crown texture
Colour IR	0 - 1	deep red	compact, dense	crinkled
	1 - 2	pinkish red	dense	crinkled
	2 - 3	light pink	somewhat porous	crinkled
	3 - 4	pink-grey	sparse	fussy
	4 - 5	grey	transparent	smooth
Natural Colour	0 - 1	deep green	compact dense	crinkled
	1 - 2	light green	dense	crinkled
	2 - 3	yellow green	somewhat porous	crinkled
	3 - 4	muddy green	sparse	fussy
	4 - 5	brown	transparent	smooth with radial lines

Both natural colour (NC) and colour infrared (CIR) photos were interpreted using mirror stereoscopes with 3X magnification. Individual tree crowns within and around the OME test plots were delineated on transparent overlays which were then graphically enlarged for numbering of individual trees. Using the original overlay over one of the photos and the numbered enlargement for tree identification, the decline level of each tree was determined with the aid of the key. Decline was first interpreted on all NC photos and then repeated on CIR photos. Two interpreters worked independently on the same overlays. Interpretation was carried out for main plots which contained the OME research plots and, in several cases, for 1-3 sub-plots chosen in the vicinity of the main plot.

RESULTS AND DISCUSSION

Video Data Analysis

Two plots, STEED and MILLER, were selected for the statistical analysis. Fifty-seven and twenty-two trees were sampled from the two plots, respectively. Correlation analysis of colour photo interpretation results and video analysis results showed that the second principal component (PC2) and the contrast (CON) had the highest linear relationship with the photo decline index (PDI).

The linear regression analysis results are shown in Table 2 and Figure 1.

Table 2. Summary of regression analysis.

Site	Ind. Variable	Model	N	r^2	RMSE
STEED + MILLER	PC2	PDI=3.74+1.88 PC2	78	0.31	1.51
	CON	PDI=3.73+0.04 CON	78	0.62	1.12
	PC2+CON	PDI=2.93+1.16PC2+0.03CON	78	0.73	0.96

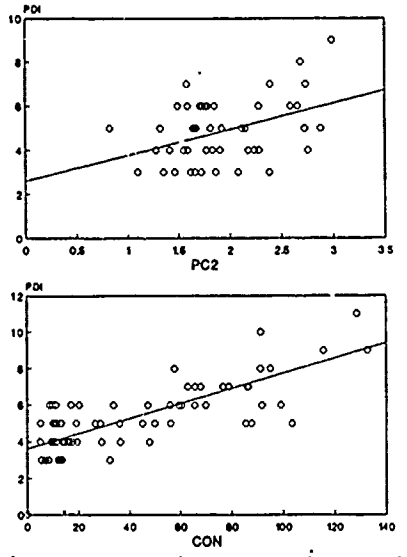


Fig.1. Linear regression analysis of video data spectral and texture variables with the photo decline index from STEED and MILLER sites (PDI=Photo Decline Index, PC2=Principal Component 2, and CON=Contrast, r^2=Determination Coefficient, n=No. of Trees, RMSE=Root Mean Square Error).

The linear correlation (r^2) of the photo decline index (PDI) with the second principal component (PC2) was 0.31. The PDI correlation with the contrast (CON) texture was 0.62. Even higher correlations were obtained when both PC2 and CON were combined in a two-variable linear model (r^2=0.73). The root mean square errors were also reduced from 1.51 (for PC2) and 1.12 (for CON) to 0.96 in the two-variable model.

Five trees, ranging from healthy to very severely declining, were selected from the STEED site (based on the ground information) for further detailed study. Average grey levels of each tree crown (7 x 7 pixel window) in four spectral bands are plotted in Figure 2. It is evident that the grey levels are higher in bands 1 and 2 for the declining trees than the healthy trees. In bands 3 and 4, the trend is not clear because the radiometric image quality in these two bands was not very good. Examination of the variances of the trees showed that declining trees usually had higher variance than healthy trees. These observations suggest that spectral and textural changes from healthy to declining should be considered as two basic

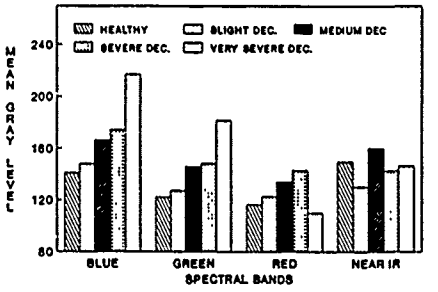


Fig. 2. Spectral comparison of five maple trees.

indicators of decline and therefore should be included in the proposed maple decline index model.

Average video decline index values for the 6 selected plots were calculated using the proposed model and ranked as shown in Table 3. It was found that the STEED site had the lowest VDI (5.2), the DAVIES site had the highest VDI (8.2), and the remaining sites were not significantly different. Overall, the trees in all the test plots showed medium decline.

Table 3. Ranking of the plots by VDI values.

Plot Name	No. of trees Sampled	Averaged VDI	Standard Deviation
DAVIES	63	8.2	1.7
FINCHAM	26	6.4	1.5
MACLACHLAN	24	6.5	2.1
MILLER	23	7.0	1.5
STEED	47	5.2	0.7
VEITCH	56	6.3	1.5

Photo Interpretation Results

Table 4 gives the maple decline level for each site based on photo interpretation of colour IR photography. The main variable used

Table 4. Colour IR photo interpretation results.

Plot	Number of trees sampled	Mean PDI ± STD	
		interpreter 1	interpreter 2
DAVIES	70	2.8 ± 0.8	2.8 ± 0.8
FINCHAM	58	2.5 ± 0.7	2.4 ± 0.8
GRIFFITH	78	2.6 ± 0.7	2.7 ± 1.1
MACLACHLAN	50	2.8 ± 1.2	2.6 ± 1.0
MILLER	59	2.4 ± 1.0	2.4 ± 1.1
STEED	52	2.5 ± 0.8	2.3 ± 0.9
VEITCH	63	2.5 ± 1.0	2.8 ± 0.9

in this index was chlorosis but overall crown condition was also taken into account.

On natural colour (NC) photos, declining maple trees appear in various shades of green to yellow-brown. On colour infrared (CIR) photos the colours are red to whitish-grey. Apart from these obvious differences there exist subtle differences between colour and colour IR photography in image rendition.

(a) On CIR the individual tree crowns are more sharply delineated.
(b) On NC bare branches are better resolved.
(c) In shadows green vegetation is better resolved on CIR but dying or dead vegetation is better resolved on NC.
(d) Biomass is much better discriminated by colour differences on CIR. This is especially evident within the crowns and is an important clue to decline assessment.

On the whole then, the CIR photos were found to be superior to NC photos in assessing decline.

Similarities between interpreters were evident in the results. The general ranking of plots from best to worst was similar (ie. with the DAVIES and GRIFFITH plots having greater decline than the STEED, FINCHAM and MILLER plots). However, there was disagreement on the status of the VEITCH and MACLACHLAN plots. For all plots and sub-plots, the determination coefficient between interpreters was $r^2 = 0.53$. The subjectivity of visual photographic interpretation of decline was apparent in the methodology and results. Two interpreters performing the same tasks under the same conditions have produced results which are similar but not highly correlated.

In comparison of the photo-interpretation index to the ground-based decline index produced by the OME, it was found that there was little correlation between them. This is partly because the OME index was determined in late June and early July in the middle of infestations of tent caterpillar at some of the sites. In August, the affected sites had recovered to some degree, producing second foliage which was small and often yellow-green. Both decline indices were therefore affected in dissimilar ways. Also, many more trees were sampled in the remote sensing study than the OME study.

CONCLUSIONS

The results of this research show that both multispectral aerial video and colour and colour IR photography are complimentary and efficient means for forest decline assessment on an individual tree basis. Digital multispectral video data analysis is quantitative and unbiased producing repeatable results. However, representation of the spectral and texture image characteristics in a decline index model involved time consuming image processing and transformation. In contrast, photo interpretation is a subjective process which requires only film processing and interpreter training. Both methods yielded maple decline indices which were well correlated with each other. This may be mainly attributed to the fact that both processes involved spectral (chlorosis) changes and texture (crown structure) changes.

An alternative approach to using spectral and texture measures of a tree crown would be to develop a hypothetical healthy crown model based on physiological and morphological characteristics of the maple species. Decline of a tree could be described by the degree it deviates from the healthy model under certain environmental conditions. Such model-based methodology should be explored in future study.

Individual tree assessment using either large scale video or colour/colour IR photography could be adopted to replace ground truth in multistage sampling. Integration of high altitude photography (or even satellite imagery), large scale photography or video imaging, and subsequent ground checking could be beneficial in terms of accuracy and cost.

REFERENCES

Conners, R.W. and C.A. Harlow., "A theoretical comparison of texture algorithms", IEEE. Trans. on Pattern Analysis and Machine Intelligence. Vol. PAM 1-2, No.3, pp204-222, 1980.

Harlick, R.M., "Statistical and structural approaches to texture", Proc. of the IEEE, Vol. 67, No. 5, pp786-804, 1979.

Irons, J.R. and G.W. Pattersen, "Texture transforms of remote sensing data", Remote Sensing of Env. 11:359-370, 1981.

Vlcek, J. and D. King, "Spectral and spatial classification of land cover types in in multispectral aerial video imagery", Technical Paper of the 5th Int. Conf. and Exposition on Electronic Imaging Society of Photographic Scientists and Engineers, Crystal City, Arlington, VA. Oct. 1986.

Vlcek, J., D. Jayasinghe, D.King, and X. Yuan, "A 4-camera video sensor and its applications in natural resources", Proc. 19th Int. Symp. on Remote Sensing of Env., Ann Arbor, MI. pp 483-489,1985.

Weszka, J.S., A.Rosenfeld, E.J. Carton, R.L. Kirby and J.M. Mohr, "A comparative study of texture measures for terrain classification", TR-361, Computer Vision Laboratory, Computer Science Centre, Univ. of Maryland, 1975.

Yuan, X., D. King, F. Cadeau, and J. Vlcek, "Multispectral video survey of a northern Ontario forest", The Proc. of 1987 ASPRS-ACSM Ann. Convention, Vol. 1, pp131-140, Washington DC., 1987.

MEIS FM: A MULTISPECTRAL IMAGER FOR FORESTRY AND MAPPING

R.A. Neville. S.M. Till

Sensor Section, Data Acquisition Division
Canada Centre for Remote Sensing
2464 Sheffield Road, Ottawa, Canada K1A 0Y7
Telephone: (613) 998-9060 Fax: (613) 993-5022 Telex: 053 3589

ABSTRACT

The MEIS II airborne multispectral imager has been providing data to the remote sensing community since 1983. During that time MEIS II has provided data for more than 200 remote sensing missions. One of the first and most successful applications of MEIS imagery has been in forestry for inventory, juvenile stand assessment, and insect damage detection and mapping. The addition to MEIS of a fore-aft stereo capability, coupled with the development of accurate techniques for the correction of geometric distortions in airborne imagery, has led to the possibility of producing high resolution data bases for topographic maps by totally digital methods.

The success of the research performed in forestry and mapping with MEIS II data combined with the advances in geographical information systems (GIS) has opened the way for the development of a new system which will include data acquisition and ground processing components, and which will provide truly operational capabilities. This report describes the sensor, MEIS FM, its conceptual design and functional specifications. MEIS FM will be a wide angle, high resolution, multispectral imager based on state-of-the-art linear arrays and custom designed high resolution optics. The design has been specifically tailored to the requirements of the mapping and forestry industries with the emphasis on the cost effective production of the end-user products.

Key Words: MEIS, pushbroom imager, mapping, forestry

1. INTRODUCTION

MEIS II is a second generation multi-spectral linear array pushbroom imager that has been in use since 1983 (Neville et al, 1983; Till et al, 1986). It has eight separate coregistered channels each filtered by one of 22 different spectral filters. Since 1987 this sensor has been operated commercially by Innotech Aviation Enterprises Ltd. During its six years of operation MEIS II has acquired data for more than 200 projects for researchers throughout Canada. Of the many applications two have the potential for wide spread operation: topographic mapping and forestry (Ahern and Leckie, 1987; Ahern et al, 1986; Gibson and Chapman, 1987; Gibson et al, 1983; Kneppeck and Ahern, 1987; Till et al, 1987).

This particular sensor design is attractive for use in mapping because it provides continuous fore-nadir-aft stereo imagery. The sensor geometry can be accurately calibrated and is stable. Any lens distortions can be readily compensated digitally and in real time. The output data are digital and hence in a form that is compatible with computerized processing systems for terrain feature extraction.

The final link required to make such line image data useful for cartography is a fast processor which will correct the image data for aircraft motion induced distortions. A prototype of such a system has been developed and is being operated at CCRS; this system uses inertial navigation data recorded simultaneously with the image data (Gibson, 1986). With National Research Council of Canada support, further developments are being undertaken by Rem/Sense Mapping Technologies, Inc. to speed up the processing and to develop a capability for the extraction of digital elevation models from line imager data.

Forestry application studies have shown MEIS II to be useful in inventory mapping, high resolution sampling, and in insect damage assessment (Ahern, 1986). Many of the system capabilities required for topographic mapping are also required for the forestry applications. In addition, the precise radiometric and multispectral aspects of the MEIS imagery are needed for species discrimination and insect damage mapping. Of course, these same sensor characteristics will benefit the mapping application by enhancing the feature extraction process.

2. SYSTEM OVERVIEW

As presently envisaged, the system will consist of an airborne segment and a ground-based segment. The airborne segment includes the sensor, navigation subsystems, real time display and image data recorder. The ground-based processor will be multi-level consisting of image data tape reader, radiometric correction subsystem, geometric correction subsystem, stereo data processor and output image recorder utilizing a computer compatible medium appropriate to the anticipated large data volumes. The output products will range from raw uncorrected imagery, to imagery that has been corrected for aircraft motion induced distortions, and finally to a fully geometrically corrected and georeferenced digital elevation model and ortho image. Each of these products will be available optionally with radiometric corrections for effects resulting from the atmosphere and solar incidence angle variations.

3. AIRBORNE SEGMENT

We present here the functional requirements imposed on the sensor by the mapping and forestry applications, how these impact sensor design, and the resulting demands on the navigation and recording subsystems. It is assumed that the sensor will utilize linear detector arrays and will be modelled on the MEIS II design. As its principal applications will be in forestry and mapping this imager is called MEIS FM.

3.1 Sensor Subsystem

The sensor characteristics considered to be necessary for the sensor to fulfil the objectives of this mapping system are summarized in Table 1. We discuss the reasons for these choices below.

To do continuous single pass stereo mapping requires a minimum of two looks at the target, one forward, the other aft. Accuracies of elevation determination can be improved by the addition of a nadir looking channel. Fore-aft look angles of $\pm 30°$ and $\pm 35°$, relative to the nadir channel, have been used successfully in the studies involving MEIS II imagery.

The number of nadir channels has been determined by spectral studies of tree species discrimination and insect damage assessment to be at least four. It is suggested that it may be wise to consider the addition of two extra channels to provide flexibility for adapting new techniques and applications. This would give a total of 8 channels, one forward, one aft and 6 nadir.

These studies also indicate that spectral bands useful to forestry applications range from 395nm to approximately 3μm, and that for some of these discrimination is enhanced by narrowing the bandwidths to 6nm and less. At the present time we are constrained to consider only the 380nm to 1000nm region by the availability of detector arrays that satisfy requirements yet to be discussed.

In order to provide the same efficient aerial coverage as is presently realized with photographic mapping cameras, the new digital systems must have comparable angular fields-of-view and resolutions. These requirements impose major demands both on the lenses and on the detector arrays. The angular field-of-view should be sufficient to give a swath width of 15km from an achievable altitude, e.g. 10-12km; this results in FOV's of 64°-74°.

The spatial resolution issue becomes somewhat more complicated in a multi-purpose sensor. Topographic mapping via established photogrammetric techniques requires high resolution aerial photography which is provided by high quality mapping cameras using fine-grained black and white films. Point features in the target are used for the photogrammetric measurements. It is adequate just to be able to distinguish these features, which are often high contrast.

In the forestry applications, on the other hand, it is necessary to be able to make accurate radiometric measurements of the various target features. In the context of the design of MEIS FM it is, therefore, the forestry application that places the greater demand on the resolution related properties of the lens-detector combination. This situation is accentuated by the fact that narrow spectral bands are required for forestry, whereas relatively broad bands can be used for mapping. This means that for the multi-purpose MEIS FM the lens is required to be large aperture and at the same time have a modulation transfer function (MTF) sufficient to give accurate single pixel radiometric measurements. (The radiometric accuracy increases with increasing system MTF and decreases with increasing scene modulation.) This combination of wide angle FOV, large aperture and high MTF is difficult to achieve and is not provided even by the best existing mapping camera lens designs. We are currently having a lens designed at the National Research Council, Ottawa, specifically for use in a long linear array camera with just this combination of characteristics.

The forest mapping application requires a resolution-swath width combination that can be achieved only by a detector array that is at least 4300 elements long. The cartographic application addresses a range of scales with a commensurate range of position accuracies. The greater the number of elements in the detector array, the better the position accuracy achievable from a given flying altitude. One should therefore choose the longest available array that satisfies the sensitivity and MTF requirements. A 6000 element array would provide a swath width of 3km with 0.5m pixels for forest inventory or a swath width of 15km with 2.5m pixels for 1:20 000 scale topographic mapping.

Another parameter that impacts the resolution achievable by such a sensor is the image line or scan line rate. One of the forestry applications, high resolution sampling, is best served by a resolution of 0.25m. For this particular mode of operation in an aircraft that is also capable of altitudes of 10-12km (for topographic mapping), line rates as high as 320 lines per second will be required. This in turn demands digitization rates up to 1.92MHz per channel.

Twelve bit digitization is chosen to be commensurate with the demonstrated dynamic ranges of such sensors and with the capabilities of modern digitizers. While the state-of-the-art image analysis of varied forest targets does not yet demand 1 part in 4000 radiometric precision, the extra range will benefit the analysis of imagery having large intrascene dynamic ranges resulting from solar incidence angle variations. The price for this is a bit rate of 5.84Mbits per second per channel for 1:20 000 scale mapping to 23.3Mbits per second per channel for high resolution sampling at a pixel size of 0.25m.

In addition to the foregoing performance requirements, there is the usual demand that the target reflectance be measurable to some specified accuracy for a given solar elevation angle. This condition imposes a lower limit on the signal to noise ratio (SNR) which in turn is achieved first by minimizing sensor readout noise and then increasing system étendue or optical throughput to meet the accuracy criterion. The latter task becomes the challenge as both large aperture and wide angular field of view work to reduce the MTF. However, the MTF must be high to give accurate single pixel radiometric and reflectance measurements. At the same time a poor SNR results in a reduced effective MTF.

In the recent development of arrays 4096 elements and longer, the trend has been to smaller element size. While this enables one to design a more compact sensor, it reduces detector area and hence optical throughput and signal-to-noise ratio. Clearly the preference is for the larger detector element. One other potential problem with long arrays is the increased potential for smearing resulting from less than perfect charge transfer efficiency (CTE). As an example, the Nyquist rate MTF for the last pixel out as compared to the first

out for a 6000 element array with two 3-phase shift registers having a respectable CTE of 0.99996 will be only 0.49. One must, therefore, place a premium on CTE when working with the longer arrays.

3.2 Critical Ancillary Subsystems

Of the various ancillary components in the airborne segment we select two that are critical to the success of the overall system. One is the inertial navigation system (INS), the other the image data recorder.

The INS data is used to correct the imagery for distortions induced by aircraft pitch, roll and yaw. The accuracy of these corrections determines the accuracy of the resulting mapping data. If one requires relative accuracies approaching the sensor pixel size, for example, 2.5m planimetric and 2m elevation accuracies from an 11km altitude, then the aircraft attitude angles must be measured to accuracies approaching 10 arc seconds.

The image data recorder is another challenging area in view of the high data output rates achieved with this sensor design. Eight channels, each with 6000 image pixels and 80 ancillary data words digitized to 12 bits when output at the highest line rate of 320 lines per second, give a total data rate of 186.8Mbits per second. Standard 14 track high density tape recorders commonly used today have a maximum rate of 3.5Mbits per second per track or 49 Mbits per second, falling far short of what will be required. Fortunately, various manufacturers are currently developing helical scan magnetic tape recorders, some of which have design goals of recording rates twice that required by this line imager.

4. SUMMARY

The basic specifications of a linear array imager for operational use in topographic mapping and forest inventory have been presented. This sensor, based on the MEIS design, should have 8 channels, of which one is forward (+35°) and one aft looking (-35°). The imager should have an angular field-of-view of 70° in order to achieve a 15km swath width from an altitude of approximately 11km. The detector array should have 6000 elements to give 2.5m pixels at this altitude.

An investigation of the implications of these requirements indicates that both the lenses and the arrays are critical components. The lenses must be relatively wide angle to achieve the desired field-of-view; they must have a relatively large numerical aperture to achieve the optical throughput and the desired signal-to-noise ratio; they must also have a high MTF across the image field to achieve reasonable single pixel radiometric accuracies. Even the best mapping camera lenses fail to satisfy all these criteria; a development effort is underway to design, fabricate and test a custom lens for the purpose.

The detector arrays should be more than 4300 elements long and preferably the longest available (6000). Special consideration should be given to detector element area, to increase optical throughput, and to shift register charge transfer efficiency to optimize MTF. It should be possible to clock these at a rate sufficient to give 320 lines per second, this to achieve the spatial resolution desired for forestry inventory sampling.

Other critical subsystems include the aircraft attitude sensing device and the image data recorder. In order to achieve the desired geometric accuracy for topographic mapping, the aircraft heading, pitch, and roll angles should be measured to 10 arc seconds. The imager specified here will output data at rates up to 186.8Mbits per second. Presently available tape recorders are unable to keep pace. However, a number of manufacturers are developing helical scan magnetic tape recorders that will take these and indeed higher data rates.

Data from this system will be radiometrically and geometrically corrected, the stereo imagery analysed, and georeferenced digital terrain models and ortho images produced on a ground-based fast digital processor. These will meet the criteria for the production of topographic maps and for forest inventory.

TABLE 1. MEIS FM SPECIFICATIONS

No. of Channels	6 nadir 1 forward at +35° 1 aft at -35°
Angular Field-of-View	70°
No. of Pixels Per Image Line	6000
Spectral Range	380-1000nm
Spectral Band Width	down to 6nm
Digitization	12 bits
Image Line Rates	80-320Hz
Data Rate (Including Ancillary Data)	46.7-186.8Mbits/sec

5. REFERENCES

1. Ahern, F.J. and Leckie, D.G., "Digital remote sensing for forestry: requirements and capabilities, today and tomorrow", Geocarto International, 2(3), pp 43-52, 1987.

2. Ahern, F.J., Bennett, W.J. and Kettela, E.G., "An initial evaluation of two digital airborne imagers for surveying spruce budworm defoliation", Photogramm. Eng., Vol. 52, pp 1647-1654, 1986.

3. Gibson, J.R. and Chapman, M.A., "Accuracy evaluation of airborne stereo line imager data", Proc. 21st International Symposium on Remote Sensing of Environment, Vol. I, pp 147-154, 1987.

4. Gibson, J.R., Park, W.M., Hollinger, A.B., Dunlop, J.D. and O'Neill, N.T., "Geometric correction of airborne line scanner data", Proc. 11th Canadian Symposium on Remote Sensing, (in press).

5. Gibson, J.R., O'Neil, R.A., Neville, R.A., Till, S.M. and McColl, W.D., "A stereo electro-optical line imager for automated mapping", Proc. 6th International Symposium on Automated Cartography, Vol. II, pp165-176, 1983.

6. Kneppeck, I.D. and Ahern, F.J., "Evaluation of a multispectral linear array sensor for assessing juvenile stand conditions", Proc. 21st International Symposium on Remote Sensing of Environment, Vol. II, pp 955-970, 1987.

7. Neville, R.A., McColl, W.D. and Till, S.M., "Development and evaluation of the MEIS II multidetector electro-optical imaging sensor", Proc. SPIE, 395, Advanced infrared sensor technology, pp 101-108, 1983.

8. Till, S.M., Neville, R.A., Leckie, D.G. and Strome, W.M., "Advanced airborne electro-optical imager", Proc. 21st International Symposium on Remote Sensing of Environment, Vol. I, pp 41-47, 1987.

9. Till, S.M., Neville, R.A., McColl, W.D. and Gauthier, R.P., "The MEIS II pushbroom imager - four years of operation", Progress in Imaging Sensors, Proc. ISPRS, ESA SP-252, pp 247-253, 1986.

COOPERATION IN REMOTE SENSING RESEARCH: THE REMOTE SENSING WORKING GROUP OF FORESTRY CANADA

Dr. Y. Jim Lee

Chairman, Forestry Canada Remote Sensing Working Group,
Pacific Forestry Centre, Victoria, B.C. V8Z 1M5T

ABSTRACT

This paper describes objectives and activities of the Forestry Canada Remote Sensing Working Group (FCRSWG). The FCRSWG was established on September 19, 1985 in Ottawa. Members come from each of Forestry Canada's Headquarters, the six regional establishments, and the Petawawa National Forestry Institute (PNFI). A statement of the role of Forestry Canada's remote sensing research and development was prepared by the Working Group.

This FCRSWG meets once a year to discuss past, current and future remote sensing developments in forestry of common interest, to coordinate member activities, and to share information on individual research programs. Topics include hardware and software acquisitions of individual members, cooperative research and a variety of special topics and joint initiatives. We intend to work cooperatively to enhance forestry remote sensing in Canada.

Keywords: Forestry Canada, Remote Sensing, Working Group, Research.

1. INTRODUCTION

Forestry Canada is the main focus for forestry matters in the federal government It provides national leadership through the development, coordination, and implementation of federal policies and programs to enhance long-term economic, social, and environmental benefits to Canadians from the forest sector. The Department is a decentralized organization with six regional forestry centres, two national research institutes, and seven regional sub-offices located across Canada. Forestry Canada Headquarters is located in the national capital region.

The research programs in Forestry Canada include forest resources, forest protection, and environment. Forestry remote sensing research provides a tool to tackle problems in all three of these major areas There is also a need for technology transfer and advisory functions to assist forest managers in making optimal use of remote sensing.

2. FORESTRY CANADA REMOTE SENSING WORKING GROUP

Because of decentralization, Forestry Canada has developed a system of working groups to help bring scientists together to improve contacts and to facilitate and coordinate research activities across the country in each program area. FCRSWG was one of the first of these to be established and the first to prepare a statement of role and functions on which this paper is based.

The FCRSWG was established on September 19, 1985 in Ottawa. Members are scientists that come from each of the six regional establishments and the Petawawa National Forestry Institute (PNFI). It is chaired by a scientist who is elected by members to serve a two-year term. Forestry Canada Headquarters provide a scientific advisor to act as secretary-coordinator. The FCRSWG meets once a year and has a supportive role in Forestry Canada remote sensing research and development, primarily in the areas of program development and coordination, and in information dissemination. Scientists take turns hosting meetings, and usually assemble a 1-day workshop to highlight ongoing remote sensing activities in their regions. The 1987 meeting was in Toronto and the 1988 meeting was in Victoria.

3. THE OBJECTIVES OF THE WORKING GROUP

The objectives of the FCRSWG are to contribute to Forestry Canada remote sensing research program development and direction, and facilitate cooperative studies among Forestry Canada's remote sensing scientists. The FCRSWG also provides expert advice on remote sensing technology to Forestry Canada headquarters, regions, and institutes. It provides coordination in acquisition and use of Forestry Canada remote sensing equipment and facilities, and makes recommendations on the purchase of new equipment. It facilitates Forestry Canada contacts in remote sensing with other agencies, and provides coordinated input in remote sensing research and development to Forestry Canada and interdepartmental planning processes. It also promotes application of advanced remote sensing technology to various disciplines within Forestry Canada.

4. FORESTRY REMOTE SENSING RESEARCH

FORESTRY APPLICATIONS: Forestry applications are the core of the Forestry Canada remote sensing program. This generally involves defining, identifying, modifying, and developing methodologies for applying remote sensing to forestry problems. Forestry Canada endeavours to maintain a strong remote sensing applications research component and cooperates with other agencies in the development of forestry applications. Forestry Canada promotes and facilitates forestry-related research and research impacting forestry applications conducted by other agencies and groups involving remote sensing. Two examples of this type of activity are Jean Beaubien's Quebec North Shore vegetation mapping with Hydro Quebec using Landsat Thematic Mapper data, and Y. Jim Lee's digital analysis assessment of root disease damage with Vancouver Forest Region using MEIS data.

RESEARCH INTO THE PRINCIPLES OF REMOTE SENSING: Research into the fundamentals and physics of forestry remote sensing is essential if major advances are to be made, and if remote sensing systems are to be optimized for forestry. Forestry Canada maintains expertise in this area and conducts research to build up an understanding of the physics of forestry remote sensing. However, this activity is not a major priority and is pursured through cooperation with Canada Centre for Remote Sensing (CCRS) or other groups. One example of this type of activity is the spectral reflectance studies at PNFI in cooperation with CCRS.

SENSOR DEVELOPMENT: Forestry Canada has a role in the development of sensors where they are relevant to forestry problems. That role is to specify the sensor design parameters and to facilitate and/or help to fund the development of the sensor. An example of this type of activity is the development of the Advanced Linear Array Imager for Forestry (MEIS F/M) at PNFI.

TECHNOLOGY TRANSFER: Forestry Canada supports the implementation of new remote sensing technologies in user agencies through advice and consultation, pilot projects, workshops, and demonstrations.

5. ACTIVITIES OF THE FCRSWG

The FCRSWP has proposed collective initiatives to various agencies. These have involved: (1) a federal interdepartmental space plan, (2) the European Space Agency's ERS-1 program, (3) detection of early signs of acid rain damage to Canada's forests, using ARNEWS (Acid Rain National Early Warning System) permanent sample plots established by the Forestry Canada's Forest Insect and Disease Survey, (4) development of an Advanced Linear Array Imager for forestry (MEIS-F/M), and (5) participation in NASA's Second ISLSCP (International Satellite Land Surface Climatology Project) Field Experiment (SIFE). The FCRSWG is currently working on remote sensing contributions to a five-year Forestry Canada strategic plan for forest management systems.

6. HIGHLIGHTS OF MEMBERS' CURRENT ACTIVITIES

Newfoundland Forestry Centre (NeFC):
Dr. B.A. Roberts is cooperating with Dr. Bajzak of the Memorial University to survey the Churchill Falls reservoir using Landsat TM data. A contract report on the use of SPOT and Landsat TM for monitoring the Hawkes Bay hemlock looper defoliation was completed in 1989 for the NeFC.

Maritimes Forestry Centre (MFC):
Mr. D. Ostaff is working in cooperation with PNFI on airborne MEIS data to determine the spectral signatures of tree species, degree of defoliation, tree components, and herbicide injury on tamarack and white pine in New Brunswick.

Laurentian Forestry Centre (LFC):
Mr. J. Beaubien is working on several aspects of image analysis, including the creation of "pseudo" spectral bands by linear combinations of the original bands, development of software for image slicing, and development of classification methodologies based on colors of simplified enhancements (a pseudo-classification algorithm).

Petawawa National Forestry Institute (PNFI):
The remote sensing research project (Dr. D. Leckie, Project Leader) at PNFI is concentrating on the use of MEIS data to study species identification, crown diameter, tree counts, stereo tree height determination, single tree classification, budworm spectral reflectance, budworm damage, and single tree budworm defoliation assessment. Other PNFI studies include Landsat TM and imaging spectrometer assessments for budworm damage, combining airborne radar with MSS for forest mapping, conifer regeneration with Landsat TM and SPOT spring imagery, conifer regeneration monitoring with Landsat MSS multidate imagery, survey of costs of inventory map production, and survey of forest inventory update procedures.

Great Lake Forestry Centre (GLFC):
Dr. R.A. Sims is working on forest cover type and site classification assessment using Landsat TM integrated with a Geographic Information System (GIS) for cutover updating, identification of new road networks, and development of spruce budworm vulnerability ratings. Large-scale 35 mm photo sampling and conventional air photo interpretation are being used for forest ecosystem mapping in northwestern Ontario. A contract study is being conducted by the Ontario Centre for Remote Sensing on regeneration assessment using Landsat TM and SPOT imagery.

Northern Forestry Centre (NoFC):
Mr. R.J. Hall is working on (a) the statistical accuracies of mapping recent cutovers using Landsat TM and MSS data, SPOT imagery, and airborne BAR imagery in cooperation with the Alberta Forest Service and CCRS, (b) development of digital enhancements from MEIS data, (c) determination of optimum large-scale photo exposure/processing parameters for forest inventory, and (d) investigation of remote sensing and GIS methodologies for spatial modelling of jack pine budworm defoliation in Saskatchewan. Mr. W.C. Moore has been involved in arranging a demonstration flight of the NASA ER-2 (U2) aircraft over Alberta in 1988 and 1989 for acquisition of airborne BAR imagery.

Pacific Forestry Centre (PFC):
Dr. Y.J. Lee is working on remote sensing applications to pest damage appraisal and forest weed control strategies, a microcomputer database system for updating forest resource inventories using satellite digital data, and computer digitization and digital

image analysis for determining the areal extent and degree of soil disturbance on cutover areas.

7. CONCLUSIONS

The FCRSWG intends to continue various forestry remote sensing research activities. Research in the integration of remote sensing system with GIS, with expert systems, as part of forest management systems, and with digital terrain models for forestry will be emphasized in the years to come. New sensors, digital stereo imagery, and better enhancements for specific forestry applications are also important.

8. ACKNOWLEDGMENTS

The author wishes to express his sincere appreciation to members of the Working Group for their review and comments on this paper. Working Group members are: Mr. Jim Richardson (ForCan HQ), Dr. Bruce Roberts (NeFC), Mr. Don Ostaff (MFC), Mr. Jean Beaubien (LFC), Dr. Don Leckie (PNFI), Dr. Richard Sims (GLFC), Mr. Ron Hall (NoFC), Mr. Walter Moore (NoFC), and the author.

COUNT OF BRAZILIAN PARANÃ PINE CROWNS IN 35 MM AERIAL COLOR PHOTOGRAPHS

Attilio Antonio Disperati
Júlio Skalski Júnior
Universidade Federal do Paranã
Depto de Silvicultura e Manejo
Caixa Postal 2959, 80.001 - Curitiba,Paranã- Brazil
Phone (041)252-7022, Telex 041-5100 UFPR

ABSTRACT

The brazilian Paranã pine is one of the most important trees of the subtropical forest in southern Brazil. In the mature stage, the tree presents a dominant crown with an umbrella like shape. This paper deals with the crown counting of Paranã pine trees as observed on 35 mm aerial vertical color photos,in scale 1:2.033, from a 3,8 hectares forest area located in the urban area of Curitiba, Paranã State. The photo-interpretation consisted mainly of crown delineation and was carried out by three interpreters with different technical knownledgement in the subject. According to the data collected in the field work, a map was prepared showing the spatial position of 194 Paranã pines, of which 145 could be delineated from the photos. The interpreters did not sketched the same crowns, the differences are related with the size, shape and number of crowns and the degree of generalization in the dense crown areas. The best results acquired by two interpreters was 68,3%. As seen on aerial photos by the interpreters, 42,3% to 50,5% of the vegetation on the area is covered by crowns of Paranã pine.

KEY WORDS: Paranã pine counting,35 mm aerial photos

1. INTRODUCTION

Paranã pine (Araucaria angustifolia(Bert.) O. Ktze) is one of the most important trees of the Araucaria forest which is a special association of the wet tropical forest. Araucaria forest occurs in several countries of South America,but the largest area is located in the states of Paranã,Santa Catarina and Rio Grande do Sul in the southern region of Brazil.

In the mature and old age natural forest the Paranã Pine is the most common and dominant species. It usually form the main part of the stand and has a comparativelly well developed and circular crown, with a candelabra shape. Taken in account these caracteristics it is easy to identify Paranã pine in the aerial photographs (Disperati,1981).

Even though the considerable advantage for photo-interpretation studies, no research was carried out to evaluate the accuracy for counting Paranã pine in aerial photos. The present study was set up in order to fill out the mentioned gap. Large scale 35 mm aerial p tos were used, the results from three diferents interpreters were compared in order to evaluate the precision of such methodology.

2. MATERIALS AND METHODS

2.1. Study area

The study area, known as "Capão da Imbuia", is located in a suburb of Curitiba city,capital of Paranã state, and ten kilometers east from the downtown. The topography of the study area is flat and the natural forest covers near all its 3,8 hectares. The area was transformed in a " preservation area" and it can be considered as unmanaged for the last 20 years.

The vegetation of "Capão da Imbuia" is typic of Araucaria forest.

According to the social structure of the forest, three different canopies are found in the study area. The first is formed by the crowns of the dominant trees , mainly Paranã pine. The second canopy, formed by the codominant and intermediate trees is also occupied by crowns of Paranã pine united with other few species,as "imbuia"(Ocotea porosa),"canela preta"(Ocotea catharinensis),"canela guaiacã"(Ocotea puberula) and "caroba"

(Jacaranda micrantha). The third and last canopy is formed by the crowns of dominated and suppressed trees and represented mainly by mirtaceas trees.

According to the size of some Paranã pine and other dominant trees, the age of the forest can be estimated as close to 200 years old. The highest tree found in the area is one Paranã pine with 30,0 meters height , measures 3,3 meters of d.b.h. and 18,8 meters of mean crown diameter.

2.2. Aerial photos

The 35 mm aerial photographs were taken at 3:45 pm on October 3rd, 1985, using a profissional 35 mm camera Canon F1 equipped with 50 mm focal lens, in a Meyer camera mount support (Meyer,1972) adapted in a right door of a high wing aircraft.The flight height was 400 m above the terrain and after the enlargements (10cm by 15cm) made from the conventional color negatives with 100 ASA, the photos had a scale 1:2.033. The forest area was entirely covered by only one flight line with six stereoscopic photos with 80% of longitudinal overlap.

2.3. Photo-interpretation

The photo-interpretation was carried out by three different interpreters. The first (A) had a theoretical and practical experience in the subject;the second (B) was envolved only one year with aerial photo-interpretation and the third interpreter(C) had no experience at all with the subject and was trained specifically for the task. All the interpreters were envolved in some aspects with the Forest School. Each interpreter was asked to perform the same task, as follow: to sketch the perimeter of the study area,the roads,buildings and lake that are found in the area and also the single crowns of Paranã pine and where this was not possible,to delineate the dense crowns area of Paranã pine.All the interpreters used the same mirror stereoscope with three times magnification binocular. The time spent by each interpreter was recorded and no limit was given for the task.

2.4. Field work

An intensive field work was carried out to produce a map showing the spatial position of all the Paranã pine in the area. After that map was done, another map was produced showing the position of each crown of Paranã pine appearing in the following canopies: dominate,codominante and intermediate. It was assumed that all the mentioned crowns could be seen on the 35mm photos.

2.5. Analysis of the data

The resulted maps were analysed by qualitative and quantitative methods. In the first, the maps were visually compared with the ground crown map. In the second two different ways were used. In the first, the study area was divided in 5 stands and in each stand it was compared each crown delineated with the crown map. This would make easy the comparison analysis in order to verify if each single crown sketched on the photos was representing only one or more Paranã pine. In the second way,it was determined the crown surface of the Paranã pine using dot grid with 16 points per square centimeter placed over the transparencies containing the photo-interpretation.

3. RESULTS AND DISCUSSION

In the present study it was used enlarged 35 mm aerial photos in a scale 1:2.033. The use of large scale photos is not frequent in studies of photo-interpretation carried out in Paranã state. The majority of the photogrammetric aerial surveys in the state, result in conventional aerial photographs in scales ranging between 1:8.000 to 1:25.000.

With the field work data, a map (fig 1.) in scale 1:500, was produced showing the spatial position of 194 Paranã pine trees occuring in the study area. From this total number, it was assumed that only 145 crowns of Paranã pine could be seen on the 35 mm photos, since 42 trees were considered as suppressed and seven presented no crown at all.

The crown maps resulted from the photo-interpretation are showed in fig 2,3 and 4. These maps are not exactly the same,notwithstanding of some similarities ; the interpreters B and C produced more detailed crown maps.

The size and the shape of the crowns were not sketched igually by the interpreters. For the interpreter B near all the crowns were near circular. Although these differences are not important in the context of this study and they can be partially explained by the variations in the skill of the interpreters, carefull analysis about these variations must be taken in studies related with the crown diameter and crown area measurements on the aerial photographs.

The analysis of the quantitative data reported on table 1 confirm that the interpreter A did a less detailed photo-interpretation resulting less crowns sketched but with larger surface of crowns.

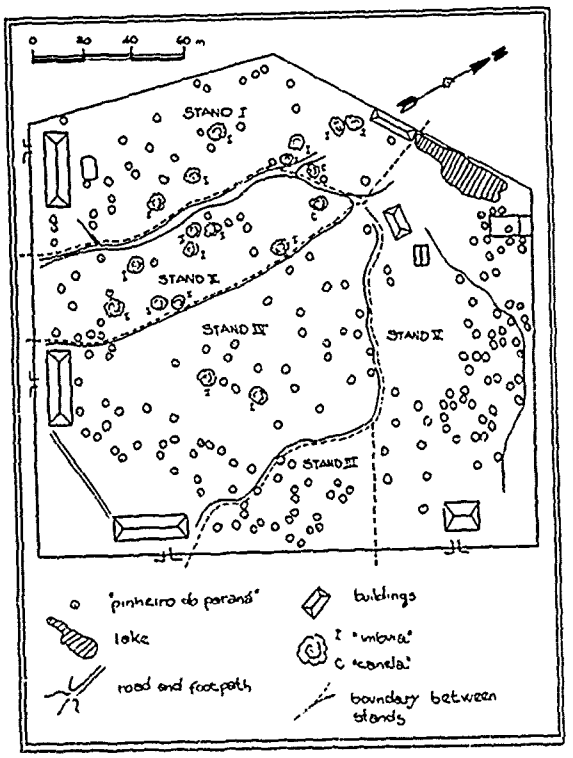


Fig 1. Location of Paranã Pine in "Capão da Imbuia"

Fig 2. Photo-interpretation crown map(Interpreter A)

From the 82 single crowns delineated by the interpreter A, 73 were corrected and represented individual trees. Each of the other 9 single crowns was representing a different cluster of trees. Using the same analysis, the interpreters B and C had a better performance than interpreter A.

Taken in consideration the relationship between the

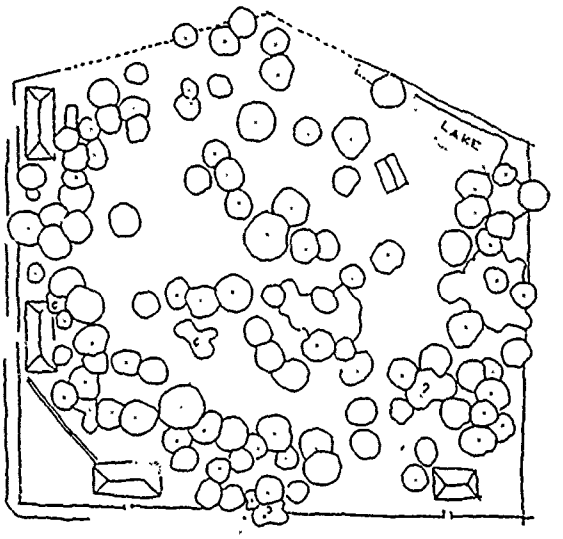


Fig 3.Photo-interpretation crown map (interpreter B)

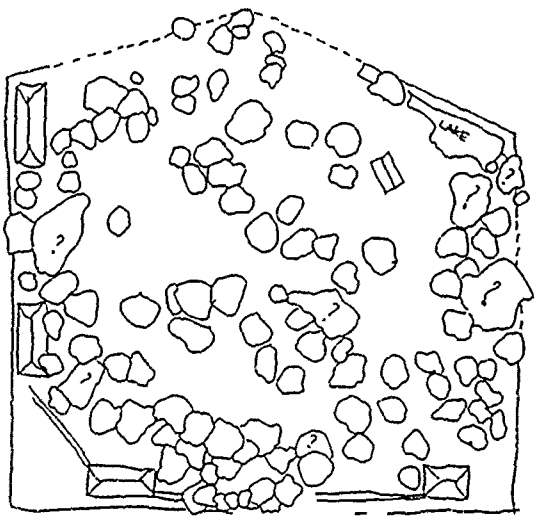


Fig 4. Photo-interpretation crown map (interpreter C)

total numbers of single sketched crowns and the possible 145 trees to be mapped, the performance of the interpreter A was 82 ÷ 145 = 56,5%. These values for the interpreters B and C were 75,2% and 72,4% respectivelly.

For the interpreter A, 11% of its total sketched crowns were representing different number of trees. For the interpreters B and C those values were 9,2% and 5,8% respectivelly. The comparison between the number of Paranã pine and the numbers of sketched crowns in each of the five stands of "Capão da Imbuia" is reported by Disperati&Skalski (1988).

With the data from the interpreter A, 50,5% of the

forest area in "Capão da Imbuia", as observed from the 35 mm aerial photographs, is covered by crowns of Paranã pine, either in individual or clustering of crowns. The values acquired from the others two interpreters were near the same, but smaller than the interpreter A.

Table 1. Number of crowns and percentage of crown surface sketched by the interpreters.

INTERPRETER	NUMBER OF CROWNS		% CROWN SURFACE		
	SKETCHED	CORRECT	SINGLE	CLUSTER	TOTAL
A	82	73	34,3	16,2	50,5
B	109	99	36,0	7,3	43,3
C	105	99	35,1	6,5	42,3

As reported in table 1, the total results acquired by the interpreters B and C are near the same, but slight different from the interpreter A, who had more experience with the subject. This can be explained by the time spent in the photo-interpretation, since that for the interpreter A 35 minutes was need, while the interpreters B and C used twice more time, 70 and 75 minutes respectivelly. Probably, due his experience, the interpreter A did not give much attention to the interpretation, as did the other interpreters. The photo-interpretation is subjective, but if the interpreter A had spent more time in the interpretation , probably the results would be near the same obtained by the interpreters B and C , at least with variations of few crowns sketched and small differences in surface of crowns.

The results obtained in this paper are in accordance with Spurr (1960) that pointed out that crown counts are almost universally low in aerial photos. It can be recalled that the present results could be improved considering in the analysis only the dominant and co-dominant Paranã pine and/or to count only the crowns instead of sketched its boundaries.

4. CONCLUSIONS

The main conclusions of this study are:

A) The three interpreters did not delineated the same Paranã pine crowns ; the differences were observed in shape, size and level of interpretation,

B) The two less experienced interpreters observed near the same quantitative values of the photo-interpretation. The more experienced interpreter spent near half of the time used by the others for the photo interpretation and because that presented less precise results,

C) The interpreters sketched correctly from 50,3% to 68,3% of all the 145 possible Paranã pine to be seen on the aerial photographs, which were 74% of all the Paranã pine counted in the field,

D) 42,3% to 50,5% of the forest area of "Capão da Imbuia", as seen on the photographs, is covered by Paranã pine crowns,

E)Using large scale aerial photographs, there is no need for skilled interpreters to sketched crowns of Paranã pine.

5. BIBLIOGRAPHY

1. Disperati,A.A. The mapping of stands of Paranã pine (Araucaria angustifolia (Bert.) O.Ktze) in the forest of south-west of Paranã State using computer-aided analysis of Landsat MSS data. Bedford College, University of London, PhD thesis,250 p,1981.
2. Disperati,A.A.;Skalski Jūnior, J. Contagem de copas de Pinheiro do Paranã em fotografias aēreas verticais 35 mm. IN: Proceedings of the V Simpōsio Brasileiro de Sensoriamento Remoto, Natal (RN),INPE,Vol 3, pp 619-626,1988.
3. Meyer,M.;Grumstrup,P. Operating manual for the Montana 35-mm aerial photography system. Sec. revision. Univ. Minn. IAFHE RSL Res.Report 78-1,62p, 1878.
4. Spurr,S.H. Photogrammetry and Photo-Interpretation. Second Edition,The Ronald Press Company, New York, 471 p,1960.

COMPARISON OF SATELLITE FOREST COVER MAPS TO TRADITIONAL FOREST INVENTORY MAPS

by

J.E. Polson
D.R. Ens
J.W. Whiting

Saskatchewan Research Council
15 Innovation Boulevard
Saskatoon, Saskatchewan
S7N 2X8

ABSTRACT

The use of digital Landsat MSS data provides a means of rapidly mapping major cover types in northern Saskatchewan. Aerial assessment of the final classification indicates good general agreement between the classification results and actual ground data for most cover types.

Accuracy assessment on a pixel by pixel basis is difficult and requires the location of very specific sites on the ground.

The majority of disagreements between the FIM and the thematic maps were the result of differences inherent in the two maps.

1. Species resolution of the FIM's allows for up to 25% of undefined commercial forest species.
2. The mapping resolution (minimum area mapped) of the two maps differed resulting in some loss of information on the FIM (4 pixels) compared to the thematic map (1-2 pixels).

The discrepancies between the FIMs and a digital classification of two burnovers indicated that:

1. Areas with living tree canopy but "burnt" ground were classified as burnover on the FIM but not on the LTM.
2. Approximately 90% of the discrepancies occurred within the non productive classification between exposed mineral soil, old burnovers with little regeneration and more recent fires.
3. There was confusion in the digital classification between exposed mineral soil, old burnovers with little regeneration and more recent fires.

Mapping natural features with the aid of digital classification of Landsat MSS imagery results in reasonable accuracy. A comparison of digitally classified thematic maps and existing cover maps must take into consideration inaccuracies and resolution differences between the two maps.

INTRODUCTION

The use of digital Landsat Multi-spectral Scanner (MSS) data provides a means of rapidly mapping major cover types in northern Saskatchewan. Assessment of how accurately the resulting classification represents the actual distribution of the cover types or other forms of mapping require investigation.

This paper compares the accuracy of a digital classification and existing forest inventory maps (FIM) in the boreal forest region of Saskatchewan, Canada. The paper also investigates the use of Landsat MSS to locate and map burnovers and compares this method to other methodology currently in use.

STUDY AREA

The Wapawekka map sheet area 72-I was chosen after consultation with wildlife and forestry personnel. The region is located in east central saskatchewan between 54 and 55 north latitude and 104 and 106 west longitude (Figure 1). The area lies in the boreal forest region of Saskatchewan.

In the more southerly portion of the map area trembling aspen (Populus tremuloides) is common on till soils. The lighter soils support Jack Pine (Pinus contorta). Both black and white spruce (Picea mariana and P. glauca) occur throughout the area, although black spruce is generally more common (Head et al. 1981).

The most dominant big game animal is the moose (Alces Alces). A wide variety of small mammals and mammalian predators occur within the area as well.

METHODS

Satellite data in colour transparency and computer compatible tape (CCT) form were purchased from the Prince Albert Satellite Station (PASS). All digital analysis were conducted using ARIES II software from the DIPIX company, Ottawa. Analysis were done at the Saskatchewan Remote Sensing Centre in Saskatoon, Saskatchewan.

Two MSS images, June 6, 1984; October 8, 1984 were chosen as working images. The images were geometrically corrected to the Universal Transmercador coordinates (UIM) and resampled to 50 m. An unsupervised classification on the June 6 image resulted in 8 cover types. Manually theming was used to generate the remaining four classes. Colour thematic maps (CTM) were printed on an ACT II INKJET plotter.

The accuracy of the cover mapping was assessed in two ways:

1. Similar cover types on FIM at a scale of 1:12,000 were colour-coded and overlain on CTM in order to look at gross agreement between the two forms of cover type assessment. Analysis were subjectively reported.
2. A more detailed assessment of agreement between the CTM and the FIM consisted of individual pixels at the intersection of UTM coordinates located on the FIM.

Are calculations of burnovers derived from Landsat digital analysis were compared with dot grid counts of similar information contained on FIM. Accuracy assessment compared total areas and georeference accuracy between the geometrically corrected Landsat thematic maps and the FIM.

RESULTS

A subjective comparison of the Landsat thematic cover map and FIM indicated good general agreement between the two data sets with most of the disagreement occurring at the pixel level.

At the pixel level, disagreement occurred in three general areas.

1. Confusion between cover types.
2. Class confusion resulting from differences in the resolution of the FIM and the Landsat thematic map (LTM).
3. Differences in class descriptions between the FIM and the LTM.

Table 1. A Comparison of Sample Units from the Landsat Thematic Maps to Forest Inventory Maps.

	Forest Inventory Categories % Agreement						
Landsat Theme	Water	Soft-wood	Hard-wood	Mixed-wood	Treed Bog	Open Muskeg	Brush-land
Lakes	95.0	1.4	0.0	0.7	0.7	0.7	1.4
Coniferous Forest	0.1	64.5	1.6	8.6	20.5	2.5	2.1
Deciduous Forest	0.0	10.8	65.7	16.6	2.9	0.0	3.9
Deciduous Mixed Forest	0.0	23.4	21.0	27.5	10.2	1.2	4.2
Treed Bog	0.4	17.5	2.5	3.2	51.1	23.2	2.1
Open Bog Marsh	0.0	0.0	0.0	0.0	52.5	40.0	7.5
Willow/Shrub	0.0	2.2	6.5	8.7	45.7	34.8	2.2
Open Jack Pine	0.0	29.5*	0.0	0.0	34.1	34.1	2.3

* not distinguished from the general softwood category.

The major sources of disagreement occurred when undetected areas of deciduous or coniferous forest occurred in the mixed wood class (according to FIM) and where upland coniferous forest was confused with densely treed bog and vice versa. The willow-shrub class identified in the CTM consisted of shrubby components which could occur in three classes on the FIM, i.e. Treed bog, Open muskey and Brushland.

The discrepancies between the FIMs and a digital classification of two burnovers indicated that:

1. Areas with living tree canopy but "burnt" ground were classified as burnover on the FIM but not on the LTM.
2. Approximately 90% of the discrepancies occurred within the non productive classification between exposed mineral soil, old burnovers with little regeneration and more recent fires.
3. There was confusion in the digital classification between exposed mineral soil, old burnovers with little regeneration and more recent fires.

The GEM tables show the dot counts of FIM and digital classification classes and the calculated percent similarity of the classes (Tables 2 and 3). Class-area calculations for the FIM's, the digital classifications and the percent similarity of the classes are presented in Table 4.

Table 2. GEM Table of Dot Counts (D) and Percentage Comparison for Montreal Lake and Holiday Fire Subarea Classes by Data Source of October 8, 1984 Imagery.

	DIGITAL CLASSIFICATION			
FOREST INVENTORY CLASS	Class Ac6	Not Ac6	Class Ac7	Not Ac7
6 (Ic6)	5292	488	*	*
Not 6	554	*	*	*
7 (Ic7)	*	*	762	292
Not 7	*	*	84	*
% Similarity	+82		+64	
% Accuracy	84		67	

Table 3. GEM Table of Fire Perimeter Dot Counts (D) and Percentage Comparison for the Holiday Fire Subarea Classes by Data Source of October 8, 1984 Imagery.

	DIGITAL CLASSIFICATION			
FOREST INVENTORY CLASS	Class Ac6	Not Ac6	Class Ac7	Not Ac7
6	5774	371	*	*
Not 6	341	*	*	*
7	*	*	954	124
Not 7	*	*	81	*
% Similarity	+88		+81	
% Accuracy	89		82	

Table 4. Area Calculations in Hectares for the Montreal Lake and Holiday Fire Subarea Classes by Data Source of October 8, 1984.

FOREST INVENTORY CLASS	Forest Inventory Map	Digital Classification	% Similarity
6	2330.6	2355.8	+98.9
7	425.0	345.5	-81.3

NOTE:
(+) Digital Classification Area Oversized
(-) Digital Classification Area Undersized

HOLIDAY FIRE SUBAREA

The discrepancies between the FIM's and the digital classification using aerial photography indicated that:

1. Areas with living tree canopy but "burnt" ground were classified as burnover on the FIM's.
2. All areas with living tree canopy were not classified as "bur. '' on the digital classification.
3. Approximately 90% of discrepancies occurred within the nonproductive classification of the FIM's.
4. There was confusion in the digital classification between exposed mineral soil (fire guard), old burnovers with little regeneration and more recent burns.

The geo-reference accuracy of the Holiday Fire (Ac6) was 82.0%. This improved to 88.4% (Ac6(P)) when the perimeter of the fire was used. Similarly, when looking at the accuracy of the Bird Dog Fire (Ac7 and Ac7(P)) using its perimeter, an increase from 64.3% to 81.0% was found. In addition, another reason for the changes in accuracy is the relationship between the year of burn and the date of the Landsat CCT. Since the Bird Dog Fire burned in 1981 and the Landsat CCT date was October 8, 1984 a portion of the burn had regenerated. This explained why the area was smaller in the digital classification. Another factor of concern was the level of accuracy for the FIM's in this area. A comparison of the 1:50,000 NTS maps and FIM's indicated that the UTM grid location differed but topographical features in relation to each other were accurate to 98.2%, which explained in part the difference between the area percentage accuracy of 81.3% and the geo-reference accuracy at 64.3%.

CONCLUSIONS

Mapping natural features with the aid of digital classification of Landsat MSS imagery results in reasonable accuracy. A comparison of CTM of the classification and existing cover maps (FIM) must take into consideration inaccuracies and resolution differences between the two maps.

LITERATURE CITED

1. Harris, W.C., A. Kabzems, A.L. Kosowan, G.A. Padbury and J.S. Rowe, 1983, "Ecological Regions of Saskatchewan", Technical Bulletin No. 10, Saskatchewan Parks and Renewable Resources, pp57.

2. Head, W.K., D.W. Anderson and J.G. Ellis, 1981, "The Soils of the Wapawekka Map Area 73-I Saskatchewan", Extension Publication 303, Saskatchewan Institute of Pedology Publication SF5, University of Saskatchewan, Saskatoon, pp108.

3. Piper, S.E., 1983, "The Evaluation of the Spatial Accuracy of Computer Classification", University of Natal/Department of Surveying and Mapping, Natal, South Africa.

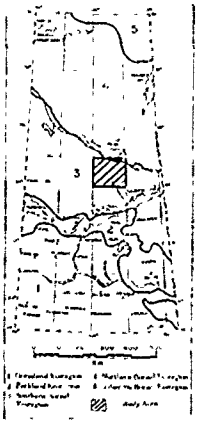


Figure 1. Location of the Study Area

CHARACTERISTICS OF FIRST ORDER INTENSITY ENTROPY AS A MEASURE OF IMAGE REGION TEXTURE

GORDON E. CARLSON

Electrical Engineering Department
University of Missouri - Rolla
Rolla, Missouri 65401 USA

ABSTRACT

The characteristics of first order image intensity entropy in measuring intensity randomness as an indication of texture in an image region are considered. These characteristics indicate that the first order intensity entropy measures less about the randomness of the region intensity values than about the dynamic range of the intensities in the region if the region contains only a limited portion of the quantized dynamic range. Thus, the modification of the first order entropy computations to include only the range of intensity levels in the image region is considered since then the entropy measure concentrates on the intensity randomness in the region. Example results of both types of entropy computations for target and background regions are shown for several FLIR images containing targets of different sizes.

Keywords: Entropy, Image Intensity, Texture Measurement

INTRODUCTION

Computation of features related to the intensity texture of an image or of some region in an image is of interest. Regions of particular interest may be targets or their surrounding backgrounds. Measures of the intensity texture of these regions are useful in region segmentation or region classification (Haralik, 1979) (Weska, 1976) (Pavlidis, 1977). Another use for texture measurements is as an image metric for the comparison of the characteristics of different images or different images of targets with regard to the difficulty which might be encountered in performing target segmentation or target classification. Consideration of this use led to the considerations discussed here.

A number of different texture measurement algorithms have been defined (Pavlidis, 1977), (Haralik, 1979), (Conners, 1980), (Rosenfeld, 1982). In some of these a texture feature used is the second order entropy associated with image intensities in pairs of pixel locations in the image region of interest. This feature is

$$E_2 = -\sum_{i=1}^{N}\sum_{j=1}^{N} p(i,j|\theta,d)\ln[p(i,j|\theta,d)] \qquad (1)$$

where $p(i,j|\theta,d)$ is the probability that a pair of pixels with relative locations as defined by the angle θ and distance d have intensities i and j and N is the number of intensity quantization levels available for the image. Entropy is also used as a feature for the difference in image intensities at pairs of pixel locations in the image region of interest. This feature is

$$E_d = -\sum_{i=1}^{N-1} p(i/\theta,d)\ln[p(i|\theta,d)] \qquad (2)$$

where $p(i|\theta,d)$ is the probability that the intensities of a pixel pair differs by i and θ, d, and N are defined as above. Actually, image intensity entropies of various orders (using various signal groups of pixels) could be used as texture features. Such entropy measurements are measures of the intensity randomness in the image region being considered since, for the defined pixel relative geometry, they are maximum when all possible combinations of pixel intensities are equally likely and zero when only one combination of pixel intensities occurs.

To compute an n^{th} order intensity entropy for an image region, conditional n^{th} order joint probability mass functions conditioned on the defined pixel relative geometry must be estimated from the image intensity values in the region being considered. The simplest intensity entropy measure is the first order intensity entropy. It only requires an estimate of the first order image intensity probability mass function for the region being considered which does not depend on any relative pixel geometry. It is a measure of the randomness or information content in the image region intensity values without regard to their spatial distribution characteristics. Thus its texture describing characteristics are limited, but it can be quite useful for small regions since joint probability mass functions are more difficult to estimate effectively with the few number of pixels in small regions. The characteristics of first order intensity entropy are considered here since this was an image comparing metric proposed; however, the concepts discussed are also applicable to higher order intensity entropies.

FIRST ORDER ENTROPY CHARACTERISTICS

It is assumed that a quantized dynamic range is quantized into M grey levels for an individual image of interest where the quantized dynamic range and/or M may not be the same for each image. This would occur for example if the images came from different sources. Some images may have dynamic ranges which are less than their quantized dynamic ranges. That is, some images may have grey levels which do not encompass the entire set of M grey levels corresponding to the entire quantized dynamic range.

To compute the first order intensity entropy for a region of an image, the intensity probability mass function encompassing all possible image grey levels (i.e. M levels) must first be estimated for the region. This is done by computing

$$p_i = \frac{n_i}{N} \tag{3}$$

where N is the number of pixels in the region, n_i is the number of times the i^{th} grey level occurs in the region, and p_i is the estimate of the probability that the i^{th} grey level occurs. The normallized first order intensity entropy for the region (called region entropy from here on for simplicity) is

$$E_n = -\frac{1}{\ln(M)} \sum_{i=1}^{M} p_i \ln(p_i) \tag{4}$$

where normalization is used to produce a value which lies between 0 and 1, and the total entropy depends only on the randomness of the intensities within the quantized dynamic range and not on the number of image grey levels used in a particular image.

If the dynamic range of the image region of interest is less than the quantized dynamic range used, then the maximum region entropy for the image region is less than 1. This maximum region entropy, E_{nM}, occurs when the grey levels spanned by pixels in the region are uniformly distributed. Therefore, for M_r grey levels spanned,

$$E_{nM} = -\frac{1}{\ln(M)} \sum_{i=k}^{k+M_r-1} \frac{1}{M_r} \ln\left(\frac{1}{M_r}\right) = \frac{\ln(M_r)}{\ln(M)} \tag{5}$$

where k is the minimum grey level in the image region. This maximum region entropy is plotted in Figure 1 as a function of the ratio of the region dynamic range to the quantized dynamic range, $\rho_D = M_r/M$.

If the region dynamic range is a small fraction of the quantized dynamic range, then the region entropy measures less about the randomness of the intensity values in the region than about the reduced dynamic range of the region. This is apparent from Figure 1 since most of the entropy change from its maximum value is taken up by the dynamic range change. In this case, the region entropy is partially measuring the same characteristic as the intensity standard deviation feature for the region measures. In some cases it may be desirable to separate these two effects and

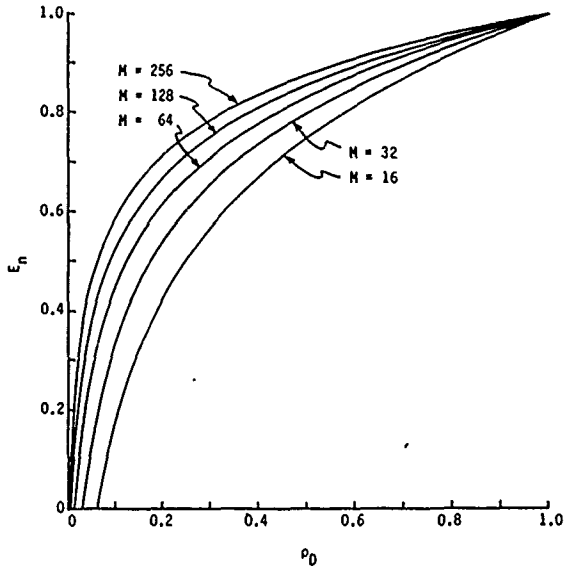


Fig. 1. Maximum Region Entropy vs ρ_D.

then use the two features together to characterize the region intensity variation. To do this, the region entropy calculation can be modified to be

$$E_{nr} = -\frac{1}{\ln(M_r)} \sum_{i=k}^{k+M_r-1} p_i \ln(p_i) = \frac{E_n}{E_{nM}} \tag{6}$$

where the last form of E_{nr} follows since $p_i = 0$ for all grey levels outside the region dynamic range (i.e. $i < k$ and $i > k + M_r - 1$). This modified region entropy considers only the intensity randomness inside the dynamic range of the region since its maximum value is 1 when the grey levels are uniformly distributed within the region dynamic range.

EXAMPLE RESULTS

As an example of region entropy, E_n, and modified region entropy, E_{nr}, values corresponding to several forward-looking infrared (FLIR) images are shown. The images used are of a set of vehicle targets at various ranges. The background is rather bland, being, for the most part, grass and roads. Some of the images contain tree clusters, extraneous man-made bright objects, power poles, small buildings, and apparent rock outcroppings. The images were taken at nominal ranges to the targets of 2.5, 3.5, 5, and 10 km. A 128x128 pixel segment surrounding the targets was selected from each of two images at each range. These segments encompass two or three targets in each segment. The average size of the targets on the images varies from 194 pixels at a range of 2.5 km to 17 pixels at a range of 10km. The quantized dynamic range is 128 grey levels. Only one of the images (one of the two at 2.5 km) has intensities which span the entire 128 grey levels.

Estimated probability mass function data for one of the image segments taken at a range of 2.5km and one of the image segments taken at a range of 10 km are shown in Figure 2. The estimated probability mass values shown are averages of two adjacent values and are connected by straight lines to make visualization easier. Data are shown for all the target pixels and for all the background (image segment minus targets) pixels. The target pixels contain a wider variation of intensity values and have a less peaked probability mass function than the background pixels. This is particularly true for the image from the short range where the resolution is great enough to show some target detail.

Figure 3 shows the average of the region entropies, E_n, and modified region entropies, E_{nr}, computed over all targets and backgrounds at each range. Since there are only 4 to 6 targets and 2 backgrounds considered at each range, the data is somewhat spread. Therefore, a linear regression line was plotted through each set of data to more clearly show trends. It is clear that the region entropy values, E_n, are significantly effected by the dynamic ranges of the targets and background since they are considerably smaller than the modified region entropy values, E_{nr}. This is as expected since the region dynamic range is less than one half than the quantized dynamic range in most target and background regions considered. The region entropy for the targets decreases more rapidly with range than that for the background. This is expected due to the reduction of the number of target pixels with range which leads to fewer possible grey levels in the target regions. The modified region entropy gives a more consistent measurement of target or background intensity randomness over the images considered as it should since the dynamic range change is being ignored. This is shown by the relatively constant values of E_{nr} as a function of range to the image segment. The targets intensities have greater randomness than the background. This is also apparent from Figure 2 where the background probability mass functions are more peaked than the target probability mass functions.

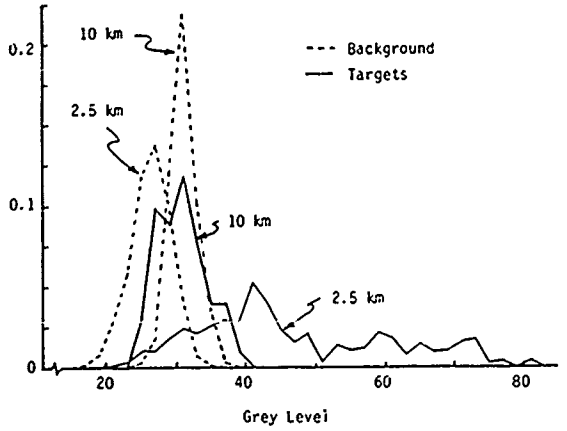


Fig.2. Example Region Probability Mass Functions.

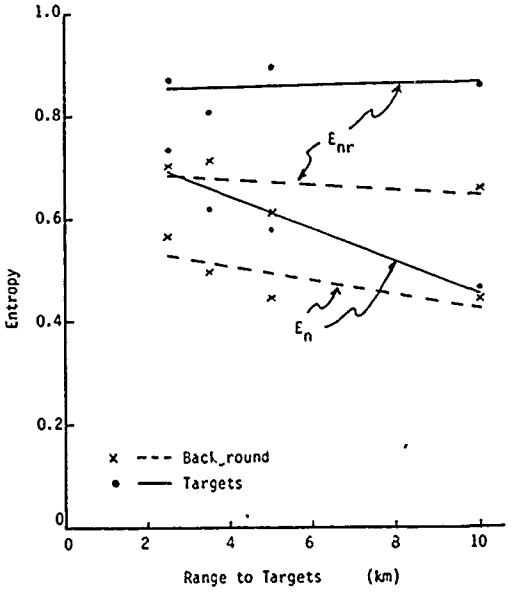


Fig.3. Entropy Values for Example Image Segments.

SUMMARY

The characteristics of first order intensity entropy values in representing intensity randomness in a textured image region have been discussed. Such intensity randomness is a texture feature and may be useful in comparing characteristics of different images or different image regions. It was indicated that if the region dynamic range is significantly less than the dynamic range used to calculate the region entropy, then the entropy value measures less about the intensity randomness in the region than about the reduced dynamic range in the region. The region entropy can be computed using only the dynamic range of the region (called the modified region entropy here) so it will concentrate on measuring intensity randomness in the region. An example is shown to illustrate these characteristics. Images used are FLIR images of targets on a rather bland background. The same concept is applicable to higher order entropy measures such as the second order entropy features used in some texture measurement algorithms.

REFERENCES

1. Conners, R.W. and Harlow, C.A., "A Theoretical Comparison of Texture Algorithms," IEEE Trans. on Pattern Analysis and Machine Intelligence, Vol. PAMI-2, No. 3, pp 204-222, 1980.

2. Haralik, R.M., "Staatistical and Structural Approaches to Texture," Proc. of IEEE, Vol. 67, No. 5, pp786-804, 1979.

3. Haralik, R.M., Shanmugam, K., and Dinstein, I., "Textural Features for Image Classification," IEEE Trans. on Systems, Man, and Cybernetics, Vol. SMC-3, No. 6, pp. 610-621, 1973.

4. Pavlidis, T., Structural Pattern Recognition, Springer-Verlag, New York, 1977.

5. Rosenfeld, A. and Kak, A.C., Digital Picture Processing, Academic Press, New York, 1982.

6. Weska, J.S., Dyer, C.R., and Rosenfeld, A., "A Comparitive Study of Texture Measures for Terrain Classification," IEEE Trans. on Systems, Man, and Cybernetics, Vol. SMC-6, No. 4, pp269-285, 1976.

Applications of Geostatistics to Image Analysis

J.A. MacDonald
J.R. Carr

Mackay School of Mines
University of Nevada - Reno
Reno, NV 89557
USA

Abstract

Experimental semi-variograms of a portion of a Landsat Thematic Mapper scene dominated by a dendritic texture (bajada) were calculated for bands 1,2,3,4,5 and 7. Complex spherical models were then fitted to the semi-variograms, consisting of two sets of parameters. It was found that the complex models match the experimental semi-variograms very closely up to the second range in the model, after which a drift in the experimental values takes over. The first range in all cases was consistent between bands as expected. The second range for VNIR spectral bands measured on the uncooled primary focal plane (Bands 1,2,3 and 4) remained constant between bands, whilst the SWIR bands on the cooled secondary focal plane (bands 5 and 7) were identical but different from the VNIR bands.

Image degradations using four-pixel and nine-pixel averages were performed to investigate the effects of regularization in image data by simulating coarser resolution imagery. As expected the variance decreased continually with increasing degradation. A relationship between the variance of the data and the sum of the sills was determined by plotting variance vs sill sums. For the VNIR bands a strong linear relationship was observed with a slope of 1.3, passing through the origin. The SWIR bands plotted with a similar slope, but offset to shorter separation values.

Overall results show promising potential for spatial enhancement using a cokriging estimation as explored by Glass et al (1988).

Introduction

The purpose of this paper is to present some results in the investigation of the spatial domain of remotely sensed images using semi-variograms (hereafter referred to as variograms). The use of variograms for this purpose has been studied by several authors (Carr and Myers, 1984; Woodcock et al, 1988a & b) who have demonstrated the potential value of this technique. Glass et al (1988) presented some preliminary results on the use of a geostatistical estimator (cokriging) to improve spatial resolution of Landsat MSS images.

The current study deals with a Landsat Thematic Mapper image of an alluvial bajada several miles north of Tonopah, Nevada, acquired on June 30, 1984. The data has been geometrically corrected using the MacDonald Dettwiler GICS system, and radiometrically calibrated (Kepper et al, 1986). This data set was chosen as a result of its predominantly dendritic texture and its availability.

Variogram analysis was invented in the late 1950's by Georges Matheron (Matheron, 1963) and is well described by Clark (1979). Based on the theory of regionalized variables developed by Matheron (a regionalized variable is one which has a fixed location in space), variograms have become a common tool in the analysis of spatial data. Digital imagery presents a classic example of such a data set.

Variograms

A variogram is a plot of half of the average square of the difference in value (DN) between two values a given distance h apart, as a function of h. In an ideal case the variogram will start at zero (there is no difference between any point and itself) and rise sharply, becoming level after some value of h. The value of h at which it becomes level is called the *range* and represents the range of influence the variables will have on each other. The value of the variogram function at which it becomes level is called the *sill* and represents the variance of the data.

The variogram function is calculated as follows:

$$G(h) = \frac{1}{2n} \sum_{i=1}^{n} [z(x) - z(x+h)]^2$$

where G(h) is the variogram function (gamma), n is the total number of data pairs separated by a distance h, z(x) is the value at x and z(x+h) is the value a distance h away from x.

Several models have been developed to model variograms. An ideal variogram, as described above, is approximated by a *spherical* model. The spherical model is calculated by the following equation:

$$G(h) = Co + C * \left[\frac{3h}{2a} - \frac{h^3}{2a^3} \right] \quad \text{when } h < a$$

$$G(h) = Co + C \quad \text{when } h >= a$$

where C is the sill, a is the range, and Co is the nugget (a representation of the random component of the data set). A complex spherical model is one which contains two ranges (a_1 and a_2) and two sills (C_1 and C_2). ie:

$$G(h) = Co + C_1 * \left[\frac{3h}{2a_1} - \frac{h^3}{2a_1^3} \right] + C_2 * \left[\frac{3h}{2a_2} - \frac{h^3}{2a_2^3} \right] \quad \text{when } h < a_1$$

$$G(h) = Co + C_1 + C_2 * \left[\frac{3h}{2a_2} - \frac{h^3}{2a_2^3} \right] \quad \text{when } h < a_2$$

$$G(h) = Co + C_1 + C_2 \quad \text{when } h >= a_2$$

When $h < a_1$ the model is influenced by both sets of parameters. When h is between a_1 and a_2 the model changes only as a result of the second set of parameters. (Clark, 1979)

Several other types of models have been proposed (Clark, 1979; Woodcock et al, 1988a & b), however, with the exception of Woodcock's disc model, they will not be considered in this study.

Procedure

Experimental variograms were calculated for a 150 pixel by 180 line sized image in each of the six reflective bands of the thematic mapper. The band 3 and 7 images were degraded by using both a four-pixel and a nine-pixel average and assigning the average value to a pixel in a new image. Thus the four-pixel degraded image was one quarter the size of the original image, and the nine-pixel image was one ninth. Experimental variograms of the degraded images were then calculated. Cross-variograms (Glass et al, 1988) were calculated between the nine-pixel average degraded band 3 image and subsampled band 2 and 4 images, as well as between the full sized band 2 and 4 images.

Complex models were fitted to each of the experimental variograms and cross-variograms using a program written by the authors which displays the model instantly as defined by assigned values for ranges and sills.

Results

In all bands it was possible to achieve very close approximations of the experimental variogram using complex spherical models. The first range was consistent from band to band, with a value of 5 pixels (one pixel is the base measurement for h). The second range in all images (including degraded) except bands 5 and 7 was also consistent with a value of 43 pixels, whilst bands 5 and 7 had a value of 23. The cross variograms between the nine-pixel degraded bands 3 image and a three times subsampled bands 2 and 4 images both had a second range of 39. First sill values ranged between 7 and 30 (with the exception of 75 for the bands 2 and 4 cross variogram), whilst second sill values ranged between 11 and 74 (again, the band 2 and 4 cross variogram was much higher at 154). See table 1.

Table 1.
Variogram parameter results

	Variance	Sill1	Range1	Sill2	Range2
Band 1	40.86	11.0	5	21.8	43
Band 2	71.29	18.5	5	37.5	43
Band 3	79.85	20.7	5	42.3	43
Band 4	73.25	20.7	5	38.2	43
Band 5	123.26	28.0	5	39.0	23
Band 7	85.00	18.0	5	20.0	23
4 pixel band 3	73.89	15.0	5	41.0	43
9 pixel band 3	69.68	12.0	5	39.0	43
4 pixel band 7	83.81	20.0	5	11.0	23
9 pixel band 7	73.57	7.0	5	20.0	23
cross band 3,2	147.74	22.0	5	72.0	39
cross band 3,4	150.57	22.0	5	74.0	39
cross band 2,4	286.18	75.0	5	154.0	43

Discussion

The extremely close fit of all the complex spherical models to this data set indicates it as being a very good model to use. However, it is to yet be determined if such is the case for all types of Thematic Mapper scenes. All indications from this study show that the disc model studied by Woodcock et al (1988a & b) is not practical for examining this type of data.

The consistent first range between bands shows that it is not spectrally dependent, and thus can be used in the analysis of spatial information with confidence. As expected all the variograms began to demonstrate a drift-like characteristic at longer separation distances (an indication of influence by a global trend). In all cases the separation distance value at which this drift begins to affect the variogram coincides with the second range; a phenomenon which is not fully understood at the moment. The consistency in the second range between bands on the same instrument focal plane leads us to believe that it may be system dependent, and not so much scene dependent. The second range should be used when determining a window of estimation influence to be used for later cokriging processing.

Upon first inspection, there did not appear to be much of a relationship between the sills and the variance of the data. As expected, the variance and sill values increased substantially for the cross-variograms. In an attempt to determine whether any relationship between the sills and the variance of the data sets did exist, a plot of the variance vs the sum of the sills was produced. The values for bands 1,2,3, and 4, including cross-variograms not involving a subsampled image and variograms of degraded images, plotted along a straight line with a slope of 1.?, passing through the origin. The values for bands 5 and 7 followed a similar slope, but were offset to the left (ie. they had a higher Y-intercept). This offset may be useful as a system evaluation tool.

As expected, the regularization effect was observed for the degraded images in both bands 3 and 7. Thus this form of degradation does simulate coarser spatial resolution imagery.

Conclusions

Variograms of Landsat Thematic Mapper data is best modeled using a complex spherical model. In this study a model using two sills and ranges modeled the experimental variograms of all Thematic Mapper reflective bands. The image used in this study was dominated by a dendritic texture. Other images with more than one dominant texture may require models with greater than two sets of parameters.

All bands generated a modeled first range of five pixels, indicating that the first range is independent of spectral information. Drift behavior in all bands commenced at the second range of the model. The second range was consistent between bands on each focal plane (43 and 23), but was different between focal planes. This second range may be possibly used as a tool for evaluating a system. Further investigation into this phenomenon is recommended.

A relationship between the variance of the data and the sum of the sills was found by plotting them against each other. The values from the primary focal plane followed a very linear relationship passing through the origin. The secondary focal plane showed a similar relationship with the inclusion of a Y intercept.

The results indicate that effective models can be obtained for use in image restoration using geostatistical estimators (cokriging).

References

Carr, J.R., and Myers, D.E., 1984, "Application of the Theory of Regionalized Variables to the Spatial Analysis of Landsat Data," IEEE 1984 Proceedings of the PECORA 9 Spatial Information Technologies for Remote Sensing Today and Tomorrow, pp. 55 - 61.

Clark, I., *Practical Geostatistics*, Applied Science Publishers Ltd., Essex England, 1979

Glass, C.E., Carr, J.R., Yang, H.-M., and Myers, D.E., "Application of Spatial Statistics to Analyzing Multiple Remote Sensing Data Sets," *Geotechnical Applications of Remote Sensing Data and Remote Data Transmission, ASTM STP 967*, A.I. Johnston and C.B. Petterson, Eds., American Society for Testing and Materials, Philadelphia, 1988, pp. 138-150.

Kepper, J.C., Lugaski, T.P., and MacDonald, J.S., 1986, "Discrimination of Lithologic Units, Alteration Patterns and Major Structural Blocks in the Tonopah, Nevada area using Thematic Mapper Data," Proceedings from the Fifth Thematic Conference: Remote Sensing for Exploration Geology, Reno, Nevada, September 29 - October 2, 1986, Environmental Research Institute of Michigan, pp. 97 - 115.

Matheron, G., 1963, "Principles of Geostatistics," Economic Geology, Vol. 58, pp. 1246 - 1266.

Woodcock, C.E., Strahler, A.H., and Jupp, D.L., 1988a "The Use of Variograms in Remote Sensing: I. Scene Models and Simulated Images," Remote Sensing of Environment, Vol. 25, pp. 323 - 348.

Woodcock, C.E., Strahler, A.H., and Jupp, D.L., 1988b "The Use of Variograms in Remote Sensing: II. Real Digital Images," Remote Sensing of Environment, Vol. 25, pp. 349 - 379.

Defining Subtle Features that are Not Confined to Reflection Domains

Joseph E. Robinson
Bronya Oldfield

Department of Geology
Syracuse University
Syracuse, New York 13244-1070

Abstract

Physical features such as tectonic structures, outcroping formations, surficial features and geochemical anomalies that are not confined to identifiable reflection domains often have their presence obscured by more prominent reflectance patterns. Many of the processing methods that tend to delegate closely related reflectances to specific classifications, or to smooth out minor variations in the image, may discriminate against subtle outlines that indicate the presence of coherent but poorly defined physical features. These subtle anomalies, which may be apparent only through continuity of shape or texture, often are of major importance to the geologic interpretation of the area. However, because they may not respond to the usual enhancement techniques, special processing methods that improve the display of subtle features are necessary for their analysis.

Spatial filtering, using operators that are designed to display features according to characteristics such as size and shape while rejecting conflicting features, can be an effective aid to interpretation. Edge detectors that display the boundaries of features have been used effectively in automatic pattern recognization techniques and for the locating of subtle outlines and lineations. However edge detectors respond only to changes in slope and may not indicate continuity of the features within the outlines. Unless the body has a recognizable shape, it may be wrongly interpreted. Conversely, band pass filters display features based on their continuity of form or texture. Also, they can eliminate many of the conflicting larger or smaller features and bring out a singular expression of individual anomalies. A series of band pass filters, each designed to display features of a specific size, can be used to analyze an image feature-by-feature until all hidden anomalies have been discovered.

Digital Fourier transforms provide an effective approach for the design and application of band-pass spatial filters. The portion of the scene that is to be analyzed is first transformed to the frequency domain by a two-dimensional fast Fourier transform program. The amplitude spectra permits the interpreter to note any dominate trends in the image and to select the frequency components for detailed examination. Filtering is accomplished by setting all unwanted components to an amplitude of zero then returning the scene to the spatial domain by the inverse Fourier transform. If the retained frequency components are confined to a band with radial symmetry about the zero frequency origin, the filtered output will have zero phase characteristics and the output features will be in their correct position. In the same manner, any number of frequency bands can be transformed to display a wide range of distinctive features. Unfortunately, filtering does require a large number of mathematical operations, however, it can be performed quickly on modern high speed work stations and is ideally suited to parallel processing systems.

Spatial filters can be applied to any digital image including those that have undergone considerable preliminary processing. Enhanced data such as the output from deconvolution or principal components analysis respond well to band pass filtering. Examples illustrate how secondary features that are hard to recognize in the original scene can be displayed prominently and correctly in the filtered output. Because the filters only retain features from the original data, all displayed anomalies can be located in the original scene.

Key words

Image analysis, Spatial filtering, Selective enhancement, Computer processing

Introduction

A remotely sensed image of the earth's surface displays the sum of all features that contribute to the recorded reflectances. They include the foreground features such as forests, farm crops, wetlands and other natural or manufactured structures that dominate the reflectance spectra as well as the background features that only make a modest contribution to the reflection patterns. The background features may be important for they often are the key to geologic deposits or other buried artifacts. They are background features because their surface expression is masked by weathering, vegetation overgrowths and soil or other surficial cover. They may be large scale such as tectonics structures, outcroping formations, non-metallic mineral deposits or geochemical anomalies; or they may be medium to small scale features such as salt domes, buried channels, solution cavities or abandoned farmsteads and archeological

sites. The presence of these background features may be indicated only by subtle changes in the more dominant foreground reflectance patterns. Where the subtle features are not concordant with the more prominent reflectance domains and are either larger or smaller in area, or have a different trend direction, then they can be enhanced selectively by the application of band pass spatial filters.

Because the image displays the sum of all component features, individual components can be deleted from the image to clarify the remainder or individual features can be extracted and displayed singularly for examination and interpretation. Also the amplitude of desired features can be increased until they are clearly defined in the image. Band pass spatial filters, designed to operate on features according to their physical size, alter the image to selectively enhance the desired characteristics. The spatial filters are designed to extract or suppress specific features according to the Fourier principle that any function of time or distance can be described by a component set of sinusoids each with a specific frequency, amplitude and phase. If the amplitudes of the components are increased the feature becomes more prominent and if the amplitudes are made equal to zero, the feature is deleted. Operations on the image can be either on the original data in the distance domain or on the component frequencies in the frequency domain. The information pertaining to the image is identical in both domains which are related by Fourier transforms (eg. Zurflueh, 1967, Robinson, Charlesworth and Ellis, 1969).

Spatial Filtering

The filtering process can be in the distance domain or in the frequency domain by use of the operational transforms. Distance domain filtering is by convolution in which the filter operator is convolved with the original image matrix. The Fourier transform of the convolution operation is multiplication and addition so that in the frequency domain the amplitude spectra is multiplied by the filter amplitudes and the phase spectra is added to the filter phases. If the frequency spectra of the filter has radial symmetry about the zero frequency origin then phase can be neglected and the position of objects in the image will not be altered. With the modern use of high speed parallel processing computers and the computational algorithm known as the Fast Fourier Transform (Cooley and Tukey, 1965), often filter design and filtering is faster and easier in the frequency domain, however both approaches are effective. The distance domain band pass filter operator has the form:

$$\frac{\sin x \cos y}{x}$$

Where x refers to the highest frequency in the band and y to the center frequency of the band. A low pass filter does not contain the cosine term. The highest recognizable "Nyquist" frequency in any image has a wavelength of two pixels and the lowest frequency that can be defined within the image by digital Fourier methods is directly related to the number of pixels in the operational subset. A band reject filter is obtained by subtracting the band pass filter from the all pass filter and a high pass by subtraction of the low pass filter. The resulting filter operator is then convolved with the matrix of values representing the image. On the other hand, the original image can be transformed to the frequency domain, the unwanted frequencies deleted and the inverse transform taken to display the filtered image. Spatial filters, whether they are in the distance or the frequency domain should retain at least an octave of image frequencies so that they will clearly display individual features without the presence of ringing side lobes. Where a number of frequency domain filters are to be tested, the frequency spectra of original images can be stored on disk, then quickly recalled and subjected to a suite of filters to obtain the optimum output quickly and easily. The filtered output can be displayed by itself, or it can be added to or subtracted from the original image to enhance the display of specific features.

Examples

Figure 1 is a portion of a Landsat Thematic band one scene from the Rock Springs area of western Wyoming. A series of east dipping Upper Cretaceous and Tertiary formations outcrop in a northerly direction while a fine grained network of drainage channels cross from east to west. Many of the valleys contain recent surficial deposits. Figure 2 illustrates the bedrock geology. Spatial filters can be used to enhance selectively either the geology or the drainage. A series of band pass filters, each covering approximately one octave of the frequency spectrum were applied to the image. Figure 3 illustrates the filtered output from a filter displaying only the low frequencies. The north-south trending geologic boundaries dominate the image and drainage patterns have been deleted. Figure 4 and Figure 5 are the outputs from higher stages in the suite of filters and add definition to the geologic features. Figure 6 displays the output from a still higher pass filter that begins to pick up some of the drainage pattern components as indicated by the presence of cross trends. The filtered images can be interpreted alone or they can be included with the original image to produce an enhanced overall picture. Image software usually contains a procedure that allows a percentage of one image to be added or subtracted from another which permits any desired degree of enhancement of the filtered subtle feature within the total picture. Figure 7 illustrates the effect of subtracting a percentage of the geology from the original image improving the clarity of the drainage display and Figure 8 has the filtered geology added to the original picture to assist the geological interpretation.

References

1. Cooley, J.W. and Tukey, J.W., "An algorithm for the machine calculation of complex Fourier series", Mathematics of Computers, Vol. 19, No. 90, pp297-301, 1965.
2. Zurflueh, E.G., "Application of two-dimensional linear wavelength filtering", Geophysics, Vol. 32, pp1015-1035, 1967.
3. Robinson, J.E., Charlesworth, H.A.K. and Ellis, M.J., "Structural analysis using spatial filtering in interior plains of south-central Alberta", Amer. Assoc. Petroleum Geologists Bull., Vol. 53, No. 11, pp2341-2367, 1969.

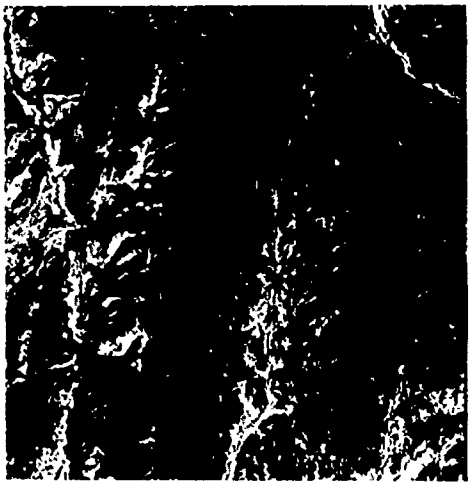


Figure 1. The original image, a 15 by 15 km subset of a Landsat Thematic band 1 scene of the Rock Springs area of Wyoming.

Figure 3. Band pass filtered image displaying only the large scale geologic features.

Figure 5. Band pass filtered image displaying smaller scale geologic features.

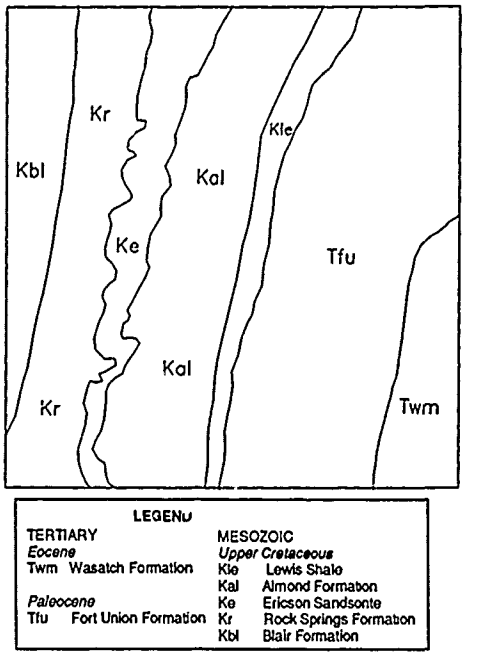


Figure 2. Bedrock geology of the Rock Springs area.

Figure 4. Band pass filtered image displaying medium scale geologic features.

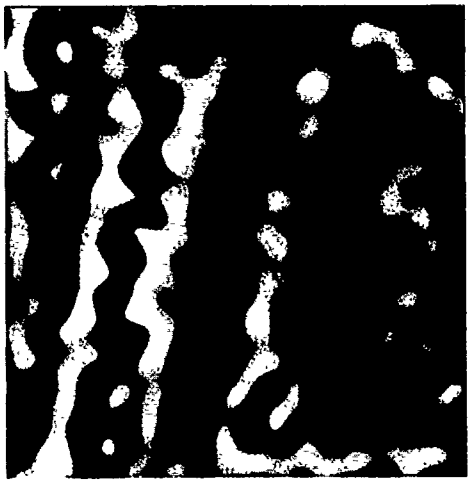


Figure 6. Band pass filtered image displaying both small scale geologic features and larger drainage features.

Figure 7. Original image with a portion of the geology deleted to enhance the drainage system.

Figure 8. Original image with a portion of the geology added to enhance the geologic presentation.

PERFORMANCE OF TWO TEXTURE-BASED CLASSIFIERS OF CLOUD FIELDS USING SPATIALLY AVERAGED LANDSAT DATA

S. K. Sengupta,[1] R. M. Welch[2] and M. S. Navar[1]

[1]Dept. of Math. & Computer Sci., [2]Inst. of Atmos. Sci., S.D. School of Mines & Technology, Rapid City, SD 57701

ABSTRACT

Using the Gray Level Difference Vector approach, classification accuracies with 1/8 km spatial resolution data are similar to those obtained using the full spatial resolution features. The implications are that there are no advantages to be gained in cloud classification accuracies by using even higher spatial resolutions obtained from LANDSAT Thematic Mapper or SPOT imagery. The optimum spatial resolution is 1/4 km. However, significant improvement in cloud classification accuracy is obtained using 1/2 km resolution data over that available from 1 km resolution of AVHRR and GOES imagery. Cirrus classification accuracy is especially compromised as spatial resolution is degraded. However, texture measures defined at the combination of pixel separations d = 1,4 improves classification accuracies by several percent even for 1 km spatial resolution data. Cirrus cloud classification accuracy is significantly improved by the use of multiple distance features.

In regard to the Max-Min Coocurrence Matrix approach, spatial distribution of thresholded consecutive extremes allow for the creation of new textural features that bring additional discriminating power to the classifier.

INTRODUCTION

There has been a great deal of interest in selection of an optimum spatial resolution for monitoring of global climate. Kong and Vidal-Madjar (1988) progressively degraded Landsat Multispectral Scanner (MSS) imagery to AVHRR resolutions. They found that MSS imagery can be linearly correlated with AVHRR data, but that the low resolution images miss the fine scale textural information. They suggested approximately 400 m spatial resolution for monitoring of land use patterns and crop growth. Townshend and Justice (1988) made a similar study by degrading MSS imagery, but used different spatial measures. Their results indicated that resolutions finer than 1 km are required for monitoring global change in vegetation. They suggested 500 m resolutions for the proposed Moderate Resolution Imaging Spectrometer (MODIS) being developed under the Earth Observing System (EOS) program. Other studies by Wielicki and Welch (1986) and Wielicki et al. (1986) have examined the effect of spatial resolution upon derived estimates of cloud fraction.

The present study examines texture-based cloud classification accuracy as a function of spatial resolution. Based on a single band of LANDSAT imagery, textural features at 1/16 km spatial resolution have been successfully employed for cloud field classification (Welch et al., 1988). These features were derived from the Gray Level Cooccurrence Matrix (GLCM) approach. Our recent work has used simplified textural measures based on vector methods such as the Sum and Difference Histogram (SADH) (Unser, 1986) and Gray Level Difference Vector (GLDV) approaches (Weszka et al., 1976). It was found that these methods have nearly equivalent accuracy while affecting significant savings in storage and computation time requirements.

For global monitoring of atmospheric and surface conditions, lower spatial resolutions on the order of 1-8 km generally are used. The present study examines the loss of cloud classification accuracy as a function of spatial resolution by degrading the imagery through progressive averaging. Textural measures are computed using two approaches: 1) GLDV and 2) Max-Min Cooccurrence Matrix (MMCM). Landsat MSS imagery is progressively degraded from a spatial resolution of 57 m to resolutions of 114 m, 228 m, 456 m, and 912 m.

TEXTURAL MEASURES

The GLDV approach is a statistical model which describes texture by statistical rules governing the distribution and relation of gray levels in local neighborhoods. More specifically, GLDV is based on the absolute differences between pairs of gray levels at a distance d apart at angle ϕ. The difference vector density function P(M) is defined for $M = |I-J|$, where I and J are the gray levels separated by distance d with relative orientation ϕ of the pixels. It is normalized by dividing the gray level frequencies of occurrence by the total frequencies.

The concept of gray level cooccurrence (Haralick et al., 1973) was extended by Davis et al. (1981) to what the authors called "generalized cooccurrence." Viewing image texture from a somewhat different perspective, Mitchell et al. (1977) proposed a new texture measure based on the statistical properties of significant local extremes in gray level. The Max-Min Cooccurrence Matrix (MMCM) approach is an extension of this notion which is based on textural features that are the statistical properties of the thresholded local extrema and their spatial separation d. These features are derived from the Max-Min cooccurrence matrix whose $(I,J)^{th}$ entry, computed for a fixed horizontal distance d represents the

frequency of the gray level pair (I,J) occurring at consecutive local extremes of opposite kinds.

Figure 1 shows a representative profile of gray levels in a cumulus image. Starting from a local minimum at the origin, the consecutive thresholded extremes are located at pixel separations of 0, 13, 17, 27, 33, 37, 43, 59, 73, 85, and 97, respectively. The dashed lines in Fig. 1 are the sliding thresholds used to locate the extrema (Hord, 1986). Although the importance of local extremes has been indicated by Mitchell et al. (1977) and implicitly by Haralick (1979), this approach has largely been overlooked in remote sensing applications.

From the normalized max-min cooccurrence matrices, the following textural measures were computed: number of pairs, contrast, angular second moment, correlation, entropy, local homogeneity, cluster shade, cluster prominence, and goodness of fit (Welch et al., 1987). The features were calculated as a function of the distance d between maxima and minima and for two thresholds, T=2 and 8.

RESULTS

Figure 2 shows the GLDV textural measures of standard deviation, local homogeneity, and entropy at five spatial resolutions ranging from 1/16 km to 1 km (912 m). The textural measures are calculated at pixel separations of d = 1, 2, 4, 8, 16, 32, and 64. At full spatial resolution, this corresponds to distances of 57 m, ..., 3.6 km, while at 1/8 km resolution these separations correspond to distances of 114 m, ..., 7.3 km, and so on.

A total of 37 Landsat MSS scenes were used in this study, 15 stratocumulus, 10 cumulus, and 12 cirrus. Each of the 37 cloud scenes is subdivided into 20 subregions. Figure 2 shows selected textural measures for a stratocumulus scene.

Textural measures for spatial resolutions of 1/16 km and 1/8 km are nearly identical. Therefore, classification results based upon texture measures are expected to be unaffected by using 1/8 km resolution data. However, there is substantial change in texture measures as it is further degraded to 1/4 km and below. Clearly, there is considerably greater homogeneity and loss of fine-scale textural information as the scene is spatially degraded.

Figure 3 shows classification accuracy for stratocumulus, cumulus, and cirrus as a function of spatial resolution and pixel separation. Cumulus has high classification accuracy and is relatively less sensitive both to pixel separation and to spatial resolution. Cirrus shows strong sensitivity both to pixel separation and to spatial resolution. Overall classification accuracies decrease with decreasing spatial resolution. Therefore one can expect better cloud classification accuracies with the proposed 1/2 km spatial resolution (EOS) MODIS imagery than with current 1 km spatial resolution imagery available from AVHRR and GOES. Further details are given by Welch et al. (1989).

Figure 4 shows representative MMCM textural features of entropy, contrast, and number of points as a function of distance between maxima and minima and of spatial resolution. With exception of the number of points, the textural measures at spatial resolutions of 1/16 km and 1/8 km are similar. Therefore, classification accuracies at the degraded 1/8 km resolution are only slightly less than at full resolution. However, as resolution is further degraded to 1/4 km, 1/2 km, and 1 km, Fig. 4 shows significant variation as a function of distance of the various textural features. As with the GLDV results, there is progressive loss in classification accuracy as the imagery is degraded in spatial resolution.

The distinctive shapes of the texture curves makes classification possible between the various cloud classes. Combining textural measures at various pixel separations leads to improvement in classification accuracies by up to several percent. One potential advantage of the MMCM approach is that it should allow better discrimination of cloud subspecies (e.g., cirrus, cirrocumulus, cirrostratus) than the GLDV or gray level cooccurrence approaches.

Acknowledgments. This investigation was conducted jointly under National Science Foundation Grant ATM-8507918 and National Aeronautics and Space Administration Grant NAG-1-542. Appreciation is extended to Joie Robinson for typing this paper.

REFERENCES

1. Davis, L.S., M. Clearman and J.K. Aggraval, "An empirical evaluation of generalized cooccurrence matrices", IEEE Trans. Patt. Anal. Mach. Intell., Vol. 3, pp214-221, 1981.

2. Haralick, R.M., "Statistical and structural approaches to texture", Proc. IEEE, Vol. 67, pp786-804, 1979.

3. _____, K.S. Shanmugam and I. Dinstein, "Textural features for image classification", IEEE Trans. Syst. Man Cybern. SMC3, pp610-621, 1973.

4. Kong, X.N., and D. Vidal-Madjar, "Effet de la résolution spatiale sur des propriétés statistiques des images satellites: une étude de cas", Intl. J. Rem. Sens., Vol. 9, pp1315-1328, 1988.

5. Mitchell, O.R., C.R. Myers and W. Boyne, "A max-min measure for image texture analysis", IEEE Trans. Comp., Vol. 26, pp408-414, 1977.

6. Townshend, J.R.G., and C.O. Justice, "Selecting the spatial resolution of satellite sensors required for global monitoring of land transformations", Intl. J. Rem. Sens., Vol. 9, pp187-236, 1988.

7. Unser, M., "Sum and difference histograms for texture classification", IEEE Trans. Patt. Anal. Mach. Intell., Vol. 8, pp118-125, 1986.

8. Welch, R.M., S.K. Sengupta and D.W. Chen, "Classification of cloud fields based on textural characteristics", SPIE Conf. on Dig. Image Proc. and Vis. Comm. Tech. in Meteor., 26-29 Oct, Boston, MA, 1987.

9. _____, _____, and _____, "Cloud field classification based upon high spatial resolution textural features. Part 1: Gray level cooccurrence matrix approach", J. Geophys. Res., Vol. D93, pp12663-12681, 1988.

10. _____, M.S. Navar and S.K. Sengupta, "Cloud field classification based upon high spatial resolution textural features. Part 4. The effect of spatial resolution". [Submitted to J. Geophys. Res. - Atmos., 1989]

11. Weszka, J.S., C.R. Dyer and A. Rosenfeld, "A comparative study of texture measures for terrain classification", IEEE Trans. Syst. Man Cybern., Vol. 6, pp269-285, 1976.

12. Wielicki, B.A. and R.M. Welch, "Cumulus cloud field properties derived using Landsat digital data", J. Clim. Appl. Meteor., Vol. 25, pp261-276, 1986.

13. _____, P. Minnis and L. Parker, "The effect of spatial resolution on satellite derived cloud cover", Preprints Sixth Conf. Atm. Rad., Williamsburg, VA, Amer. Meteor. Soc., pp301-305, 1986.

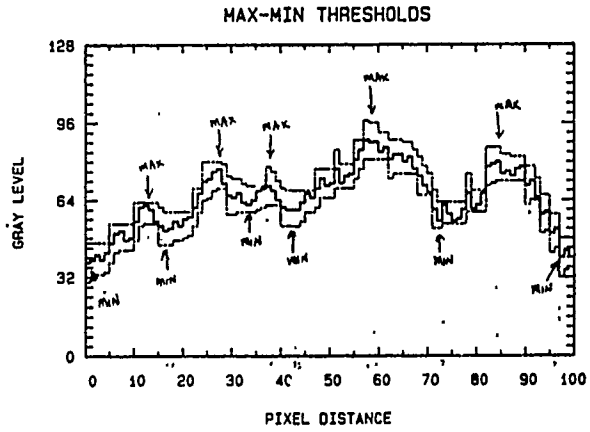


Fig. 1: Operation of the MAX-MIN algorithm for a representative cumulus cloud segment.

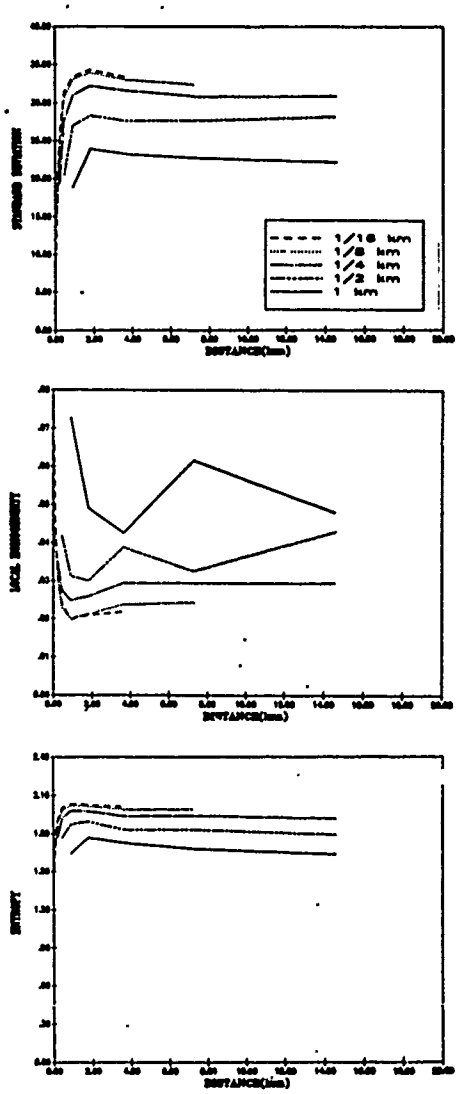


Fig. 2: GLDV textural measures of standard deviation, local homogeneity, and entropy as a function of pixel separation and spatial resolution.

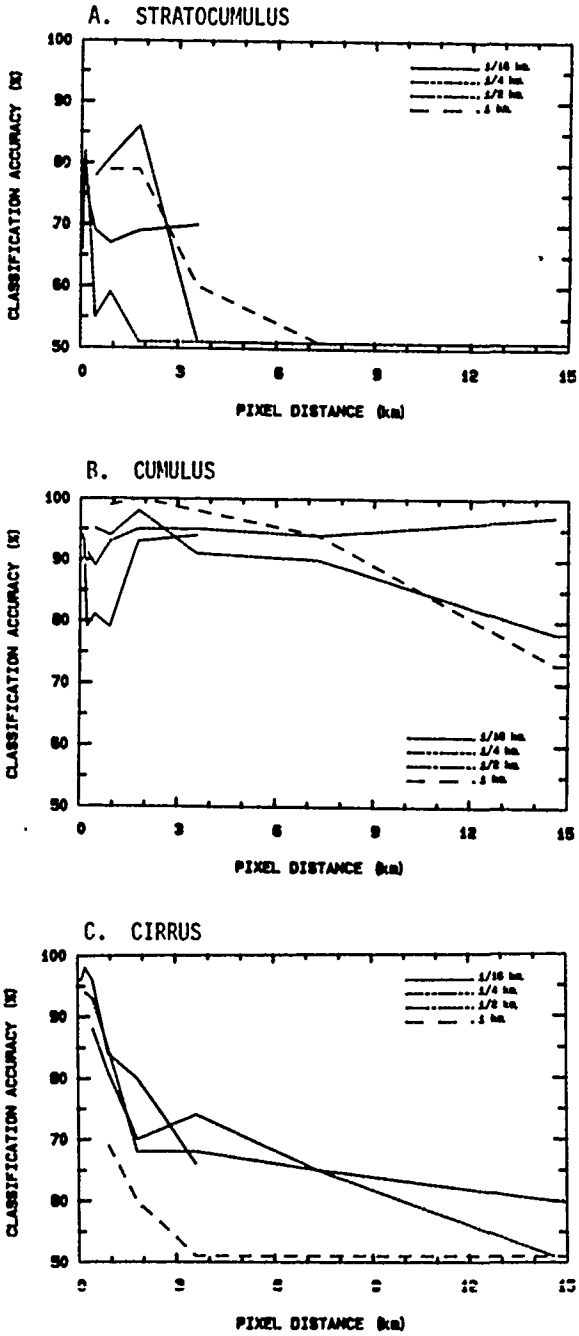


Fig. 3: GLDV classification accuracy for stratocumulus, cumulus, and cirrus as a function of pixel separation and spatial resolution.

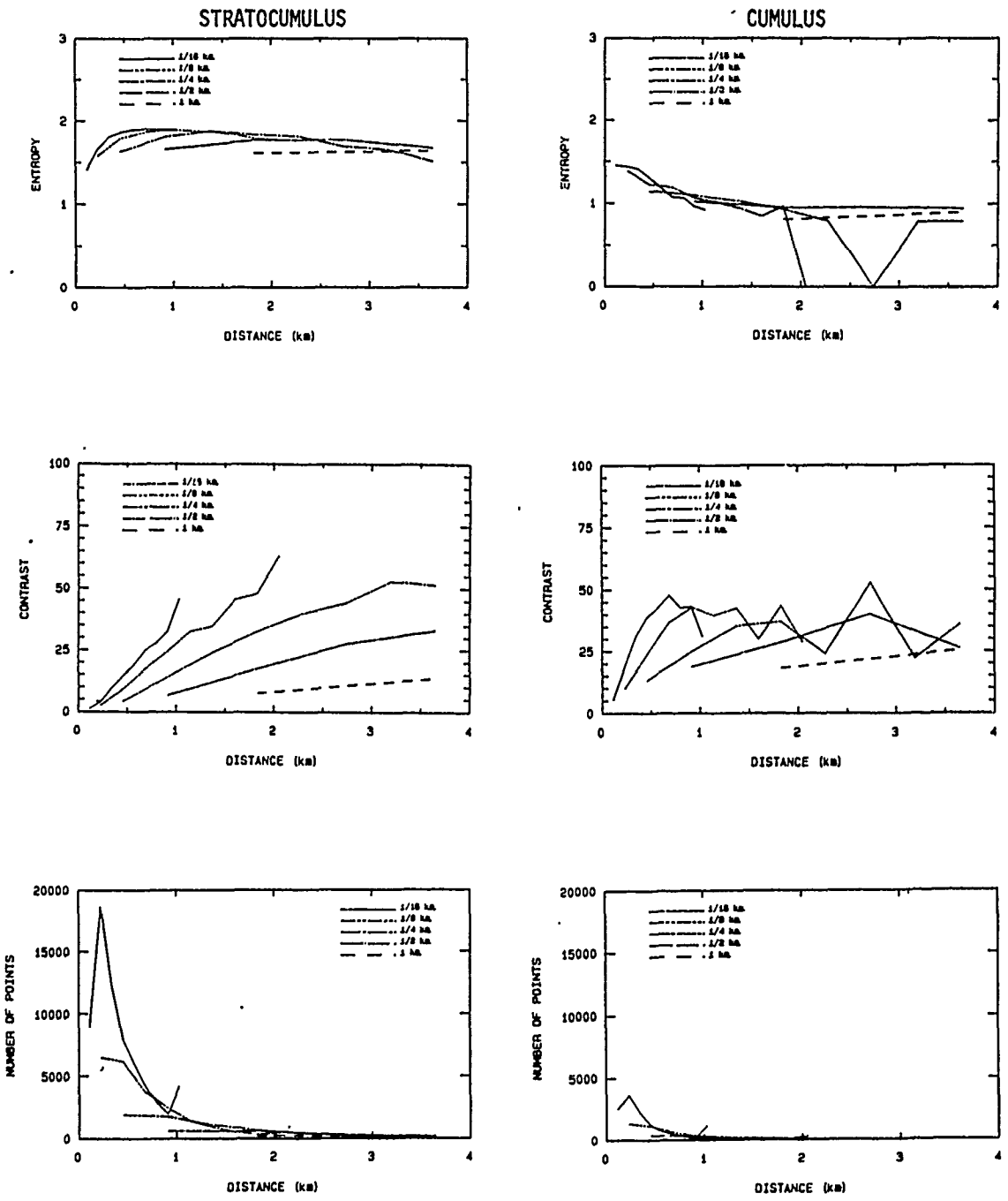


Fig. 4: MMCM textural measures of entropy, contrast, and number of points as a function of max-min separation and spatial resolution for representative stratocumulus and cumulus clouds.

IMAGE SEGMENTATION BY ITERATIVE PARALLEL REGION GROWING AND SPLITTING

James C. Tilton

Mail Code 636
NASA Goddard Space Flight Center
Greenbelt, MD 20771

ABSTRACT

Earlier reports (Tilton, 1984, Tilton, 1988; Tilton and Cox, 1983) discussed the "Spatially Constrained Clustering" (SCC) iterative parallel region growing technique, and its potential application to data compression and image analysis. The SCC algorithm is implemented on the Massively Parallel Processsor (MPP) at the NASA Goddard Space Flight Center.

Most previous region growing approaches have the drawback that the segmentations produced depend on the order in which portions of the image are processed (*e. g.* Schachter, *et al*). An ideal solution to this problem would be to merge only the single most similar pair of spatially adjacent regions in the entire image in each iteration. However, this ideal approach becomes impractical except for very small images, even when implemented on a massively parallel computer. The SCC algorithm overcomes these problems by performing, in parallel, the best merge within each of a set of local, possibly overlapping, subimages throughout the image.

A region splitting stage is also incorporated into to the algorithm. However, experimental tests show that region splitting generally does not improve segmentation results. The SCC algorithm has been tested on various imagery data, and test results from a Thematic Mapper image is summarized at the end of this paper.

Keywords: Image segmentation, Image Analysis, Data parallel analysis.

INTRODUCTION

Segmentation is the process of partitioning images into constituent parts called regions using image attributes such as pixel intensity, spectral values, and/or textural properties. Image segmentation produces an image representation in terms edges and regions of various shapes and interrelationships.

Image segmentation is a key step in many approaches to data compression and image analysis. An optimal coding of an image segmentation, such as through a region label map and region feature file, can be used to effect data compression (Tilton and Ramapriyan, 1988). Image analysis can be performed on an image segmentation by using the shape, texture, spectrum, etc. of the regions found by the image segmentation and interrelationships between the regions. This region based analysis of imagery is potentially more effective than pixel based analysis, because region based analysis exploits spatial information whereas pixel based analysis does not. It is also potentially more effective than spatial analysis based on analyzing pixels in areas defined by a fixed grid Grid based analysis essentially utilizes an arbitrarily imposed segmentation of the image, whereas region based analysis utilizes a segmentation which is derived from characteristics of the image data.

Most image segmentation approaches can be placed in one of three classes: (i) characteristic feature thresholding or clustering, (ii) boundary detection, and (iii) region extraction. Characteristic feature thresholding or clustering is often ineffective because it does not exploit spatial information. Boundary detection does exploit spatial information through examining local edges found throughout the image. For simple noise-free images, detection of edges results in straightforward boundary delineation. However, edge detection on noisy, complex images often produces missing edges and extra edges which cause the detected boundaries to not necessarily form a set of closed connected curves that surround connected regions. One way to overcome this problem is to combine region extraction and boundary detection. Manohar, *et al* (1988) report on some experiments in combining boundary detection approaches with the iterative parallel region growing approach discussed here. An early attempt at fully integrated edge detection and region growing is reported by Latty (1984).

ITERATIVE PARALLEL REGION GROWING AND SPLITTING

Early approaches to region extraction (usually by region growing) had the disadvantage that the regions produced depended on the order in which portions of the image are processed. But Schachter, *et al* (1979) suggest that implementing region growing as "an iterative parallel process" would overcome the order dependent problem. An ideal solution to this problem would be an iterative parallel process that merged the single most similar pair of spatially adjacent regions over the entire image at each iteration. This is the approach suggested by Tilton and Cox (1983) and by Beaulieu and Goldberg (1989). However, this ideal approach becomes impractical for all but very small images, even when implemented on a massively parallel computer.

A compromise solution to this problem is to merge a selected set of region pairs, in parallel, each iteration. The problem then reduces to finding an effective and efficient approach to select some best

set of region pairs to merge at each iteration. The current version of the SCC algorithm selects the set of region pairs to be merged at each iteration by choosing the best merge within overlapping subimages centered on each region, or on each pair of spatially adjacent regions. The specification of a subimage and this selection process will be given in the following sections.

Conventional wisdom in the image analysis field seems to be that all region growing algorithms must have a splitting phase in order to be as effective and/or efficient as possible. Our empirical studies show, however, for cases where the subimages controlling the region growing process are of sufficient size, region splitting rarely occurs, and when it does, it has no discernable effect on the segmentations produced. We report these results in more detail in a later section.

DEFINITION OF A SUBIMAGE WITH RESPECT TO A REGION

A subimage with respect to a particular region can be defined recursively as follows: A level 0 subimage for any region is the empty set. A level 1 subimage, with respect to a region, is the region itself. A level 2 subimage, with respect to a region, is the level 1 subimage, with respect to that region, plus all regions that are spatially adjacent to the level 1 subimage. Finally, a level n subimage, with respect to a region, is the level n-1 submimage, with respect to that region, plus all regions that are spatially adjacent to the level n-1 subimage. See figure 1 for examples.

SPATIALLY CONSTRAINED CLUSTERING (SCC) ALGORITHM

We call our region growing algorithm the Spatially Constrained Clustering (SCC) algorithm, since it is essentially a technique for clustering data in which only spatially adjacent regions can group together at a particular iteration. We could have just as well called our algorithm Iterative Parallel Region Growing, but retain the SCC name to maintain continuity with our earlier work on essentially similar algorithms (Tilton and Cox, 1983; Tilton, 1984; Tilton and Ramapriyan, 1988; and Tilton, 1988).

The basic SCC algorithm, which has essentially remained unchanged from our earliest work, is a follows:

i. Initialize the segmentation process by labeling each pixel as a separate region.
ii Calculate a similarity criterion between each pair of spatially adjacent regions.
iii. Merge pairs of regions that meet the merge constraints.
iv. Check for convergence. If converged, stop. Otherwise return to step ii.

In detail, the SCC algorithm has evolved over the years. The main elements that have changed are the similarity criterion (in step ii) and the merge constraints (in step iii). The convergence criterion (in step iv) have changed as dictated by changes in the similarity criterion.

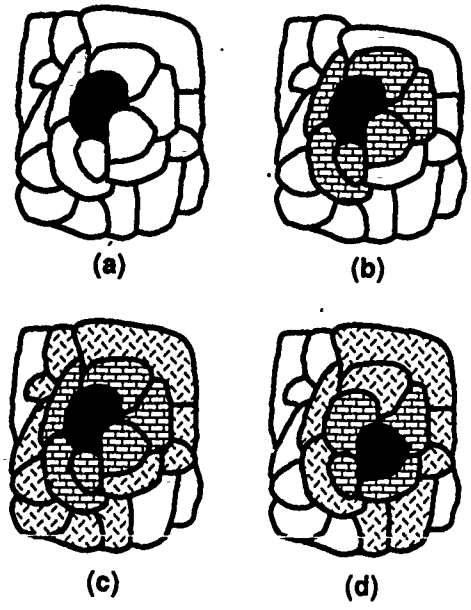


Figure 1. (a) Level 1 subimage for region R, (b) level 2 subimage for region R, (c) level 3 subimage for region R, and (d) level 3 subimage for (a different) region R.

SIMILARITY CRITERION

The original SCC algorithm used a similarity criterion based on a combination of a pair of statistical hypothesis tests. Both tests consider whether a pair of regions come from the same probability distribution. The first test checks the mean values of the regions, assuming unequal region variances, and the second test checks the variances of the regions (Tilton and Cox, 1983). (Beaulieu and Goldberg (1989) use a means test that assumes the pair of regions tested have identical region variance. This assumption is almost universally violated for many types of imagery, *e.g.* remotely sensed imagery.)

Due to program memory restrictions on the MPP, a simpler similarity criterion had to be developed for the MPP implementation of the SCC algorithm. Thus we have developed a *dissimilarity* criterion based on minimizing the increase in variance normalized mean squared error (Tilton and Ramapriyan, 1988). For completeness, we include a brief derivation of this dissimilarity criterion here. (This is a dissimilarity criterion rather than a similarity criterion because it increases with increasing dissimilarity between a pair of regions.)

The Mean Square Error (MSE) of the k^{th} band a multiband image is defined as

$$MSE_k = E[(D_k - D_k^r)^2] \cong \frac{1}{N-1} \sum_{p=1}^{N} (D_{kp} - D_{kp}^r)^2 \qquad (1)$$

where D_k and D_k^r are the data values of the k^{th} band of the original and region mean images, respectively; D_{kp} and D_{kp}^r are the values of the p^{th} pixel of the k^{th} band of the original and reconstructed images, respectively; E denotes the expected value; and N is the total number of pixels in the image. A region mean image is formed by substituting the mean vector of each region for the multispectral radiance values of each pixel in the region.

The variance normalized mean squared error for the k^{th} band ($NMSE_k$) is defined as

$$NMSE_k = \frac{MSE_k}{VAR_k} \qquad (2)$$

where VAR_k is the variance of the k^{th} band. For a particular pair of spatially adjacent regions, $\Delta NMSE_k$ is the change in $NMSE_k$ when the pair of regions is merged and the new region mean image is formed and compared to the original image. The dissimilarity criterion used in the MPP implementation of the SCC algorithm is the $\max_k(\Delta NMSE_k)$ for each pair of spatially adjacent regions, where the maximum is taken over all bands ($1 \le k \le m$). (Optionally, the dissimilarity criterion can be taken as $\sum_{k=1}^{m}(\Delta NMSE_k)$.)

The change in $NMSE_k$, or $\Delta NMSE_k$, is calculated as follows:

$$\Delta NMSE_k = \frac{MSE_k^c - MSE_k}{VAR_k} \qquad (3)$$

where MSE_k^c is the mean squared error when regions i and j are merged, while MSE_k is the mean squared error before regions i and j are merged. Using the definitions of MSE_k and the region mean, it is easy to derive a more fundamental version of equation (3), *viz*

$$\Delta NMSE_k = \frac{n_i(\bar{D}_{ki}-\bar{D}_{ki,j})^2 + n_j(\bar{D}_{kj}-\bar{D}_{ki,j})^2}{(N-1)VAR_k} \qquad (4)$$

where n_i and n_j are the number of points in regions i and j, respectively, before combining, and N is the number of points in the image. $\bar{D}_{ki}$ and $\bar{D}_{kj}$ are the mean values of band k for regions i and j, respectively, before combining, and $\bar{D}_{ki,j}$ is the mean value of band k for the region that would result from combining regions i and j.

MERGE CONSTRAINTS

The original SCC algorithm simply merged the single most similar pair of spatially adjacent regions over the entire image in each iteration (Tilton and Cox, 1983). This is also the approach taken by Beaulieu and Goldberg (1989). As noted earlier, this ideal approach is impractical for all but very small images, even when implemented on a massively parallel computer. However, in light of the persuasive theoretical justification given by Beaulieu and Goldberg (1989) in favor of performing the best single merge per iteration, we have sought to develop an approximation of this approach.

The following scheme should closely approximate the one merge per iteration approach. At each iteration first divide the image into a set of, possibly overlapping, subimages, and then perform the best merge within each subimage. After each iteration the set of subimages would be redefined appropriately. At early iterations, when the individual regions are very small, the subimages would be relatively small. At later iterations, when the average region size is larger, the subimages would also be larger. For the final iterations the subimages may equal the entire image. For large enough subimages, this approach should closely approximate the ideal one merge per iteration approach.

A suitable massively parallel implementation of this scheme would allow concurrent performance of all the merges at each iteration. Depending on the size of the image and subimages, up to several thousands of merges to occur, in parallel, during the initial iterations.

In order to better describe how the current version of the SCC algorithm implements this scheme, we first define merge constraint levels. Constraint level n signifies that merges are constrained to be the best merge within the union of the level n subimages with respect to each of the potentially merging region pairs. Constraint level n-0.5 signifies merges are constrained to be the best merge within the level n subimage with respect to only one of the potentially merging region pairs. (By the definition of a level n subimage, the level n-1 subimage of a region is contained within the level n subimage of a neighboring region. Thus, for a level n-0.5 merge constraint, one of the merging regions is the best merge within it's level n subimage, and the other is the best merge within it's level n-1 subimage. The level n-0.5 merge constraint comes from the average level of the two subimages controlling the merging process.) See figure 2 for a graphical description of the subimages over which best merges are performed at selected constraint levels.

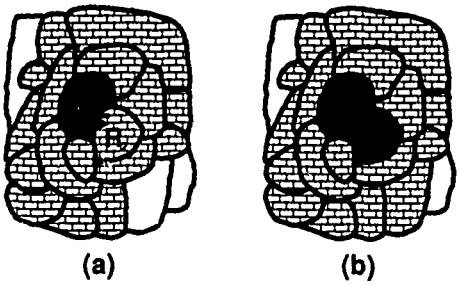


Figure 2. Controlling subimages for (a) merge constraint level 2.5, with respect to region R_1, and (b) merge constraint level 3.0, with respect to regions R_1 and R_2.

Employing merge constraint level 0.5 in the SCC algorithm is equivalent to performing the best merge for each region, without regard to what would be the best merge for any neighboring region. Note that this constraint level allows strings of regions to merge together at each iteration. A combination of a 0.5 merge constraint level with an ad hoc global threshold is the merge constraint used by an earlier version of the SCC algorithm (Tilton, 1988).

Willebeek-LeMair and Reeves (1988) describe a region growing process similar to the SCC algorithm. Besides the similarity functions and region features utilized, the essential difference between the SCC region growing process as described in Tilton (1988) and the region growing process described by Willebeek-LeMair and Reeves is a requirement that all merges be pairwise mutually best for the regions merged at any iteration. Under this restriction, only pairs of regions merge at any given iteration. When this pairwise mutually best constraint was added to the SCC algorithm it was found that the ad hoc dissimilarity threshold could be raised at a faster rate, thus achieving faster convergence, while maintaining the quality of the resulting segmentations. Under the merge constraint level scheme, this is equivalent to a combination of a 1.0 merge constraint level with an ad hoc global threshold in the SCC algorithm. (However, the merge constraint level 1.0 scheme does not break ties as does the pairwise mutually best constraint scheme. In this scheme it is possible for a region to merge with two other regions if its dissimilarity function value versus those two regions is identical.)

The current version of the SCC algorithm can also perform region growing based on merge constraint levels higher than 1.0, in addition to performing

region growing based on merge constraint levels 0.5 and 1.0.

The SCC algorithm also employs an ad hoc overall merge threshold. If a merge's dissimilarity function value is not better or equal to this overall merge threshold, it will not be executed, even if it is the best merge in its subimage. The overall merge threshold currently used by the SCC algorithm is as follows:

$$\text{Thresh} = \frac{ME - NMSE}{(1/2\ NR) * FR} \qquad (5)$$

where $NMSE = \sum_{i,j}(\max_{k}(NMSE_{kij}))$, k=1 to m, i=1 to N_1, j=1 to N_2 (m = number of bands, N_1 = number of rows, N_2 = number of columns), ME (a user defined parameter) is the maximum allowable value for NMSE, NR is the number of regions in the image at the current iteration, and FR (another user defined parameter) is the estimated fraction of regions participating in merges each iteration. This overall threshold generally does not affect the merges performed until the final few iterations

PIXEL SPLITTING

Conventional wisdom in the image analysis field seems to be that all region growing algorithms must have a splitting phase in order to be most effective and/or efficient. Until recently, we also accepted this conventional wisdom, and had for some time planned to add a pixel splitting stage to the SCC region growing algorithm with the hope of further improving results.

In our pixel splitting stage all pixels in all regions are examined to determine whether or not the overall region feature values have drifted away from the feature values of individual pixels within each region. Pixels are split out from regions that have become too dissimilar to them by relabeling these pixels as individual regions. The split out pixels may join with another neighboring pixel or region in the next ensuing region growing step.

Empirical studies of this region growing/pixel splitting combination showed, however, that pixel splitting rarely occurred, and had no discernable effect on the segmentations produced if merge constraint level was 1.0 or higher.

EXPERIMENTAL RESULTS

A 128-by-128 subset of a 7-band Landsat Thematic Mapper (TM) image collected over the USGS Ridgely, Maryland Quadrangle was used as a test data set for this study. Test results are given in Table 1.

Table 1 shows that for this case, pixel splitting never improved the segmentation results, with the only discernable difference in results being at constraint level 0.5. Increasing the constraint level generally improves the segmentation result (decreases the value of NMSE). The only exception to this is increasing the constraint level from 1.5 to 2.0 actually *increased* the measured value of NMSE. However, a visual inspection of the segmentation results shows that the segmentation produced with constraint level 1.5 merges a large agricultural field together with a large forested area. The segmentation produced with constraint level 2.0 does not merge any large agricultural fields with large forested areas or any other ground cover type. Thus the visual quality of the segmentation with constraint level 2.0 seems to be better. An improved quantitative error measure is apparently needed. Similar results have also been obtained with other TM images, Synthetic Aperture Radar (SAR) imagery, and a simulated data set.

Table 1. Test results obtained for the test image for various merge constraint levels, with and without pixel splitting, for ME = 1.0 and FR = 0.5.* (Results given for iteration where the number of regions equals 20).

Constraint Level	Pixel Splitting	NMSE*	Number of Iterations
0.5	NO	0.989	6
0.5	YES	1.008	5
1.0	NO	0.892	36
1.0	YES	0.892	36
1.5	NO	0.846	85
2.0	NO	0.857	184

* ME, FR and NMSE as defined in Equation 5.

REFERENCES

1. Batcher, K. E., "Design of a massively parallel processor", IEEE Trans. on Computers, Vol. C-29, No. 3, pp. 836-840, 1980.
2. Beaulieu, J.-M., and M. Goldberg, "Hierarchy in picture segmentation: a stepwise optimization approach", IEEE Trans. on Pattern Analysis and Machine Intelligence, Vol. 11, No. 2, 1989.
3. Chen, T. M., D. H. Staelin and R. B. Arps, "Information content analysis of Landsat image data for compression", IEEE Trans. on Geoscience and Remote Sensing, Vol. GE-25, No. 4, pp. 499-501, 1987.
4. Latty, R. S., "Scene segmentation through region growing", Proc. of the Tenth International Symposium on Machine Processing of Remotely Sensed Data, West Lafayette, IN, June 12-14, 1984.
5. Manohar, M., H. K. Ramapriyan and J. P. Strong, "Parallel algorithms for determining motion vectors in ice floe images by matching edge features", Proc. of the 2nd Symposium on the Frontiers of Massively Parallel Computation, Fairfax, VA, Oct. 10-12, 1988.
6. Tilton, J. C., "Multiresolution spatially constrained clustering of remotely sensed data on the Massively Parallel Processor", Digest of the 1984 Int'l Geoscience and Remote Sensing Symposium, Strasbourg, France, pp. 661-666, Aug. 27-30, 1984.
7. Tilton, J. C., "Image segmentation by iterative parallel region growing with applications to data compression and image analysis", Proc. of the 2nd Symposium on the Frontiers of Massively Parallel Computation, Fairfax, VA, Oct. 10-12, 1988.
8. Tilton, J. C., and S. C. Cox, "Segmentation of remotely sensed data using parallel region growing", Digest of the 1983 Int'l Geoscience and Remote Sensing Symposium, San Francisco, CA, pp. 9.1-9.6, Aug. 31-Sept. 2, 1983.
9. Tilton, J. C., and H. K. Ramapriyan, "Data compression experiments with Landsat Thematic Mapper and Nimbus-7 Coastal Zone Color Scanner data", Proc. of the Scientific Data Compression Workshop, Snowbird, UT, May 3-5, 1988.
10. Schachter, B. J., L. S. Davis and A. Rosenfeld, "Some experiments in image segmentation by clustering of local feature values", Pattern Recognition, Vol. 11, No. 1, pp. 19-28, 1979 .
11. Willebeek-LeMair, M., and A. P. Reeves, "Region growing on a highly parallel mesh-connected SIMD computer", Proc. of the 2nd Symposium on the Frontiers of Massively Parallel Computation, Fairfax, VA, Oct. 10-12, 1988.

REVIEW ON IMAGE ANALYSIS WITH MATHEMATICAL MORPHOLOGY IN REMOTE SENSING

G. FLOUZAT

Centre d'Etude Spatiale des Rayonnements
Centre National de la Recherche Scientifique, Université Paul Sabatier
9. Avenue Colonel Roche - 31029 Toulouse Cédex

ABSTRACT

This synthesis paper proposes to show the principal contributions of Mathematical Morphology (MM) to extracting information from an image.

The first part reviews the different steps of knowledge extraction according to a simulation model of visual perception and understanding. The corresponding operations in digital image processing and the place of MM are thus organized from low level to high level functions.

The second part describes the principal operations of MM and the properties of morphological operators.

The third part presents, for the low level functions, the interest of the alternated sequential filters for noise removal and textural feature extraction. In edge detection, skeletons are efficient for zero-crossing location in early vision modeling. Directionnal information is also accessible with a good choice of structuring elements.

The fourth part shows the connex components level where, after a segmentation or a classification, filtering on length and size parameters are very useful in both local and global measurements.

The fifth part concerns high level functions in structure quantification of objects image. Two families of functions are described : global measurements and symbolic description by graph construction.

The paper describes in the three latest parts methods already implemented but also some new algorithms ideas.

Key words : Mathematical Morphology, Image analysis.

INTRODUCTION

The use of Mathematical Morphology in remote sensing imagery processing is a methodology recently introduced if we consider other approaches now very well known in image analysis. For example, local statistics or global histogram handling are traditional technics in the digital imagery field. The idea of morphological analysis is interesting because it is a complementary point of view in comparison with others methods: it consists of a geometrical approach of image processing using set operators such as union, intersection and complementation. The different operations are generated by a structuring element, defined by a particular shape of pixels, acting on the objects independently of there size, shape and grey level.

I - THEORY OF ANALYTICAL INTERPRETATION

Knowledge extraction on a scene by visual analysis of an image is an efficient process which is commonly used in many application fields. The obvious results obtained by this way allows to think that human perception and understanding is a good model of methodology in digital image processing (BARLOW, 1972 MARR, 1980). In remote sensing, this idea is not new and, in this approach, different authors have produced numerous basis with respect to photointerpretation for many years (COLWELL, 1965 ; GUY, 1970). Thus, reference to the human model is very often useful to build strategy in image analysis ; we can use it to organize the tasks which give access to the different kinds of characteristics present in an image (FLOUZAT, 1982 ; WANG, 1983).

To follow this reasoning, the theory of analytical photo-interpretation (FLOUZAT, 1982) is a contribution to order the description of an image in elementary steps. In short, it must define the relationships existing between the penomena studied, their representation in the image and the corresponding digital processing for automatized feature and knowledge extraction. Three main parts are therefore necessary to set data processing functions in order to perform an advanced image analysis (FLOUZAT, 1988).

The first one includes the low level functions to map out raw primal sketch (MARR, 1980) with textural, geometric and spectral primitive features. The second one consists of objects extraction and characterization functions to make up the continuous sketch of the photointerpreters with significant textural elements (almost equivalent to segmentation and measurements on connected components). The third one is composed of the high level functions used to obtain spatial arrangement quantifications corresponding to the attempt at doing an help for realistic interpretation.

In this paper, morphological methods in remote sensing are presented in relation with these three functions categories.

II - MATHEMATICAL MORPHOLOGY

Mathematical morphology is a concept of image processing based upon set theoretical aspects principally developed at the Ecole des Mines (Paris) in the 1960's and described in MATHERON (1975), SERRA (1982, 1986). In a digital image, the set P of the positions and the set T of the grey tones defined an image function f (x,y) with $P \subset \mathbb{N}^2$ and $T \subset (\mathbb{N} \cap [0,255])$. The resulting image I, one byte encoded, is made up of a set of parts $\mathcal{P}(I)$ composed of all the subsets X. Morphological operations consist of transforming sets $X \in \mathcal{P}(I)$ into $\Phi(X)$ by action of structuring elements B modifying X by union, intersection or complementation. The structuring element B acts as a geometric probe whose shape and dimension allow the extraction of particular forms when a translate of B is contained in a part of a set X or its complement. Thuqs the geometrical properties of B (symetry, convexity, orientation, shape, size) entail an a priori knowledge of the characteristics of the extracted parts of X.

Erosion and Dilation

Let $X \in \mathcal{P}(I)$ be a set, x a point of X, $B \subset P$ a structuring element and B_x the translated of B whose origin is x. The eroded set $X \ominus B$ of X with respect to B is the set of points x such that B_x is included in X :

$$X \ominus B = \{ x \mid B_x \subset X \}$$

The erosion can be expressed also in translation terms : the intersection of all the translated of X by Ob define the eroded set when Ob described the whole structuring element B :

$$X \ominus B = \bigcap_{b \in B} X_{Ob}$$

In the same way, the erosion of a grey tone function f by B is the function $g = f \ominus B$ defined by the incersection of the translate of f. At any point x, g(x) is given by the minimum of these translated functions :

$$g(x) = \inf_{y \in B_x} \{ f(y) \} \qquad \forall x \in P$$

Taking into account the duality with respect to complementation, the morphological operation called dilation (noted $\oplus$) is the dual operation of erosion:

$$X \oplus B = (X^c \ominus B)^c$$

In translation terms :

$$X \oplus B = \bigcup_{b \in B} X_{Ob}$$

And because of the duality with respect to erosion, the dialted set of X is defined also as the following :

$$X \oplus B = \{ x \mid B_x \cap X \neq \emptyset \}$$

For a grey tone function, the dilation of f is also the dual operation of erosion :

$$g(x) = \sup_{y \in B_x} \{ f(y) \}$$

Sequential operations of erosion and dilation are interesting : the ultimate eroded point of X can be used as a mark of the origin object and opens possibilities for counting labeling technics. A sequence of dilations of set X by B restricted to an Y set is called conditional dilation of X in Y. The result is the filling of Y fron X. Note also that a structuring element can be considered to be the union of several simpler structuring elements. Erosion (dilation) by such an element is equivalent to successive erosions (dilations) by the subelements.

Opening and closing

As the eroded set of X by B is the set of origins x of B_x included in X, the opening X_B of X is :

$$X_B = \bigcup_{x \in (X \ominus B)} B_x$$

and as $\bigcup_{x \in X} B_x = X \oplus B$, we have $X_B = (X \ominus B) \oplus B$.

Opening can be also expressed in terms of unions and intersections of translates of the set X.
In a grey tone image, the opening of a function f by B is the function g such as :

$$g(x) = [(f \ominus B) \oplus B](x)$$

$$= \sup_{y \in B_x} \left\{ \inf_{z \in B_y} (f(z)) \right\} \qquad \forall x \in P$$

As erosion and dilation are dual operations with respect to complementation, the morphological closing (X^B) is the dual operation of closing :

$$X^B = (X \ominus B) \oplus B$$

For a grey tone function f, the closing g is the following function

$$g(x) = [(f \oplus B) \ominus B](x)$$

$$= \inf_{y \in B_x} \left\{ \sup_{z \in B_y} (f(x)) \right\} \qquad \forall x \in P$$

The most important property of opening and closing is idempotence. And, as the two are increasing operations and as opening is antiextensive and closing is extensive, it is possible to know ridge and valleys of a function at the first iteration : this is obtained by comparison between original and transformed function independently of the grey level.

An other interesting property comes from a sequence of openings with an increasing size of structuring element. Thus, it allows the extraction of bigger and bigger objects (binary case) or of bigger and bigger ridge by difference functions at each iteration and regardless of the local shape or the local signal intensity.

Morphological filters

Opening and closing can generate many filters in relation with their propert'es and their order in processing sequences. Morphological openings and closings have the following properties :

$f_B \leqslant f \leqslant f^B$ (1),

$f \leqslant g \Rightarrow f_B \leqslant g_B$ and $f^B \leqslant g^B$ (increasing) (2),

$(f_B)_B = f_B$ and $(f^B)^B = f^B$ (idempotence) (3).

Generally, one calls an algebraic opening (resp, algebraic closing) any transformation which satifies the (1), (2) and (3) properties. If only (2) and (3) are satisfied, the transformation is called a morphological filter. It can be proved (MATHERON, 1975) that every algebraic opening $\Psi(f)$ is the Sup of morphological opening f_B where B is the set of all the possible neighbourhood. (same theorem holds for closing by replacing Sup by Inf). Thus, if Ψ_1 and Ψ_2 are two algebraic opeings (resp. closings), then $\Psi = \text{Max}(\Psi_1, \Psi_2)$ is also an algebraic opening (resp. $\Psi = \text{Min}(\Psi_1, \Psi_2)$ is an algebraic closing).

In the following, we note γ_λ and φ_λ respectively for algebraic openings and closings ; let λ ... parameter of the neighbourhood considered as ... sing function of the size of this neighbourh...
A morphological filter can begin with an al... opening γ_λ followed by the corresponding φ_λ, t... γ_{λ_2} with $\lambda_2 > \lambda_1$ (resp. closing beginning followed by φ_λ^2). A typical sequence may be written :

$$M_n = m_{\lambda n} \cdots m_{\lambda 2}\ m_{\lambda 1}$$

with $\quad m_\lambda = \varphi_\lambda\ \gamma_\lambda$

and $\quad \lambda_1 < \lambda_2 < \ldots\ldots < \lambda_n$

This kind of operators such as M_n belongs to the general theory of the Alternated Sequential Filters (ASF) defined by SERRA (1986). Filter M_n is increasing and idempotent and :

$$\lambda > \mu \Rightarrow \gamma_\lambda \leq \gamma_\mu \quad \text{and} \quad \varphi_\lambda \geq \varphi_\mu$$

The ASF also satisfy the following law of absorption :

$$n' \geq n \Rightarrow M_{n'} M_n = M_{n'}$$

Thus, the general structure of morphological filters is well defined and the main difficulty is often to construct algebraic openings and closings adapted to the purpose : the choice of structuring element is very important to make interesting properties or the filters.

Thinnings and skeletons

Thinnings and skeletization notions are principally derived from hit or miss transformations.
For a set $X \in \mathcal{P}(I)$, $x \in X$ and a structuring element B which is in two complementary parts (B1 and B2), the hit or miss transformation of x by B (noted $X \circledcirc B$) is :

$$X \circledcirc B = \left\{ x \mid B1_x \subset X \text{ and } B2_x \subset X^c \right\}$$

with $\quad B1 \cap B2 = \varnothing$

In the binary case, such a structuring element shows that detected points x belong to the inside boundary of X. The thinning of set X (noted X o B) with B is obtained by putting in 0 the origin of B when the chosen configuration (B1, B2) is locally matched with the image. Hence, the thinning of X results from the difference :

$$X \circ B = X - (X \oplus B)$$

This operation can be iterated as the others elementary morphological operators. At a given step, if an object does not contain the part B1, the thinning does not act on the considered object and its connectivity it preserved. More generally, thinning is also possible on grey tone images.

The skeleton of a set X is based on the following notion : it consits to describe the shape of a set by a more simple set. In the binary case, skeleton is corresponding to the ideas of bissector and medium axis. Skeleton can be built with iterative thinning precessed untill the idempotence is obtained :

$$X \circ {}_nB = X \circ {}_{(n+1)}B$$

An other construction of a skeleton Sq (X) of set X is realized also by iterative processing :

$$Sq(X) = \bigcup_{n \geq 0,\ \mu \geq 0} (X \ominus {}_nB) \cap \left[((X \ominus {}_nB) \ominus \mu B) \oplus \mu B \right]^c$$

Definition of skeleton of set X is also given by union of the disks centres with maximum radius included in X. Generally, this is modelled by the following distance d :

$$d(x, X^c) \quad \text{with } x \in X$$

Point x belong to skeleton if :

$$d(x, b_1) = d(x, b_2)$$

with $b_1 \in X^c$ and $b_2 \in X^c$ and b_1, b_2 on two opposite sides. Computation of this distance a makes a grey levels function where ridges are equivalent to the skeleton of X. Squeletization is a neither increasing nor decreasing transformation but it is antiextensive and idempotent and preserve connectivity and homotopy.

Skeleton exists also for X^c and it is defined by points $x \in X^c$ at the same distance d(x,X) of different subsets of X. One calls this skeleton a SKIZ : skeleton by influenced zones.

In conclusion, thinning and skeleton allow the extraction of singular lines of an image and that is very important for morphological analysis.

III - LOW LEVEL FEATURE EXTRACTION

Some results have been already obtained in remote sensing by MM at this level of modelling early vision process.
the first problem in image analysis is generally the enhancement of the observed phenomena representation.

More precisely, the perturbations generated by coherent illumination such as speckle in radar imagery are a great difficulty for visual or digital analysis. In this field, alternated sequential filters are interesting because of acting alternatively on local minima and maxima whithout modification of the local mean level.

Different families of ASF (SERRA, 1986) can be used to remove speckle. For example, multidirectionnal filter (SAFA, 1986, 1989) and comparative filter (SAFA, 1989) have been developped in the general formulation :

$$M_n = m_n\ m_{n-1}$$
$$= \gamma_n\ \varphi_n\ \gamma_{n-1}\ \varphi_{n-1}$$

In the directionnal filter, dimension n of the structuring element is determined by the mean speckle grain width. In the comparative filter, dimension n determine radius and shape of pseudo-circular structuring element.

This kind of ASF is very efficient : local mean has a good preservation and standard deviation is very decrased (SAFA, 1989).

Moreover, morphological filters are an indirect way of obtaining textural information : results of low-pass filtering provides a low frequency component of the image, and the high-frequency component contains textural information.

A complementary approach (BOUSQUET, 1986) indicates also the interest of enhancement of flat grey levels segments determined after conventional textural analysis by using morphological filters.

For many applications, edge detection methods are very often necessary to detect local contrasts and delineate boundaries betwen ground objects. The introduction of MM for this aim is based on computation of the residues of erosion and dilation operators. Morphological gradient can also contribute to edge detection by combining theese two elementary operators. Generally, the results quality is related to the existing local perturbations. To avoid this difficulty, two ways are actually used.

The first one consist of coroine blur with grey tone erosion and dilation (LEE, 1986). It seems that this blur minimum morphologic edge operator it a less expansive method to extract edges in many images. The second one consists in modelling early vision by morphological filtering in two channels size-tuned with two spatial frequencies levels. Theese two conditions define dimensions of structuring elements used to thin result of first derivative operators (MADIER, 1986). Now, it is interesting to see the interest of algorithm which use skeleton for the zero-crossings location in comparison with the others methods.

Extraction of singular lines is an other important chapter in image analysis. For exemple, ridges, valleys, roads and by hydrologic networks are important information but the elongated shapes imply more complex morphological model for a good feature extraction. This is possible by appropriate sequences applied to well adapted spatial resolution images (DESTIVAL, 1986, 1987a, 1987b). Stimulating attempts are also realized by 3D morphological approach on digital terrain models to obtain physicaly significant ridges, peaks, valleys and watersheds.

Spectral characteristics are a priori out of the domain where MM can be used. Nevertheless, if we cansider a two spectral channels image, positions of points belonging to the different classes of training fields constitute an image. This 2D space is open to classification technics derived from imagery processing. The principle of such a methodology is to group points in objects and to separate the corresponding objects by SKIZ or distances computation (MADIER 1986a, 1986b, 1987). Different methods have been tried and a particular attention is important to solve overlapping problems between different classes.

IV CHARACTERIZATION AT CONNECTED COMPONENT LEVEL

In this part of image analysis digital processing contribute to enhance results of low level functions by simplifying the scene into a continuous sketch defined by connected entities or segments.
The place of MM is fundamental to do reconstruction and segmentation from classified images : this case reduced to a set of binary images formed by a sequence where one class is considered with its complement (the union of all the other classes). An important task is to cancel out artificial texture of classified images and many algorthims are possible (DURAND, 1985, FLOUZAT, 1986).
Enhancement by agglomeration coupled with size filtering is now well known (FLOUZAT, 1983, 1984) and very useful to extract quantitative information : principle is based on closing followed by increasing openings ; the results show significant visual object and give measurements about size distribution of elementary parts of the considered class like after a sieving.
At this level, geometric analysis is therefore a contribution to image processing and MM allows to extract shape primitives to do shape recognition (SHAPIRO, 1986). Quantitative characterization of natural forms by elementary operators sequences are also directly possible (MERING, 1987a, 1987b).

More generally, MM conducts to many objects filtering based on the different labels recognized on the objects. Directional filtering belongs to this processing family. When a class of things extracted from an image is well contrasted, one acts on size and orientation of linear structuring elements. The basic operation is a sequential opening where a parameter of the structuring element is modified at each interation. Thus, if such a geometric probe is included (or not) in an object, this one has the corresponding label (or not). This method is very efficient for the principal direction analysis on a digital image (MARIEL, 1989).
Filtering on the length of a class of entities is also possible with MM sequenced operators. that is a traditional case of the use of skeleton. The interesting class of objects is reduced to a binary image and skeleton is computed. A logical transform is generated along an iterative process and recognize the terminations of skeleton branches. For one iteration, skeleton is less longer than two pixels ; therefore at the iteration λ , skeleton have lost 2λ pixels and objects with a skeleton more little than 2λ are cancelled. Conditional dilation ,using such a skeleton as a mark permits to build original objects of a length class (FLOUZAT, 1986).

Likewise, an analysis of skeleton can furnish a directional information by considering its points as a graph and computing a linear regression on it.

V HIGH LEVEL STRUCTURE QUANTIFICATION

The contribution of MM in image understanding concerns principally quantitative global description. In this step of image analysis, interesting information generally comes from structure characteristics. Thus, the essential problem to solve is to quantify an image by scene analysis corresponding to visual interpretation.
The first step of this methodological part is based on the generalization of local measurements adapted to every connected component. For example, directional filtering of labeled entities gives easily statistics about angular distribution for the whoole population of objects. Then, all the possible computations on such an histogram provide global measurements on the scene. This method can be generalized to the different labels of connected components built at the previous step. If digital processing at this level is associated with data base, global handling of entities is more easy. In the same way, intercept measurements applied to the objects level indicate global spatial arrangement in the image. Thus, statistics on intercept frequency versus polar angle or on length frequency provide information about spatial position and orientation of objetcs. (FLOUZAT, 1986)

An other kind of global measurements can be derived directly from morphological operators. A first example is the morphological compacity index (FLOUZAT, 1984) which is the ratio betwen the opening and closing of the same set. This index is interesting because it gives two global informations about shapes and edges physionomy. It have also an intersting behaviour when it is used iteratively with an increasing structuring element. Numerous others measurements have been developped to have global observations but also in order to make a monotoring of classified image enhancement (DURAND, 1985).

The second step of structure quantification is based on symbolic graph computation.
The aim is to obtain a global or a limited area graph representing spatial arrangement by a measurement of adjacency lengthes. This can be computed by a very simple morphological processing ; let L (A,N) be the

contact length betwen sets A and N :

$$L(A,N) = \text{Card}\left[(A \oplus B) \cap N\right]$$

The meaning is that L (A,N) is obtained by points dilated from A and restricted to N. By derivation, L (A, I - A) is the adjacency length between set A and all of the others sets in the image :

$$L(A, \{ I - A \}) = \text{Card}\left[(A \oplus B) \quad A^c\right]$$

Computation of L (A, {I - A}) is also an information about the outside perimeter of set A. The building of symbolic graph of an image is obtained as follows : all sets are the nodes and measurements of L provides the lengthes of corresponding ares.

Frequently, the different sets are the objects of differents classes represented in an image. Generally, it is very interesting to see a graph for an image and an other graph corresponding only to a part of this image. Stimulating inferances are also possible by comparison betwen such a symbolic graph with the corresponding graph provided by geographic information system on the same area.

CONCLUSION

This synthesis is only an indication about that it is possible with MM in remote sensing imagery processing.
Two points are importants :

- morphological approach is a complementary method in comparison with the others passibilities in digital processing.

- morphological analysis can be split up into elementary operators which have possibilities to be implamented in very high speed of execution time.

Theese two facts lead to see that large developments are conceivalble.

BIBLIOGRAPHY

BARLOW, H.B., NARASIMHAN, R., ROSENFELD, A., 1972 Visual Pattern Analysis in Machine and Animals. Science. Vol. 177, N° 4049, pp 567-575, 1972.

BOUSQUET, P., FLOUZAT, G., 1986. Segmentation by Mathematical Morphology Hand on the Image understanding Model in Remote Sensing. IGARSS'86 - ESA. SP, Vol. 1, pp 553-557, Sept. 1986, Zurich.

COLWELL, R.N., 1965. Aids for the Selection and Training of Photointerpreters. Photogrammetric Engineering, Vol. 31, N° 2, pp 327-339, march 1965, Washington.

DESTIVAL, I., LE MEN, H., 1986. Detection of Linear Network on Satellites Images. Eighth ICPR, Vol. 2 pp 856-859, October 1986-Paris.

DESTIVAL, I., 1987a. Recherche automatique des reseaux lineaires sur les images SPOT - Bull. Soc. Franç. Photogr. Teted., N° 105, january 1987, Paris.

DESTIVAL, I., 1987b. De l'extraction des reseaux lineaires à leur suivi sur des images SPOT par un système d'interprétation à base de connaissances. Coll. SPOT 1, pp 1395-1402, CNES, november 1987, Paris.

DURAND, M.A., FLOUZAT, G., 1985. Quantification de l'aspect visuel des images classées. Photointerprétation, Vol. 3, N° 5, pp 47-51, May 1985, Paris.

FLOUZAT, G., 1982. Modelisation de la compréhension visuelle des images de télédétection : essai de simulation numérique de la photointerprétation analytique. ISPRS Symp., Vol. 2, pp 7-24, september 1982, Toulouse.

FLOUZAT, G., MERGHOUB, Y. 1983. Modelisation de l'extraction manuelle d'éléments texturaux par la morphologie mathématique. Photointerprétation Vol. 6, N° 6, pp 59-64, November 1983, Paris.

FLOUZAT, G., SERRA, J., MERGHOUB, Y., 1984. Aide automatique à la compréhension des images de télédétection : modelisation des caractères morphologiques de la texture. First Coll. on Flectronic Image, CESTA, Vol. 2, pp 735-740, may 1984, Biarritz.

FLOUZAT, G., MOUEDDENE, K., 1986. Computer aided interpretation of complex geological patterus in Remote Sensing. IGARSS 86, ESA-SP 254, Vol. 2, pp 783-786, september 1986, Zurich.

FLOUZAT, G., 1988. Morphologie mathematique et analyse des images de teledetection. Photointerpretation, Vol.5, N°1,pp 1-16, October 1988, Paris.

FLOUZAT, G., DURAND,. M.A., 1988b. Reconstruction morphologique de structure sur des images classées. Photointerprétation, Vol. 5, N° 4, pp 33-38, October 1988, Paris.

GUY, M., 1970. Traitement des images et methodes d'exploitation Geoforum, 3, pp 47-62, 1970.

LEE, J.S.J., HARALICK, R.M., SHAPIRO, L., 1986. Morphologic Edge Detection. Eighth ICPR, Vol. 1, pp 369-373, October 1986, Paris.

MADIER, M., FLOUZAT, G., BOUSQUET, P., 1986. Extraction des structures floues par un modèle morphologique simulant les propriétés psychophysiques de la voie visuelle primaire. Snd Coll. on Electronic Image. Vol. 2, pp 749-756, april 1986, Nice.

MADIER, J.P., FLOUZAT, G., JOURLIN, M., 1986a. A Non-Parame tric Supervised Multispectral classification Method Using Binary Morphological Operators. IGARSS' 86. ESA-SP-254, Vol. 1, pp 547-552, september 1986, Zurich.

MADIER, J.P., FLOUZAT, G., BOUSQUET, P., JOURLIN, M., 1986b, Partition Morphologique avec apprentissage dans un espace bidimensionnel. Snd Coll. on Electronic Image, Vol.,1, pp 90-96, april 1986, Nice.

MADIER, J.P., FLOUZAT, G., JOURLIN, M., 1987. Deux methodes de partition morphologique anisotrope. Third Coll. on Electronic Image, MARI87, CESTA, pp 334-341, may 1987, Paris.

MARR, D., 1980. Vision. A computational Investigation into the Human Representation and Processing of Visual Information. Freeman and Company 1980, New--York.

MARTEL, C., FLOUZAT, G., SOURIAU., A., SAFA., F., 1989. A Morphological Method of Geometric Analysis of Images : Application to the Gravity Anomalies in the Indian Ocean. J. Geophy. Res. Solid Eavth and Planets, Vol. 94, N° B2, pp 1715-1726, february 1989.

MATHERON, G., 1975. Random sets and Integral Geometry. Wiley, New-York

MERING, C., PONCET, Y., JACQUEMINET, C., RAKOTO, R., 1987a. Quantitative Description of Denudation Forms in the Western African Sahel. Adv. Space Research, Vol. 7, N° 3 , pp 31-39, Pergamon Press, London.

MERING, G;, JACQUEMINET, C., 1987b. An approach of quantitative Description of Sand-Hills Shapes in the West African Sahel from Remote Sensing Imagery. Acta Stereologica, pp 951-956, Ljubljana.

SAFA, F., FLOUZAT, G., 1986. Enhancement of Radar Images for Interpretation by simulating the Vision of Edges and Textures. IGARSS'86. ESA-SP , Vol. 3, pp 1629-1633, Sept. 1986, Zurich?

SAFA, F. , FLOUZAT, G., 1989. Speckle removal on radar imagery based on Mathematical Morphology. Signal Processing, Vol. 16, N° 4, april 1989, North--Holland.

SERRA, J., 1982. Image Analysis and Mathematical Morphology. Academic Press, New-York.

SERRA, J., 1986. Elements de théorie pour l'optique morphologique. Thesis Doctorat d'Etat. University of Paris VI, january 1986, Paris.

SHAPIRO, L.G., MacDONALD, R.S., STERNBERG, S.R., 1986. Shape Recognition with Mathematical Morphology. Eighth ICPR, Vol. 1, pp 416-418, October 1986, Paris.

WANG, S., ELLIOTT D.B., CAMPBELL, J.B., ERICH, R.W., HARALICK, R.M., 1983. Spatial Reasoning in Remotely Sensed Data. IEEE Geo. Rem. Sens. , Vol. GE-21, N°1, pp 94-101, january 1983.

Mathematical Morphology and Remote Sensing

Robert M. Haralick
Intelligent Systems Laboratory
Department of Electrical Engineering • FT-10
University of Washington
Seattle, WA 98195

ABSTRACT

An algebraic system of operators, such as those of mathematical morphology, is useful for remote sensing image interpretation because compositions of its operators can be formed which, when acting on complex shapes, are able to decompose them into their meaningful parts and separate the meaningful parts from their extraneous parts. Such a system of operators and their compositions permit the underlying shapes to be identified and reconstructed as best possible from their distorted noisy forms. As well they permit each shape to be understood in terms of a decomposition, each entity of the decomposition being some suitably simple shape.

Since shape is a prime carrier of the spectral information held in remote sensing imagery, there should be little surprise about the importance of mathematical morphology. Morphological operations can simplify image data preserving their essential shape characteristics and eliminate irrelevancies. As the identification and decomposition of objects and surfaces, correlate directly with shape, it is only natural that mathematical morphology has an essential structural role to play in remote sensing image interpretation by computers.

REAL-TIME MORPHOLOGICAL TRANSFORMATION FOR LINEAR FEATURE EXTRACTION ON AERIAL IMAGES

DAN IONESCU, AND ROBERT KADAMANI

Department of Electrical Engineering, University of Ottawa, 770 King Edward, Ottawa, Ontario, Canada, K1N 6N5

In image analysis problems that have to face time constraints, a convolution kernel is needed. The size of the kernel has to be as small as possible for meeting the time restriction and image reconstruction fidelity.

Arithmetic operations are computational intensive and time demanding. Morphological operators provide an alternative for fast and accurate methods to treat images locally. Combinations of binary operators can be made in order to have some global features of the image. The convolution process is achieved without using addition and multiplication.

Operations like linear feature extractions and structure location are executed by convolution with a specific structuring element. Some of the features of aerial images such as roads, airport runways, etc, are suitable for being described as linear features and consequently obtained by applying morphological transformations.

Experiments executed in real time on aerial images on an image processing computer based on a linear mesh architecture are presented

The results are highly accurate and robust.

INDEX TERM : Binary operations on aerial images, pictorial algebra, 2-D, 3-D convolution, linear feature extraction.

OPERATIONAL SATELLITE-DERIVED SST CHARTS FOR CANADIAN WATERS

Brian W Wannamaker
SEA SCAN
16065 Fifth Line, Albion
RR3 Caledon East,
Ont. Canada, L0N 1E0
(416) 880-0528

Data from the TIROS-N/NOAA A-M series of near-polar orbiting meteorological satellites is received and distributed operationally in Canada by the Atmospheric Environment Service at Downsview (Toronto) and Edmonton. Recent upgrades to these stations raise the possibility of improvements in the use of quantitative data from these satellites in operational and research applications. One operational product developed to prototype stage is an automatically produced Sea Surface Temperature Chart. This information has immediate applications to fisheries, navigation and military operations. The technique of obtaining quantitative, georeferenced polar satellite data without human intervention has much wider applications. The SST chart is created in a digital or imagery format in standard map projections. Each pixel value represents a best estimate of the sea surface temperature in a grid square over a period of time. A second chart illustrates a confidence level for the estimate.To produce a quantitative product, the data must be calibrated both in geophysical and geographical units. Instrument calibration includes the full blackbody temperature-dependent non-linearity transmission. The conversion from digital count value to equivalent black body temperature at the top of the atmosphere includes the wavelength dependence of each channel's sensor response and of the surface emissivity.In the spatial calibration, Three models of spacecraft position versus time are used. A simple circular model is used to extract a buffer of data that covers the area of interest. The classical Brouwer-Lyddane model is used to update the ephemeris parameters from an outside source to a new epoch time within the area covered. These parameters are used to define an osculating elliptical orbit for the actual bending of the image into a standard map projection. Data is accumulated in grid squares over a number of individual passes of two satellites. Contamination of surface temperature estimates by cloud and ice is avoided using both visible and infrared channels. Atmospheric effects are removed through the use of a dual infrared channel combination. For waters off the eastern coast, combination coefficients determined from North Atlantic atmospheric profiles are used. Comparisons of the results with SST charts created from ship and buoy data will be given. The trade-offs between spatial and temporal resolution and percentage cloud cover will be discussed. Precise charts of surface temperature can be produced in an operational setting.

INTEGRATING WATER QUALITY SURVEILLANCE WITH AUTOMATED BIOMONITORING FROM REMOTE WATER QUALITY PLATFORMS

E. L. Morgan, W. L. Pennington, and M. A. Eggleton
Center for the Management, Utilization, and Protection
of Water Resources and Department of Biology,
Tennessee Technological University,
Cookeville, TN 38505
615-372-3134

ABSTRACT: Advances in site-specific monitoring of fish's physiological responses to continuous stream flows now complement surveillance programs utilizing remote water quality data collection platforms. Developments in automated biosensing provide methods for measuring in situ fish breathing rate changes to stream hydrographs and add a real-time biological monitoring dimension to remote water quality networks incorporating satellite data retrieval communications.

In meeting our goal of developing automated biosensing capabilities for remote monitoring, a series of field trials was designed to test various configurations of in situ fish-holding chambers, breathing rate detectors, and system interface to streamside water quality data collection platforms for satellite data retrieval. Results were used to design groups of automated biosensing devices for detecting in situ rainbow trout breathing rate responses and to implement six of these units at each of two data collection, platform-equipped, water quality stations located along a stream subject to acid precipitation influences in the Southern Appalachian Mountains. Remote stations are being maintained for near real-time data needs as a part of the 5-year Acid Precipitation Mitigation Program initiated in 1985 by the U. S. Fish and Wildlife Service under the National Acid Precipitation Assessment Program. Our objective is to evaluate stream ecological responses to flow regulated liming for a two-year treatment period.

AIRPHOTO INTERPRETATION AND SALMON RIVER RESTAURATION

Jean-Marie M. Dubois and André Gosselin

Centre d'applications et de recherches en télédétection (CARTEL)
et Département de géographie et télédétection
Université de Sherbrooke, Sherbrooke, Québec, Canada J1K 2R1
Téléphone: 819-821-7180 Télécopieur: 819-821-7238 Télex: 05-836149

Résumé

La gestion des stocks de saumon atlantique est devenu un problème important et, devant le besoin d'agir, la protection et la restauration des habitats en rivière est primordiale. La méthode proposée par télé-interprétation vise l'inventaire physique des habitats et la quantification du potentiel salmonicole des rivières dans le but de réintroduire le saumon dans ses anciens habitats. Dans cette méthode, on utilise des photographies aériennes peu coûteuses au 1: 15 000 avec un minimum de relevés au sol.. On divise la rivière en biefs dans lesquels on évalue les critères de pente, de géomorphologie, de morphométrie et de granulométrie. On s'attarde de façon plus particulière aux fosses, frayères et obstacles à la montée du saumon. Par la suite, les différents habitats sont pondérés et on leur attribue un facteur de productivité. Cette méthode est fiable à plus de 80 à 95% selon le type de vérification au sol. Une application à une rivière du sud du Québec où le saumon a disparu depuis 75 ans à cause de la pollution et de la présence de barrages, a permis d'estimer la productivité d'un seul sous-bassin de 1 030 km2 à 34 900 saumonaux, en plus de localiser avec précision les sites de bons potentiels et d'identifier les secteurs où les activités de restauration seront prioritairement dirigées. (Grants: NSERC, The Royal Canadian Geographical Society)

Key words

Airphoto, aquatic habitat, sediments, geomorphology

1. Introduction

Over the last few years, the protection and restoration of atlantic salmon rivers has gained increasing interest. This concern stems from the need to protect the atlantic salmon from environmental pressures which increasingly threaten the survival of the species. In spite of its incredible physical strength, the atlantic salmon requires specific and stable environmental conditions in order to survive. Accordingly, efforts to reintroduce the atlantic salmon in a specific region from which it has disappeared or to introduce it in a new region require gathering information on the state of the habitats available for salmon.

A method of inventory was developed in Québec during the course of geomorphological research work carried out on the north shore of St.Lawrence River (Dubois, Clavet et Desmarais, 1975). Since 1976, this method has been applied to both small and large rivers in different physiographic regions of Québec. Gaspésie, James Bay, the Appalachian plateau and the north shore of St.Lawrence River (Dubois et Clavet, 1977). This type of hydromorphological inventory using remote interpretation has turned out to be a basic tool in the evaluation of the potential for fish production in rivers. It permits to derive a semi-quantitative evaluation of many habitat parameters and it can also be used for the basic management of fish inventories by providing a stratified sampling giving a much more reliable estimation of parr populations (Caron et Ouellet, 1985). However, the derived map proposed by Dubois and Clavet (1977) does not permit the quantitative evaluation of smolt production capabilities nor to establish the number of spawners needed to maintain the optimal production of a river (Côté et al. 1987).

To address these needs, a method for quantifying the potential of rivers for salmon production inspired by the work of Dubois and Clavet (1979) was developed by the "Ministère des loisirs, de la chasse et de la pêche du Québec (MLCP)". This method proposed by Boudreault (1984) offers a more detailed description of rivers as it provides for a simplified map compared to the one proposed by Dubois and Clavet (1977) and for the tabular compilation of data. It is however limited to characterization of the river bed described essentially according to its channel pattern and substrate.

However, a more thorough inventory is necessary in regions where anthropogenic influences on the environment are more significant. In few basins where industrial and agricultural development are more important causing specific concerns, it is essential to adapt and improve the present techniques of inventory. Intensive development has a direct influence on the aquatic environment of the salmon, which reacts strongly to any modification of his habitat. In this context, in addition to the estimation of salmon production, it is essential to acquire a description and to locate areas presenting risks for the reintroduction of the salmon and those offering the highest potential for good salmon habitats. The development of a method of inventory applicable to this type of environment is of primary importance and constitutes the first step in the reintroduction of the atlantic salmon in such regions. The main objective of the proposed methodology is to evaluate several factors (obtacles to fish migration , channel patterns, bed grain-size, shoreline land use...) which are fundamental for the thorough assessment of the salmon habitat.

2. Material and methods

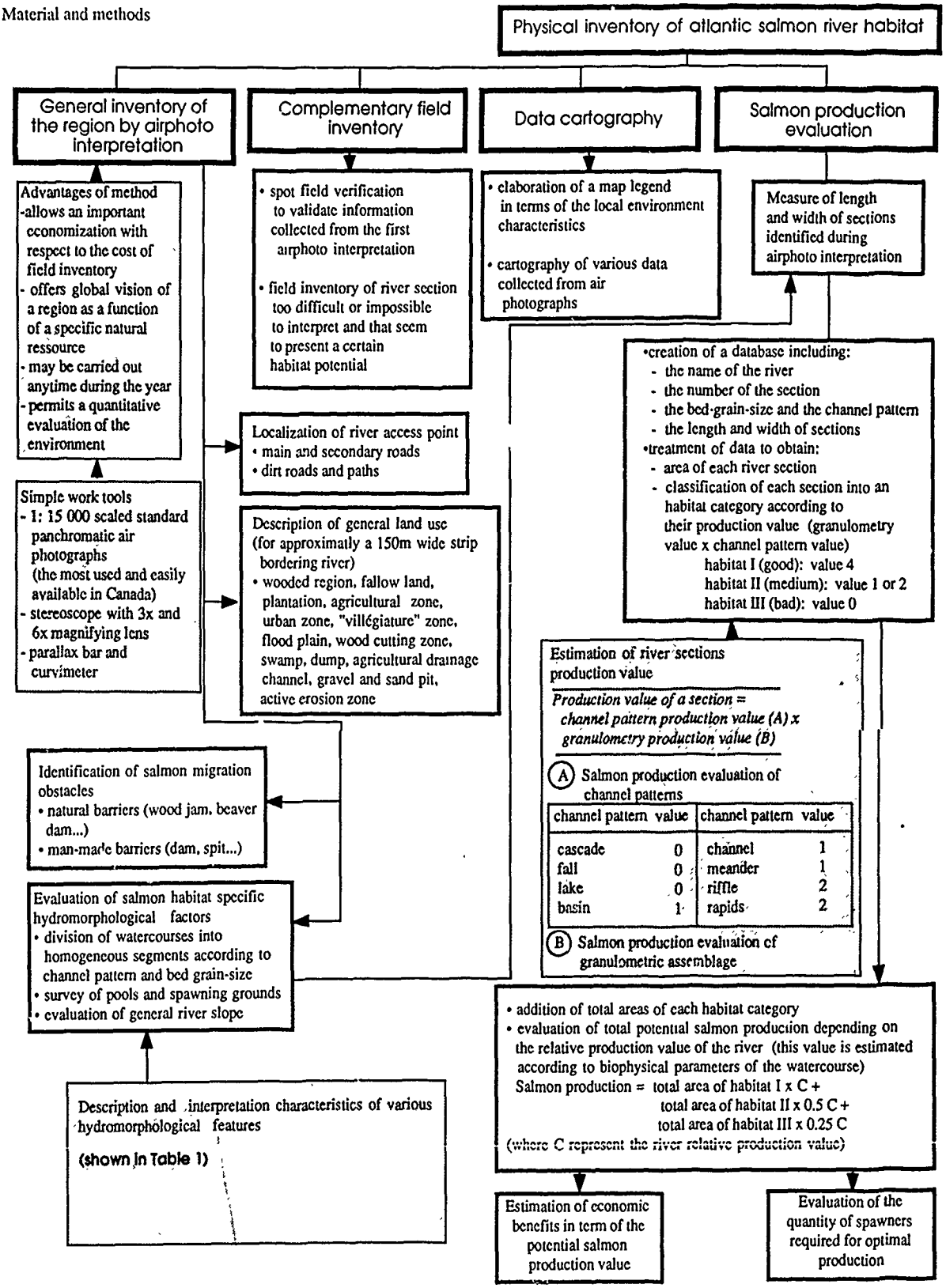


Figure 1 - Diagram of the inventory methodology

Table 1 - Description and interpretation characteristics of hydromorphological features

feature	description	interpretation characteristics
basin	• zone of deep water generally located at the base of an obstacle and corresponding to an enlargement of the channel • sedimentation due to slow currents	• easily recognizable due to its position at the base of obstacles, the considerable channel widening and by the impression of depth associated with the dark water colour • this feature generally corresponds to the presence of a main pool
cascade	• nickpoints of river bed resembling stair steps where bedrock and blocks dominate bed grain-size • obstacle which generally can be negotiated by salmon except under certain conditions	• alternation of white and grey tones that shows the succession of nickpoints often interposed with rocky basin
channel	• portion where water depth, generally 1m or more, is constant • the current is moderate to slow and the water surface is smooth • bed material varies from sand to large pebbles	• identifiable by the dark colour produce by the generally high water depth and by the smooth water surface resulting from the slow current
fall	• segment where a sudden drop occurs • composed of bedrock and sometime very large blocks • may or may not be bypassed by salmon	• sudden dislevelment of river bed • white tinted water surface
lake	• spot where no lothic characteristic are observed; the current is slow to absent and the water depth high	• lakes, even the smallest, are easily identifiable when they interupt the course of rivers and where a change of environment, from lothic to lenthic, is noticeable
meander	• section where river course describes many loops • river bed presents alternation of riffles and pools • bed grain-size varies from sand to small pebbles	• easily noticeable by its looping structure and alternation of riffles and pools
rapids	• characterized by a small change in the slope of the river bed where the current is fast and the water surface broken by a large outcrop of bedrock	• recognizable coarse texture produced due to water turbulence on coarse substrates • white and grey spots on the water surface are also a clue in identifying that feature
riffle	• sector of low depth constituting a shoal, a sucession of shoals with pools, or a small increase in the river bed slope • the current is rather fast • the bed grain-size generally varies from gravel to small pebbles with occasionally a sandy fraction • segment where best spawning areas are found • spawning areas are divided into two categories: 1-main spawning area: major bed grain-size varies from gravel to small pebbles and where the river bed shows depth variations along the whole section 2-secondary spawning area: more uniform river bed configuration and the water depth is more constant, also the grain-size may contain a sandy or a coarse material fraction	• noticeable by light tones that often indicate the presence of shoals • relatively smooth texture produced by the many undulations affecting water surface • however those ripples sometimes create a glidder effect almost identical to the case of light reflected from wavelets produced by the wind (this problem can be avoided by observing adjacent airphotos and by considering the flow up and downstream of the segment studied)

3. Results

The development of the method of inventory is the result of a study undertaken on the "au Saumon" River basin, tributary of the "Saint-François" River. The "Société Saumon Saint-François" group is actively working to reintroduce the atlantic salmon in the region. The results derived from this study have led to the overall improvement and effectiveness of the method of inventory.

The "Saint-François" River basin is located in the southern part of Québec. The basin represents an area of approximately 10 230 km2. The present project deals with a typical sub-basin, the "au Saumon" River basin.

The longitudinal profile of the streams in the "au Saumon" River basin is irregular, mostly smooth. With a length of approximately 75 km and an average slope of 0,58%, the "au Saumon" River is quite representative of good salmon rivers. A longitudinal profile constitutes a determining factor in the evaluation of the salmon potential. The longitudinal profile of a typical salmon river is, according to Elson (1975), between 0,2 and 1,15 %.

3.1. Salmon habitat inventory

The complete inventory of the "au Saumon" River basin has permitted to delineate precise areas of the territory where priority interventions should be concentrated and to define sites with high potential habitats

for the reproduction and the development of salmon. Figures 2 shows example of the mapping of data collected during the survey of the "au Saumon" River basin. Furthermore, it was possible to evaluate the entire basin and to locate potential spawning sites and various areas where river bank reforestation should be given priority . This type of inventory is useful to people in the field working for the reintroduction of salmon in the basin who need to direct their actions more efficiently in the context of managing and producing of salmon.

3 2 Potential for salmon production of the "au Saumon" River basin

Data collected during airphoto interpretation have been processed in order to evaluate the potential for salmon production of each stream of the "au Saumon" River basin (Table 2).

Table 2 - "Au Saumon" River basin potential for salmon production (in smolts)

River or stream	number of segments	salmon production (smolts)
"au Saumon" River	97	18 389
Ditton River	38	2 123
Chesham River	29	1 042
"la Loutre" Stream	7	1 143
Mining Stream	7	171
Bown Stream	14	1 030
Moffat Stream	11	9 911
Albion River	7	1 091
salmon production potential:		34 900 smolts

3.3. Discussion

The results of the inventory giving 34 900 smolts as the potential for salmon production of the "au Saumon" River watershed (1 030 km2) are interesting in that they represent a relatively high potential compared to other basins in Québec. The Jacques-Cartier River near Québec City, with an area of 2 515 km2, has recently been restored for the reintroduction of the salmon and has a potential for salmon production estimated at 58 000 smolts (GTRJC, 1979).

There is a certain margin of error to be reckoned with, given the difficulty of interpreting several physical factors describing the salmon habitat. However, comparison between field data and the results obtained by aerial photograph interpretation indicate that the degree of confidence attributable to the method of inventory is over 80%.

Furthermore, areas measured using data relating to the length and width of each segment of the river introduce a certain margin of error leading to area over-estimation when these are estimated at any other time than at low water level. However this over-estimation is partially compensated by the under-estimation resulting from the use of the average scale of the aerial photographs during the computation of segment areas. Moreover, when faced with uncertainty during the airphoto interpretation of the hydromorphological factors, the choice was always directed towards the parameters resulting in under-estimation in order to obtain more reliable final results.

4. Conclusion

The methodology developed during this study, based on airphoto interpretation, permits the inventory of the physical factors giving a description of the habitat of the atlantic salmon for a given river. Furthermore, the method facilitates the planning and management of the salmon resource and its environment. The method of inventory cannot describe biophysical factors (such as interactions between competing species or physico-chemical parameters), but it does permit the location of sites offering potential for salmon production where further study is warranted. The method of inventory represents a first step in the framework of a more elaborate biophysical survey of a watershed. Finally the method allows one to better identify and characterize sites where environmental management initiatives should be carried out in order to preserve and protect salmon habitats.

5. References

BOUDREAULT, A. 1984. Méthodologie utilisée pour la télé-interprétation des rivières à saumon de la Côte-Nord. Ministère du Loisir, de la chasse et de la pêche du Québec, 26 p.

CARON, F. et Ouellet, G. 1985. Méthodologie d'inventaire des saumons juvéniles au Québec. Ministère du loisir, de la chasse et de la pêche du Québec, 13p.

COTÉ, Y., Dulude, P., Jomphre, D., le Bel, J.-P., Ouellet, G., Rouleau, A. et Roy, L. 1987. Essai de classification normalisée des substrats granulaires et des faciès d'écoulement pour l'évaluation de la production salmonicole. Ministère du loisir, de la chasse et de la pêche du Québec, Direction régionale des opérations régionales, Direction générale de la faune, 4 p.

DUBOIS, J.M.M., Clavet, D. et Desmarais, G. 1975. Rapport préliminaire sur la cartographie des facteurs hydromorphologiques influant sur l'activité du saumon de la rivière Matamek et de ses tributaires. Matamek Annual Report 1975, Woods Hole Oceanographic Institution, WH01-75-62, p. 77-87.

DUBOIS, J.M.M. et Clavet, D. 1977. Hydromorphologie du bassin de la rivière Matamec et de la rivière Moisie et recommandations pour la cartographie des rivières à saumon du Québec: rapport final. Département de Géographie, Université de Sherbrooke, 43p.

DUBOIS, J.M.M. et Clavet, D. 1979. Télédétection de l'hydromorphologie et du potentiel de frayères à saumon de la rivière Saint-Jean en Gaspésie. rapport final. Dép. de géographie, Univ. de Sherbrooke Québec, 23p.

ELSON, P. F. 1975. Atlantic salmon rivers, smolt production and optimal spawning, an overview of natural production. Salmon Restoration Conference. International Atlantic Salmon Foundation Special Publication, no.6, p. 96-119.

GTRJC (le groupe de travail de la rivière Jacques-Cartier). 1979. Avant projet de restauration du saumon dans la rivière Jacques-Cartier. Ministère du tourisme, de la chasse et de la pêche du Québec, 64p.

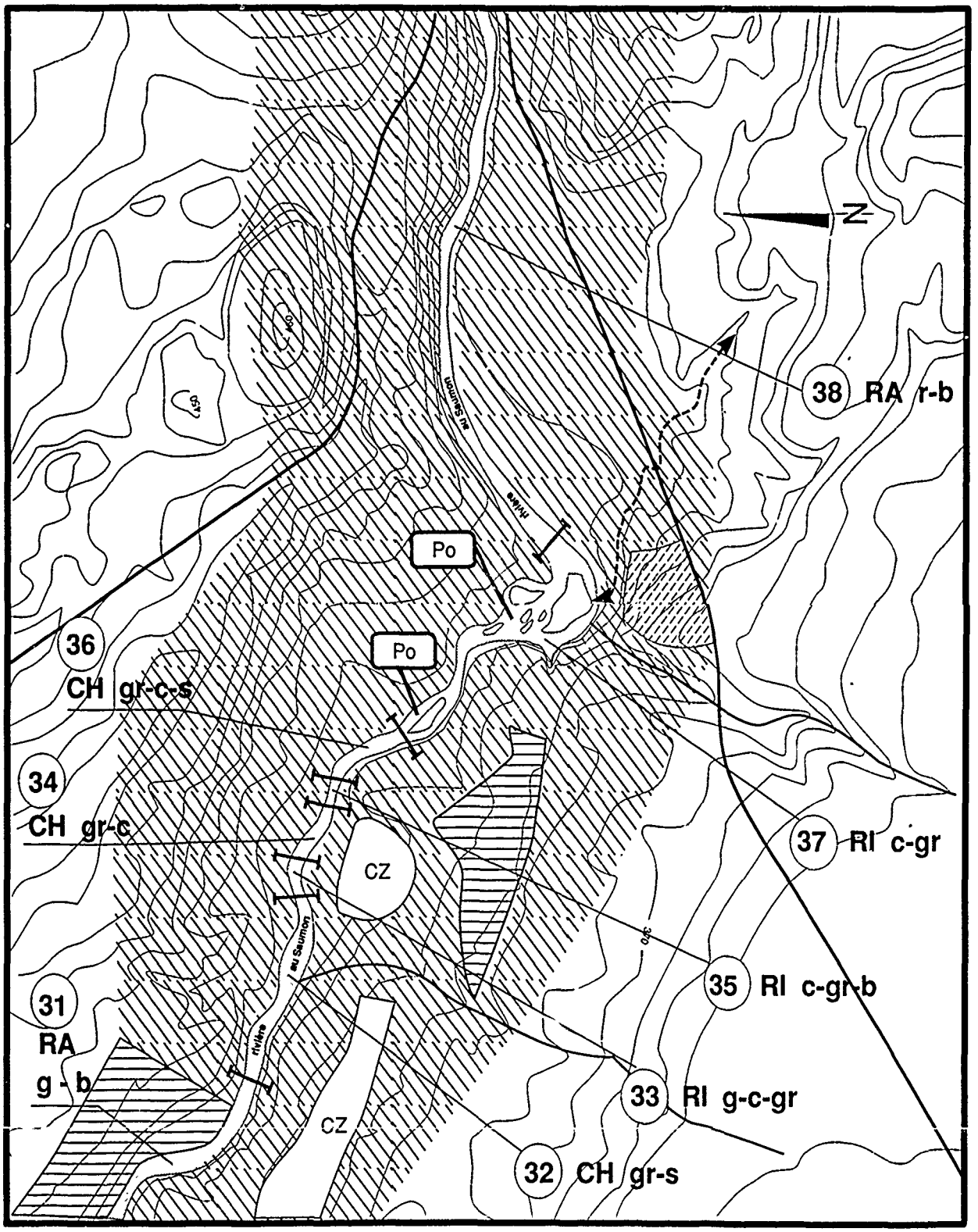


Figure 2 - Map of various data collected on the "au Saumon" River

THE INFORMATON CONTENT OF AVHRR, MSS, TM AND SPOT DATA IN THE SKAGERRAK SEA.

KAI SØRENSEN[1], TOMMY LINDELL[2] AND JAKOB NISELL[2].

[1]Norwegian Institute for Water Research,
Blindern, N-0313 Oslo, Norway.
[2]Centre for Image Analysis, University of Uppsala,
Glunten, S-75183 Uppsala, Sweden.

ABSTRACT

A investigation has been performed in the Skagerrak sea in May 1988 for evaluation of the possibilities of tracing different water masses containing algae and suspended sediment. The study has been performed during a bloom of the algae Chrysocromulina Polylepis and during heavy fresh water input from the river Glomma into the archipelago in the northern Skagerrak.

A geometric linking of AVHRR, MSS, TM and SPOT data was performed on May 13 in order analyse the information content of the different sensors. Data were also analysed to define the applicability of various sensors with different geometric scales. The AVHRR thermal data where used for identification of the water masses where the algae resided while the MSS, TM and SPOT data were used to define frontal zones in detail and to monitor the distribution of suspended sediments within the archipelagos. The image evaluations were based on within satellite calibration references and field observations. The advantages of the combination of high and low resolution data is also investigated.

It was observed that in the Skagerrak MSS and TM were both sufficiently good for identifying the different water masses. For describing distribution of suspended sediments within the archipelago the MSS, TM and HRV were similar. For the open sea the AVHRR thermal data gave sufficient thermal information, but as soon as one approaches the coastline or within the archipelagos the TM becomes superior. The reflective bands of AVHRR gave little information of the algae, partly bacause of sun glint, but a rather detailed knowledge of the Glomma water distribution. Differrent ratioing techniques showed the best discriminating possibilities and were used for the classification of the different water masses.

Keywords:Skagerrak, Sensor linking, NOAA, Landsat, SPOT

INTRODUCTION

The present investigation have been executed within an ongoing project for calibration and evaluation of different types of multispectral satellite data for operational use in water quality monitoring.

A special study has been performed within a situation with heavy freshwater input and a algal bloom of Chrysocromulina polylepis in the Skagerrak Sea which caused severe ecological consequences. During the last decades algal blooms occurres more frequently in this area and they often follow the Norwegian Coastal Current. This was also the case during May 1988 and a close correlation between the advancing watermass defined by the temperature and the massive algal bloom was observed. Thus it was possible to use thermal IR images to monitor the movements of the algae.(Johannessen et al., 1988). On May 13 the north-east winds in the northern Skagerrak, the large amount of freshwater and a weak outflow from the Oslofjord diminished the algal invasion in the surface waters of this area.

Monitoring of temperature, algal bloom and suspended sediments is probably the most important task within Aquatic Remote Sensing and it is therefore important to increase the knowledge of the limitations, usefulness and information content of different sensors.

MATERIAL AND METHODS

The linking of the different data sets have been performed on scenes collected on May 13. The data used in the study are NOAA AVHRR from morning (0854 GMT) and afternoon (1336 GMT), the Landsat scenes from 0958 (GMT) and SPOT data from 1046 (GMT). AVHRR, MSS and TM covers the Oslofjord and Skagerrak and the SPOT scenes only the archipelagos in the northern part of Skagerrak (Oslofjord). The AVHRR, MSS and SPOT data have been geometrically rectified to TM with a pixel size of 240 m, but the full information content of the different sensors have been used in the evaluation. The image processing has been made on a PC-based EBBA-GIS system, produced by the Swedish Space Cooperation.

Ground truth data have been collected in the Hvaler archipelago, but unfortunately no in situ data were available from the heavy algal area on May 13. The presence of the algae is confirmed by reports of toxic and lethal effects on fish and by data from measurements along the Norwegian coast collected a few days earlier.

Figure 1 shows the investigation area with the different test sites where the information content of the data have been analysed. On site A and B thermal infrared information has been used. The sites C and D have been used for multispectral identification of algae and suspended sediments.

RESULTS

The temperature distribution in the Skagerrak on the May 13, at 0854 GMT is shown in Figure 1. In the central Skagerrak a jet of warm water (>10 deg.C) is progressing from the Swedish coast. The cold water mass has a temperature of 6-7 deg.C. (dark areas).The warm water mass is most likely connected to the area along

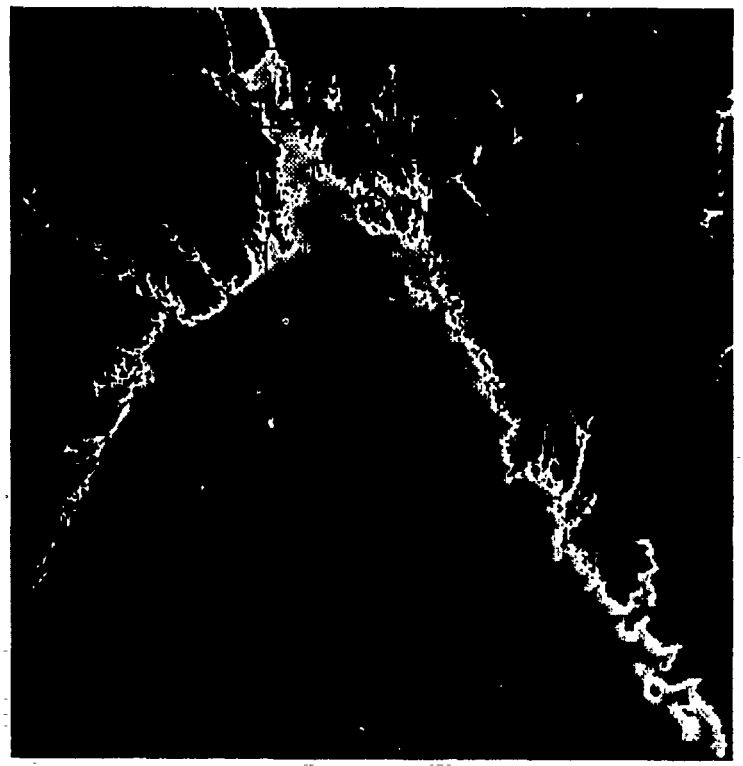


FIGURE 1.
Location of the area of investigation with four specified test sites. Map showes temperature of May 13 from AVHRR channel 4. Dark colours 6-7 and light colours 11-12 deg.C. Test site A: Central Skagerrak. Site B: The Breiangen area. Site C: At the Norwegian coast. Site D: The Hvaler archipelago.

the Swedish coast where the first toxic effects of the algae were reported. The characteristic fresh water plumes from the Glomma and Drammen rivers are seen in the temperature image (9-10 deg. C.).

Thermal information

At site A in the central parts of Skagerrak , a comparison between AVHRR and TM has been made according to Figure 2. The AVHRR can easly identify the water masses but the TM gives a better information of the frontal zones and therefore a better understanding of the transport mechanisms. The TM, however have some sensor striping disturbing the information.The images describe the warm frontal advection across the Skagerrak towards to the Norwegian coast.

In Breiangen (Site B) comparisons have been made between AVHRR and TM to study the resolution differences. The TM shows in detail a vortex which can not be clearly identified in the AVHRR. This small vortex of fresh water from the Drammens River in Breiangen is however clearly seen in the TM image. In spite of the geometric limitations, the AVHRR-image gives good information of the main distribution of the fresh wate The zone of the mixed pixels influenced by land is approximatly 2 km. The geometric error in resampling the AVHRR to TM is in the order of 1 pixel. Because of the limited resolution in the AVHRR images, waterbodys close to the coast are difficult to interpret.

Optical information

Efforts to identify the different water masses within the optical wavelengths in the Skagerrak have been performed for AVHRR, MSS and TM. The SPOT scenes do not cover the outher part of the Skagerrak area. Test site C is from the the Norwegian coast where toxic effects on fish was reported on May 13 (Aksnes et al., 1989). The area where the algae was reported could be identified in the MSS and TM. The secchidepth in the watermasses with this algae were 4-6 meters (Einar Dahl, pers.comm.). The images of AVHRR contain very little information of the algae in these parts of the scenes due to sunglint both in morning and afternoon image. For these sensors, ratioing and differrence images do not show any significant signatures. In all reflective MSS and TM channels the radiance were low in the coastal waters at both sides compared to central Skagerrak. Simple ratioing techniques and chromaticity transformation (Lindell et al. 1986) is showing the best discriminating possibilitis. Figure 4 show ratio images for MSS and TM at the Norwegian coast.

The identification of the water masses within the archipelago (site D), as shown in figure 5, was tried for all sensors. AVHRR, MSS, TM and HRV all showed the main freshwater plume outside the archipelago, indicating the westerly tendency of the water movement due to the coriolis force and the general circulation pattern. The loading of suspended sediments from the Glomma River at May 13 was in the range of 10- 20 mg/l. The differentiation of the small scale distribution of the fresh water is, however, much better described by the high resolution sensors. The radiometric resolution is however not critical in the definition of the plume in the inner archipelago. Here the geometric resolution is more important. The SPOT HkV sensor seems to give more dynamic information than MSS and TM.

CONCLUSIONS

AVHRR and TM data in open areas give similar information of large scale distribution of water masses. However frontal zones and thus dynamic features are considerably more distinct in the TM data. This is to some extent compensated by the frequent passages of the AVHRR. In the archipelago and near-shore areas there are naturally many advantages due to the better spatial resolution of the earth resources satellite than the weather satellite. A combination of the sensors is therefore recommended in many cases.

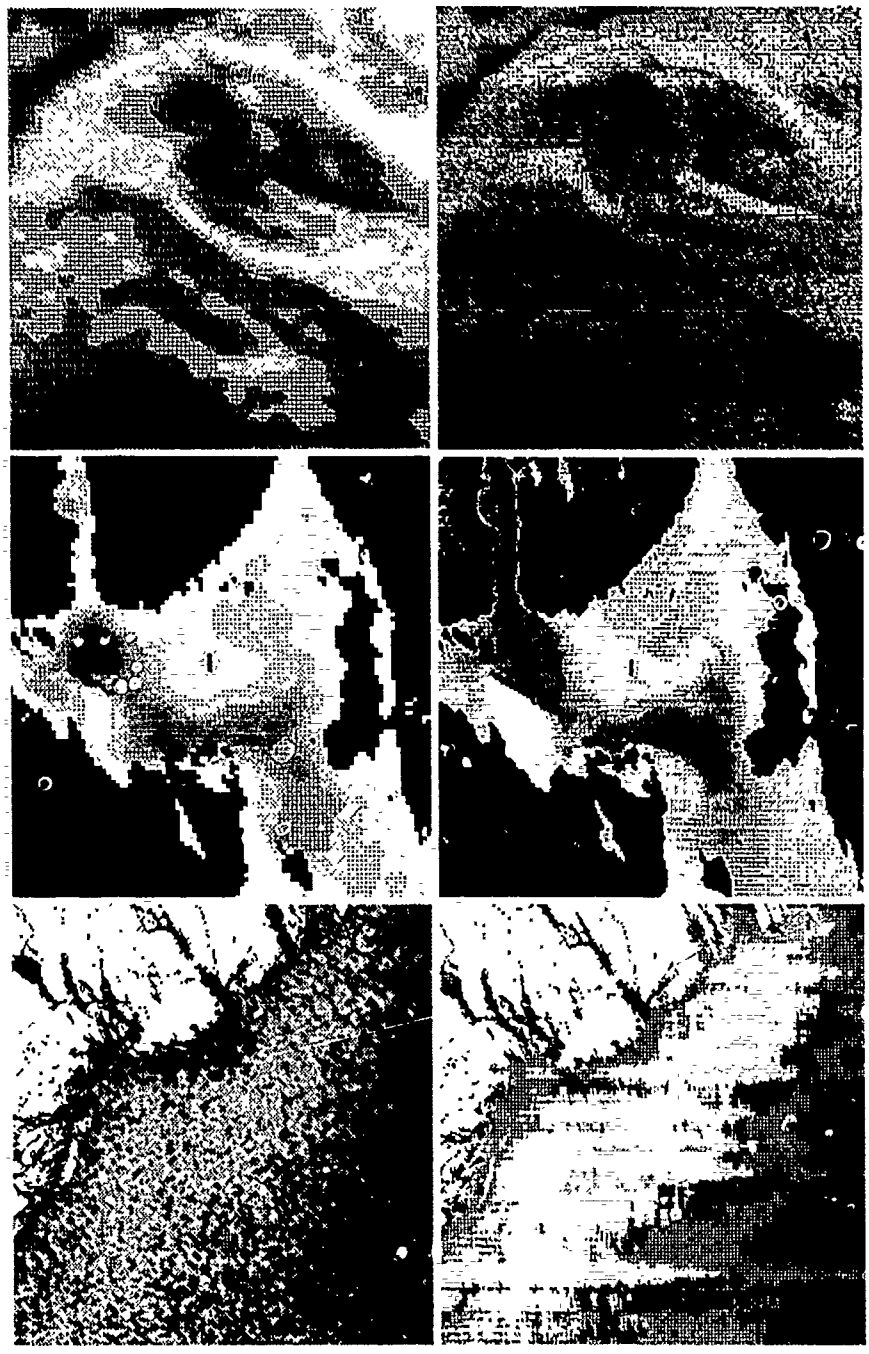


FIGURE 2.
A comparisons between AVHRR (left) and TM (right) thermal image of central Skagerrak (site A). Images are resampled to 240 m pixels. May 13 at 0854 (AVHRR) and 0956 (TM) GMT. The images showes relative temperatures with warm water as light colours.

FIGURE 3.
AVHRR thermal images (left) and TM6 (right) of the Breiangen area (site B) indicating the difference in the geometric resolution of TM and its superiority in identifying small scale phenomena. The images showes relative temperature with warm as light colours.

FIGURE 4.
Ratio images of MSS (left) and TM (right) on May 13, showing the supposed areas (site C) of Chrysocromulina Polylepis (light colours). MSS image is the ratio of channel 1/3 and the TM image is channel 2/4.

The reflectance wavelength of the AVHRR shows very little information in the central Skagerrak, in this particular case due to sunglint. Experiences from other scenes show limitations in the use of those AVHRR-wavelengths, but can be used for discribing heavy loading of suspended sediments. The reflective information of the MSS,TM and HRV is often similar in the water environment, but spatial resolution improvement of TM and HRV are sometimes important in the definition of water masses and water movement. Sometimes the combination of optical and thermal wavelength is essential, but generalization of the use of TM vs. AVHRR in combination with MSS, can not be made due to the geometric differences. The similarities in MSS, TM and HRV in describing water quality is important when calibrations between different sensors are desired. This is important for monitoring, when a large number of availible scenes is necessary.

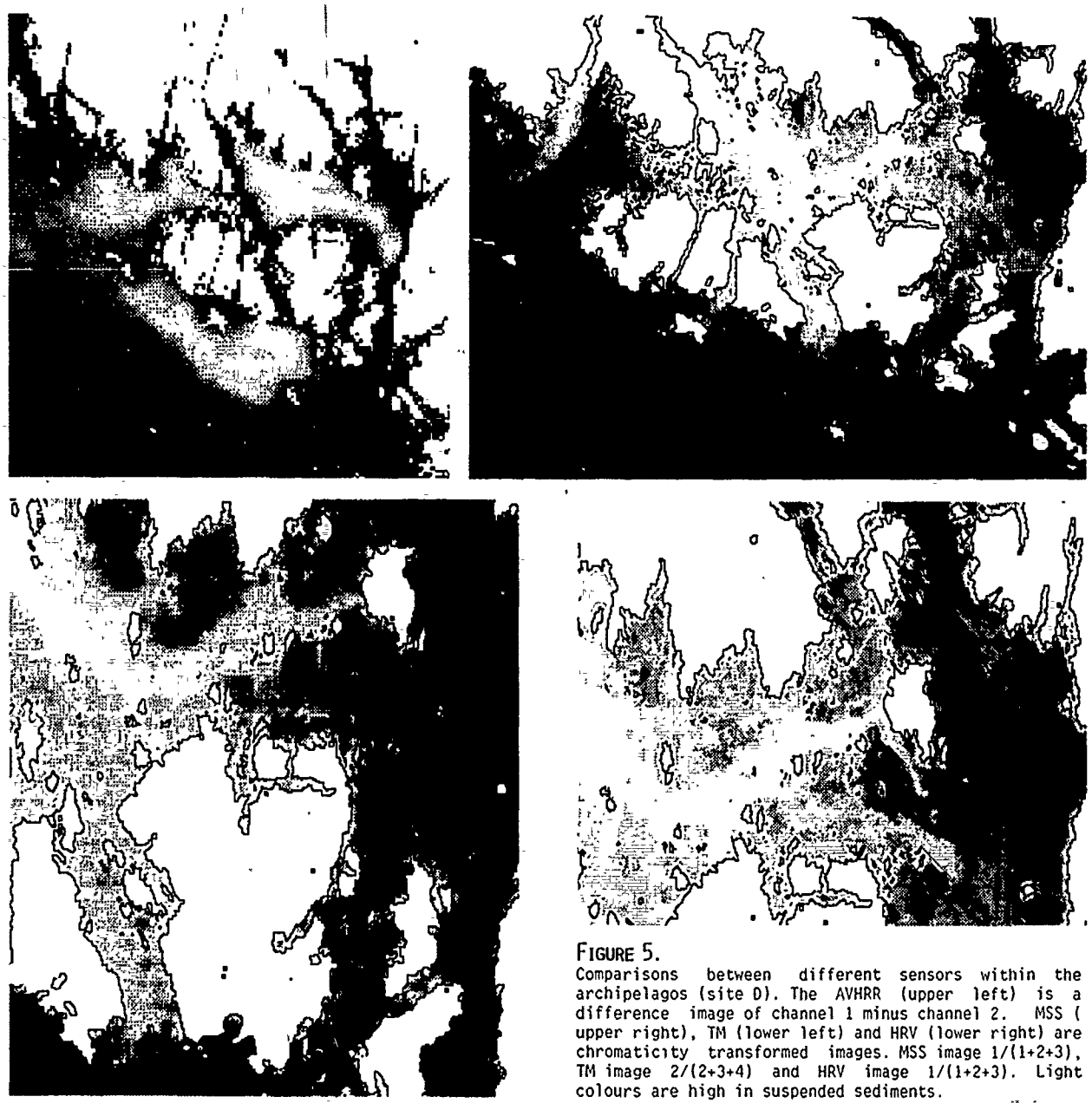


FIGURE 5.
Comparisons between different sensors within the archipelagos (site D). The AVHRR (upper left) is a difference image of channel 1 minus channel 2. MSS (upper right), TM (lower left) and HRV (lower right) are chromaticity transformed images. MSS image 1/(1+2+3), TM image 2/(2+3+4) and HRV image 1/(1+2+3). Light colours are high in suspended sediments.

ACKNOWLEDGEMENT

This project has been supported in parts by a number of different agencies. We acknoweledge the support by the Norwegian Institute for Water Research, the Royal Norwegian Council for Scientific and Industrial Research, the National Swedish Scientific Research Council , the National Environmental Protection Board, Sweden (SNV), the State Pollution Authority of Norway the Swedish Board for Space Activities and Swedish Space Cooperation (SSC). We are grateful to Mats Rosengren (SSC) for his technical support. The AVHRR data were bought from the University of Dundee, England and processed by a program constructed by Mr. B. Karlsson, SNV.

REFERENCES

Aksnes. D.L., J.Aure, G.K. Furnes, H.R. Skjoldal and R.Sætre. "Analysis of the Chrysocromulina Polylepis Bloom in the Skaerrak, May 1988. Environmental Conditios and Possible Causes". Bergen Scientific Centre, BSC 89/1, 1989..

Johannessen, J A , O M Johannessen and P M Haugan, "Remote sensing and model simulation studies of the Norwegian Coastal Current during the algal bloom in May 1988". The Nansen Remote Sensing Center Technical report no. 16, 1988.

Lindell, T., B.Karlsson, M.Rosengren and T.Alfoldi "A further development of the chromaticity technique for satellite mapping of suspended sediment load", Photogram. Eng. and Rem. Sens., 52, 1521-1529, 1986.

THERMAL REMOTE SENSING FOR WALRUS POPULATION ASSESSMENT IN THE CANADIAN ARCTIC.

David G. Barber[1], Pierre R. Richard[2] and Klaus P. Hochheim[3].

[1] Earth Observations Laboratory, Institute for Space and Terrestrial Science. University of Waterloo, Ont.
[2] Department of Fisheries and Oceans, Marine Mammal Management Section,Winnipeg, MB.
[3] Electromagnetic Sensing and Interpretation (E.M.S.I.), Winnipeg, MB.

ABSTRACT

We present results of an experiment to determine the utility of thermal remote sensing for detection and enumeration of walrus (*Odobenus rosmarus*). Concurrent aerial photography and emitted thermal infrared (10.6 μm) imagery were acquired over walrus hauled-out on pans of sea ice in Foxe Basin, NWT, Canada. Digitizing the thermal imagery provides a method for obtaining an objective measure of walrus abundance. The thermal imagery consist of *blobs* of higher digital brightness on a relatively homogeneous background (ocean water and sea ice). The relationship between altitude, maximum and mean pixel brightness and pixel frequency within a particular blob, is compared with photographic *ground confirmation*. Results clearly indicate that thermal remote sensing can provide a cost effective method for obtaining a stratification variable and the relationship between the number of animals and the thermal signature can provide a population estimate.

Keywords: Walrus, Thermal remote sensing, Population surveys, Macintosh II.

1. INTRODUCTION

Walrus are a resource facing pressure from an increasingly large native population, whose expectations for development rely heavily on wildlife . Population assessment of walrus stocks is complicated by the gregarious nature of these animals. Walrus concentrations are typically large and sparsely distributed with rocky shore lines and ice pans being the preferred haul-out habitat.

Aerial censusing is the most appropriate method for obtaining population estimates. Visual surveys are hampered by the fatigue and boredom of surveying large expanses of ice and water almost devoid of walrus and, when groups are encountered, by the difficulty of counting large numbers of walrus in tight clumps. In addition, when surveying land haulouts, accurate counting is hampered by a poor object-background ratio. Photographic surveys can minimize these problems but they are only practical and affordable if sampling schemes can be implemented to reduce survey costs. In most cases, the problems described above lead to imprecise or very costly walrus population surveys.

We address two questions, directly related to these sampling problems.

- Can high altitude thermal surveys provide sufficient resolution to detect walrus concentrations?
- Can high altitude thermal surveys provide a quantitative index of walrus abundance?

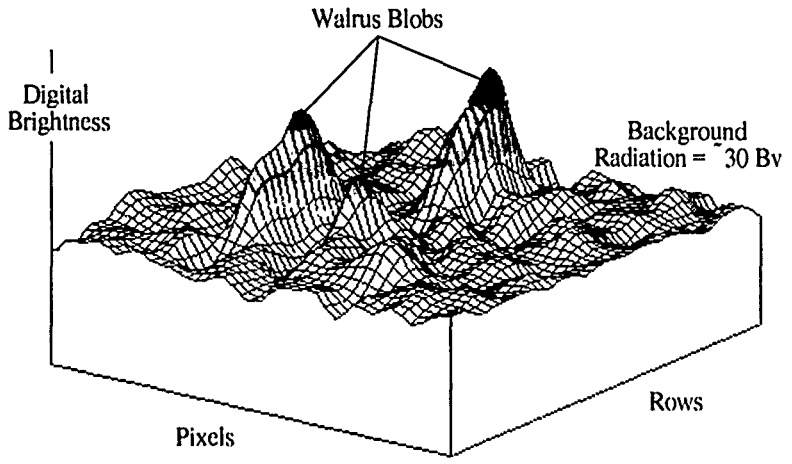


Figure 2. Three dimensional surface of object (walrus) and background (ocean and ice) thermal emissions (from inset of Fig 1).

2. METHODS

On 21 August 1988, we obtained airborne thermal and photographic images of walrus, hauled-out on ice in Foxe Basin, NWT, using the Department of Fisheries and Oceans Remote Sensing System (Yaremchuk and Barber 1985). Thermal imagery was collected with a Forward Looking Infrared (FLIR) thermal imager (peak wavelength =10.6 μm). The FLIR provides real-time display on a video monitor and signal recording on a Super-VHS video cassette recorder (VCR). The FLIR is housed in a specially designed nose cone of a DeHavilland Twin Otter, and can be adjusted at angles of 10° to 90° (down) from the flight plane. Photography was obtained using a large format mapping camera with Kodak Aerocolor negative (2445) film.

During the experimental flight, the FLIR was adjusted at a 20° depression angle. This viewing angle results in a keystone shaped image of the ocean surface. At the center of the keystone (principal point) the swath width (perpendicular to flight direction), ranges from 1.8 km at 610 m (2000 ft) to 8.9 km at 3048 m (10,000 ft) altitude. We use the term range to denote distance from sensor to principal point of the oblique image and altitude for flying height of the aircraft.

Several repeat passes were flown over walrus, starting at an altitude of 610 m (2000 ft) and in increments of 305 m (1000 ft) up to 3048 m (10,000 ft). Low altitude photographs were taken during the 610 m (2000 ft) pass, to obtain an accurate count of each group imaged (ground confirmation). FLIR Super-VHS video records were digitized on a Macintosh II[1] (Mac II) using a DT2255 QuickCapture[2] video board. The board captures a 640 by 480 pixel frame at an 8 bit dynamic range and 1:1 pixel aspect ratio (i.e., square pixels).

The digitized imagery was interpreted using Mac II image analysis software (Barber *et al.* 1989). Walrus thermal emissions appear as bright *blobs* on a relatively homogeneous background (Fig. 1). By examining digital brightness values (Bv) in a three dimensional plot (Fig. 2) of blobs and surrounding background (inset from Fig. 1), is was possible to isolate the walrus emissions (Bv≥30) from the background (Bv<30). This threshold was consistent over all altitudes and abundances of walrus imaged. Statistics from each blob (pixel count, mean Bv {mean}, maximum Bv {max} and Bv standard deviation {SD}) were compared to photographic counts of walrus in each blob.

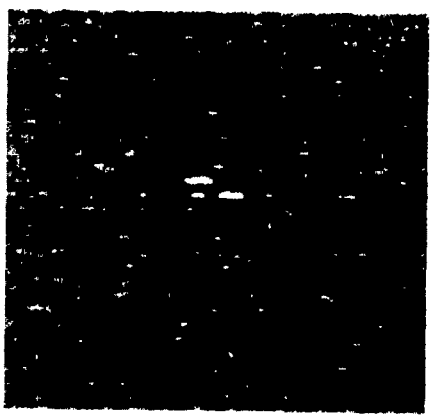


Figure 1. Thermal image of 3 walrus groups on a single ice pan (1524m altitude; 3524m range).

[1]Macintosh II is a trademark of Apple Computer.
[2]QuickCapture is a trademark of Data Translation Inc.

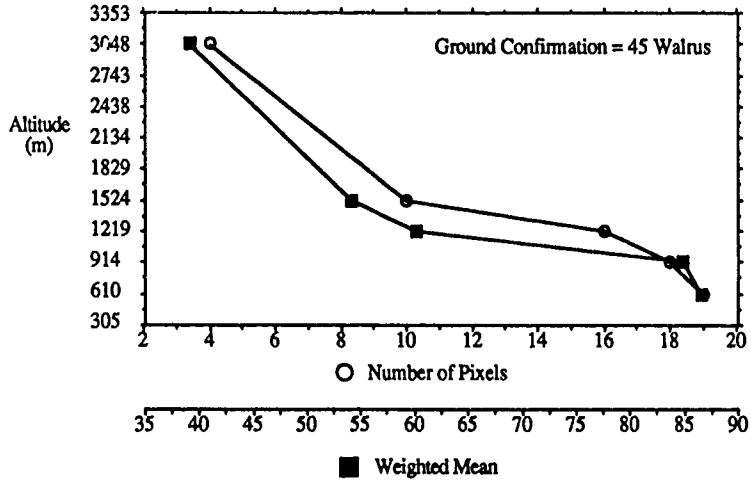


Figure 3. Relationship between altitude, pixel frequency and mean brightness (Bv) above 30.

3. RESULTS AND DISCUSSION

Can a high altitude thermal survey provide sufficient resolution to detect walrus concentrations?

Aerial passes were conducted between 610 m (2,000 ft) and 3048 m (10,000 ft) altitude over three groups of walrus on a single ice pan (Fig. 1). It was possible to discriminate three distinct blobs up to 1524 m (5000 ft) and a single blob thereafter up to 3048 m (10,000 ft) altitude. Descriptive statistics from the blobs show a linear trend between the number of pixels and mean brightness with altitude (Fig. 3). Similar but less distinct trends were also observed for maximum brightness and standard deviation (SD).

The smallest group imaged (9 walrus) was detectable up to an altitude of 1219 m (4000 ft). The group was not identified visually (on the analog recording) but could be distinguished on the digital data (object:background = 33:28 Bv). Given the same atmospheric conditions, we can expect to detect a minimum group size of 9 at a maximum range of 3564 m (2.2 mi) and a minimum group size of approximately 30, at an 8911 m (5.5 mi) range.

Can high altitude thermal surveys provide a quantitative index of the number of animals imaged?

We compared blob statistics from a series of passes over several groups of walrus hauled-out on different pans of ice. Although we have a very small sample, our results suggest that within group blob statistics (i.e., thermal image of a constant number of walrus) are less variable at higher altitudes. This result is reasonable, since at high resolution the FLIR will record individual concentrations of walrus and patches of ice between walrus. At low resolution the sensor integrates over the total radiation emitted from the ice pan, thereby decreasing variability.

Data from a pass at 1219 m (4000 ft) altitude, provided the largest sample (13 observations) with walrus group sizes ranging from 9 to 117. Although these data display higher within group variability than blob statistics at higher altitudes, this is the only altitude with a sufficiently large sample to develop a multiple regression model. Analysis suggests that a linear regression of thermal blob statistics (mean , max, SD, and pixel frequency) can be used to estimate walrus abundance. Inspection of residuals showed no departure from assumptions of the linear multiple regression model (Neter and Wasserman 1979). The model [1] is significant at the 95% level. The coefficient of multiple correlation is 0.874 at 12 degrees of freedom. An analysis of variance on the regression parameters (Table 1) shows that a significant relationship exists between walrus counts and the blob statistics ($F_{[.05, 4, 8]}$; $P = 0.0125$).

4. CONCLUSIONS

Our results suggest that the relationship between thermal blob statistics and walrus concentration can be used to obtain a quantitative estimate of abundance. Specifically, our analysis shows that thermal imagery can be used to:

- detect walrus hauled-out on sea ice at a scale sufficiently coarse to allow complete coverage of a large survey area in a single flight.

Table 1. Analysis of Variance for regression of walrus number and blob statistics.

Variation	DF:	Sum of Squares:	Mean Squares:	F statistic:
Regression	4	12356.196	3089.049	6.496
Error	8	3804.112	475.514	
Total	12	16160.308		

$$\hat{Y}_i = 14.97 + 6.718X_1 + 1.909X_2 - 3.107X_3 + 7.091X_4 \quad [1]$$

Where: Y: walrus count
X1: mean Bv
X2 : SD
X3: pixel freq.
X4 max. Bv

- obtain a walrus abundance estimate from this complete coverage survey.

The multiple regression model provides us with a means of obtaining a walrus abundance estimate from blob statistics. Because of the number of independent variables, we are cautious in our interpretation of the coefficient of multiple correlation. We consider the results promising, but preliminary. Knowledge gained from this experiment will be used to design surveys that specifically address the problem of calibrating the thermal imagery. Precision estimates, for predicted walrus abundance, will be obtained from coefficients of the resulting multiple regression model.

Our ability to detect walrus at a range of 5.5 miles is encouraging from the viewpoint of walrus censusing and habitat related research. Because the FLIR imagery is inexpensive (compared with digital thermal sensors) our approach represents an exciting new application for quantitative thermal remote sensing in resource management.

ACKNOWLEDGEMENTS.

This research was supported by the Department of Fisheries and Oceans and by a Centre of Excellence grant from the Province of Ontario to the Institute for Space and Terrestrial Science.

LITERATURE CITED

1. Barber D.G, J.D. Dunlop, J.M. Piwowar and E.F. LeDrew. Image Analysis on a Macintosh II. IGARS'89, July 10-14, Vancouver, BC. Canada.
2. Neter J. and W. Wasserman. Applied Linear Statistical Models. Richard D. Irwin, Inc. Homewood, Illinois. 1974.
3. Yaremchuk G. and D. Barber. An Aerial Remote Sensing System for Arctic Marine Mammal Research. Remote Sensing in Manitoba. Oct. 1985. pp. 10-13. 1985.

DETERMINATION OF PROBABLE FISHING AREAS FOR THE ALBACORE (Thunnus alalunga) IN CHILE'S CENTRAL ZONE.

María Angela BARBIERI B.(1), Eleuterio YAÑEZ R.(1), Martín FARIAS S.(2) and Raúl Aguilera H.(2).

1: Escuela de Ciencias del Mar,
Universidad Católica de Valparaíso,
Casilla 1020,
Valparaíso
Chile.
Phone 214435. Telex 30389 UCVAL CL.

2: Centro de Estudios Espaciales,
Universidad de Chile,
Casilla 411-3,
Santiago,
Chile.
Phone 6981702. Telex 240523 CENET CL.

ABSTRACT

Starting in 1986 satellite sea surface temperature (SST) imagery, obtained by AVHRR radiometer of NOAA satellites is being used to determine areas of probable fishing for the albacore (Thunnus alalunga). This tuna is fished by artisanal fishermen from Chile's Central Zone.

In order to make such determinations several factors such as SST, upwelling frequency and duration, upwelling strength and chlorophyll A, are taken into account. Information on the first two is obtained from satellite imagery; upwelling intensity is estimated from wind strength data, and chlorophyll A from samples taken on the site.

The Centro de Estudios Espaciales (CEE) provides SST data which is processed in a low cost system and then sent to Universidad Católica de Valparaíso (UCV) through a MODEM connected to an IBM PC/AT computer. A probabilistic model is used to integrate all data and determine how high is the probability of catching an albacore. This result is then indicated on the sea surface temperature satellite image.

The process generates a probable fishing area chart which is distributed to the fishermen for their use to select a fishing zone. In return they provide UCV with environmental and catch information.

KEY WORDS: Satellite, fisheries, albacore, Chilean cost.

INTRODUCTION

Small-scale fishing of albacore exists in the zone of Valparaíso, Chile (32° to 34° S Lt.). The production is mainly for internal consumption. Level of landings is highly variable due to several factors such as environmental changes that affect the resource distribution and the changing price in the internal market. This fishery is important since there are several small fishermen villages that live from albacore capture, besides, its development will grow if the market price increases.

One of the factors that affects and determines the distribution of the albacore is the Sea Surface Temperature (SST). Several investigators have worked in the use of thermal charts of the sea, obtained with satellites, since the late 70's for tuna fishing, achieving good results in determination of resource distribution (CITEAU et al. (1981) and LAURS et al. (1985)).

NOAA AVHRR satellite data usage was introduced for albacore fishery in Chile in 1986 as a mean to help in determining probable fishing zones. This work is a joint effort carried by the Universidad Católica de Valparaíso (UCV). The Centro de Estudios Espaciales (CEE), and the artisanal fishermen of Quintay and Valparaíso.

The CEE has developed a system that allows quick and easy NOAA-AVHRR data reduction and distribution. The ground station is located in Peldehue, an area located 40 Km. north of Santiago, Chile.

The information collected by NOAA satellites is received and processed daily at the ground station and sent to Universidad Católica de Valparaíso (UCV) via MODEM. There, it is merged with the rest of the information to produce the probability chart, in time, for the departing fishing boats.

Results achieved are satisfactory. It was established that albacore is captured in the outer boundary of thermal fronts that are formed as part of the upwelling process. High catch rate was attained when fishing at those locations, specially when fronts were stable for several days and the water was high in chlorophyll concentration wiht predominance of diatoms (BARBIERI et al, 1987). This results are similar to the findings of SVEJKOYSKY & LASKER (1985) in the California coast.

Estimation of probable fishing zones in this work is done using the decision tree to combine some of the environmental variables associated with albacore distribution.

MATERIALS AND METHODS

Environmental Data: two types of information are used: oceanographic and meteorologic. Sample collection trips are carried on board of albacore fishing boats. Samples are taken at different levels, surface and deep waters, analysis is performed to establish oxygen contents, salinity, chlorophyll A, water color and temperature. Temperature of the surface samples is used for calibration of satellite data. Meteorological data is obtained from the station at Punta Angeles lighthouse (Lat. 33°01'S and Long. 71°39' W) collected by the Meteorological Center of Viña del Mar of the Chilean Navy.

Biological and Fishing Data: it is collected at the landing points. Size and weight of the captured albacore is recorded in addition to the fishing effort carried in the operation as well as any other data from the fishing area.

Since the beginning of this project is 1986, biological and fishing data were collected by UCV personnel, starting in 1989, fishermen have been trained to gather the data and fill out the forms by themselves. They are able, also, to use now the SST charts. Charts are sent to the associated fishermen through their organizations, the same channel is used in return to obtain the data forms.

Satellite Data:

1. Hardware and Software.
The reception of the satellite signal is performed through a 12-meter parabolic dish antenna. The antenna is computer driven for first acquisition of signal and it also has capability for autotracking the satellite. The signal goes then through a receiver-demodulator that removes carriers and possible noises that could have been collected as part of the transmission-reception process. A Bit-synchronizer takes care of the rest of the analog noise yielding a reconstructed digital signal.

The reconstructed signal is then fed to a decommutator which separates the 5 AVHRR channels and the TOVS data, these 6 streams are recorded on analog tape for archive and backup and, at the same time, 3 AVHRR channels are input to the computer for real-time storage in its hard disk.

Decommutation, selection and processing of NOAA data is achieved using locally generated hardware and software. The specialized hardware has been kept to a minimum to reduce the maintenance and replacement costs and the main effort put in a software system that would run in a commercially available computer.

The IBM-type PC/AT computer was chosen as the sustaining hardware core of the system. Its characteristics: speed, memory size, available peripherals, arithmetic coprocessor and its open architecture made it a good choise at the time. The software was developed using C language for speed and maintainability, eventually it will also allow portability (AGUILERA, 1988).

C language, being used by a vast majority of professional programmers and system developers, has a very good suport in terms of pre-built routines and compilers. There are several good manufacturers of all kind of tools for software development using this language which yields to costs reduction, time and money wise, that are by temselves enough reason to choose it.

2. Processing Programs.
Data stored in the computer's hard disk includes all ancillary information such as calibration, timing, etc. Files are in standard MS-DOS format. No attempt is made, at this level, to compress, validate or verify the data, except for very basic selfconsistency checks to avoid filling the storage space wiht noise. A full validation is performed post-pass by a pre-processor program which also performs coarse geometric corrections using a pixel repetition technique. This software module displays on the computer screen the complete image obtained during the contact with the satellite and allows the selection of areas for further processing. The user, interactively, can specify any area using a movable window controlled by the keyboard standard cursor keys. Window size is also user selectable.

The pre-processor program creates files containing all the available information related to the selected area. In general, the amount of information stored on these files is independent of the window size as pixel and line sampling is used to produce always a 256 x 256 image.

The information stored in the new files can then be displayed, filtered, converted, distributed in electronic form and printed for hardcopy distribution. The data format is maintained throughout the process to allow the different modules of the system to be used at any stage of it. Thus, spatial filtering, line fixing and/or cloud/land masking can be done to the raw data, the intermediate uncorrected temperatures or the final moisture absorption corrected temperature charts. This capability is very useful specially to detect and mask out low level clouds and fog.

Hardcopy charts are created in two possible forms, using dot patterns obtaining over 20 gray levels or, using subscript characters available on dot-matrix printers. The first form is used for quick checks of shapes and trends, the second one is for field use as it allows better reading of actual values under a wide range of light conditions. As it has turned out, it is a matter of taste, some users prefer the shades and others the little characters.

AVHRR channels 4 and 5 are converted to degrees centigrade using the standard procedure of computing the equivalent temperature from the measured radiating energy. Calibration parameters and

published coefficients (LAURITSON et al, 1979) are used in the process. Computed temperatures are then combined to compensate for atmospheric water absorption. An additional correction is performed to remove the effect of the observation angle. This correction was locally devised and it was necessary to make it latitude dependent to achieve results that are closer to field truth (the in-situ measurements).

AVHRR channel 1 is used in raw form to help in land and cloud masking. Land is masked out to help in site location by using the coast line as reference, automated georeferencing is, at the time of this writing, in the process of development. Even if the chart is georeferenced land masking is still a good help for the fishermen.

Clouds are the worst enemy of the SST remote detection as what the satellite sees is the top temperature of them instead of the sea. High clouds are rather easy to detect as their top temperature is very low compared to the sea surface but low clouds and fogs tend to have temperatures quite similar to the sea. The only way to detect them semi-automatically is with the use of visible channel data. Fog and cloud masking is done with an interactive graphics program that presents the thermal and visible images on screen as well as the scatter plot of both, the detection becomes very easy as the plot shows totally different patterns for each of the mayor features of the images: clouds, fog, open water, land, etc. The user can then mask out those pixels that belong to the particular feature of interest.

RESULTS

Probable Fishing Areas Determination.

The Decision Tree is used to determine the fishing probability within the area of study. The expected catch rate (T) is computed, expressed in number of captured fishes in a given zone by one small fishing boat, this is a discrete variable. The events are mutually exclusive and collectively exhaustive.

The source to establish the probabilities was obtained from satellite thermal data collected since the 1986 fishing season up to the current year (1989). 1986 is considered as a normal year, in 1987 the "El Niño" phenomena affected the zone of study. Two facts were noticed during the "El Niño" year: A positive anomaly of 1° C on the SST and the absence of upwelling events. As result of this, albacore capture fell to zero (BARBIERI et al, 1989).

Several relationships were established between the catch rate of albacore and environmental variables such as the SST, sea color (C) and the levels of chlorophyll A. Upwelling index is derived from wind data (BAKUM, 1975) and the shape, location and maturity of the thermal front from the SST satellite images. Four of these relationships are used for the model.

The components of the model are:

a) Decision Alternatives: four are considered,

1) The average monthly SST of the area where albacore is captured.
2) The upwelling index (UI)
3) The temperature gradient (TG), that indicates the presence of a thermal front.
4) Productivity, considering the chlorophyll A levels (CL).

b) Event: are those that can occur as a result of each decision alternative, they are shown in table I.

c) Probabilities and Results: are calculated, the figures obtained are expressed in terms of number of albacores susceptible to be captured by one "bongo" (a small engine powered fishing boat, maximum 9 meters long). The fishing technique is trolling using 4 or 5 lines per boat. Results are expressed in probable total capture per boat per day (24 hours) at sea: high probability "A" is assigned when the expected number is over 50 albacores, medium "M" for 10 to 49 and, low or null "B" for numbers below 9.

Model Implementation

SST changes throughout the fishing season, according with the results obtained in previous seasons, albacores are captured mainly in 17 C in January, 18 C in February, 17 C in March and 16 C in April and May. Fishing is possible in temperatures that go from 14.5°C to 18.7°C.

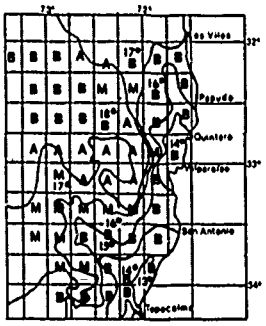
Based on historical data, the monthly upwelling index is calculated for each month (UIsm), this is compared with UIj (daily upwelling index), if the latter is higher, the upwelling process is evolving. IUsm is also compared to the average upwelling index for the last 5 days (UIj-5), if this last one is higher, the upwelling event continues. In both cases probability of high catch rate are low (B).

If IUsm is higher than UIj-2 - average of the two previous days - the upwelling relaxed. Finally, if UIsm is higher than UIj-10 - previous 10 days average - the period correspond to absence of upwelling. In this two cases, the decision tree model continues by computing the temperature gradient per cell (GTc), cells are defined as 15'x15' (Lat/Long), if GTc tends to zero for a given cell probable catch rate is low (B), thermal discontinuities appear in the form of high gradient; in this case the chlorophyll A level (CL) is used. If CL is below 0.7 mg/lt, probable catch rate is medium (M). Any higher value yields a high probability (A). The chart of harvesting probability per cell (fig 1) is superimposed over the satellite thermal image to obtain the final chart of probable fishing areas (fig 2). This last process is done manually, but a software is under development to automate it.

On January 27, 1989, the albacore fishermen caught around 32 pieces per boat, in the outer boundary of the upwelling plume, in waters 17°C to 18°C, 60 miles in front of Valparaíso. Theorically, thru the model, a high capture probability had been estimated for that zone (fig 2). Therefore, the estimation was higher than the actual capture.

On January 30, because of the weather, the small boats couldn't reach the high capture probability zones determined by the model. Fishing, instead, was performed between Valparaíso and Quintay 15 miles offshore in waters 15°C to 16°C, catching 4 to 0 pieces per boat. The latter places had been marked as low capture probability zone by the model (fig 2).

Figure 1 shows cold waters south of San Antonio (Lat 33°35'S) as a result of an upwelling event, when this happens, fishermen must stay in port due to sea conditions, afterwards, when the wind



Fish Harvesting Probability Chart Per Cell.

A= High;

M= Medium

B= Low

Figure 1.

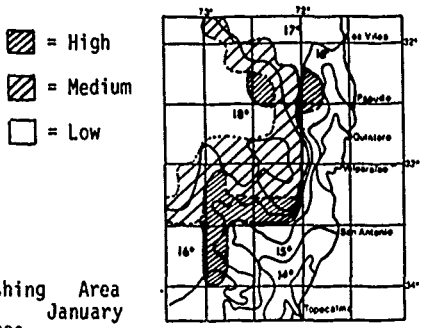


Figure 2.

Probable Fishing Area Chart for January 26 thru 31, 1989

deminishes they go out looking for the stock without knowing where they can find the favorable waters for albacore fishing. The method used before was to continually monitor the surface temperature of the sea using a regular thermometer. This could last for several days with no results, increasing the operational costs, fuel alone is 48% of the total without considering salaries. Now, with SST satellite images and the probable fishing areas charts they go directly to the fishing zone reducing searching times which, in turn, decreases the operational costs significantly and increases the success of each fishing trip, thus, days at sea without fishing are avoided. Fisherman are increasingly interested in using this techniques for the obvious advantages it has.

STRETTA & SLEPOUKHA (1983) compute the probability of fishing and its evolution at the intertropic Atlantic using cells of much larger size 2° Lat and 5° Long, forecast are good for 13 days. Predictions are given to the tuna deep sea fishing fleet. This work has been done for a fleet composed of small boats operating from shore with very short autonomy in a zone where environmental parameters vary very rapidly. The method used allows the production of probable fishing charts that are valid for a given day, in general, two charts per week are needed as a minimum.

Techniques and methods devised for developed countries are in general not applicable as a package to the various third world conditions, they have to be tested, adapted and modified to truly fill the needs that exist in those countries as is the case of Chile.

Table 1. Model Components, Decision Alternatives, Events and Results.

DECISION ALTERNATIVE	EVENT	RESULT
1) SSTm Monthly sea surface temperature	SST >3C SST ± 2C	T1
2) UIsm Monthly upwelling Index	UIsm<= UIj evolving upwelling UIsm<= UIj-5 continuous process UIsm>= UIj-2 relaxing upwelling UIsm>= UIj-10 no upwelling for several days	T2 T3
3) TG Temperature gradient	TG = 0 no thermal front TG > 0 thermal discontinuity	T4 T5
4) CL Chlorophyll	CL < 0.7 mg/lt CL > 0.7 mg/lt	T6 & T8 T7 & T9

Acknowledgments

The authors are profoundly grateful to their institutions UCV and CEE, who sponsored this work since 1986. Starting in 1988 this work was carried out with a grant from the International Development Centre, Ottawa, Canada. Special mention deserve the fishermen of Valparaiso and Quintay as well as the people of UCV and the CEE that worked hard to make this project work.

REFERENCES

1. CITEAU, J.; J.Y. LE GALL et R. PIANET, "Le satellite Météosat et la flotille thoniére intertropicale", la Peche Maritime 1245, pp740 - 744, 1981.

2. LAURITSON, L.; NELSON G.J.; PORTD F.W. "Data Extraction and Calibration of TIROS-N/NOAA Radiometers. NOAA Technical Memorandum NESS 107.1979.

3. LAURS, R.M.; FIELDER, P.C. & D.R. MONTGOMERY, "Albacore tuna catch distributions relative to environmental features observed from satellites", Deep Sea Res., Vol 31, N°9, pp1085-1099, 1984.

4. STRETTA, J.M. & M SLEPOUKHA, "Les satellites, l'ordinateur et l'aide á la peche thoniére", La Peche Maritime 1258, pp38-40, 1983.

5. SVEJKOVSKY, J and R. LASKER, "Effect of thermal structure on albacore tuna distribution", In: "Applications of remote sensing to fisheries and coastal resources", R. Amidei (Ed). Report of California Sea Grant Work Shop. San Diego (USA), 30 Nov 1983, pp9-13, 1985.

DEVELOPMENT OF ALGORITHMS FOR REMOTE SENSING OF MARINE *TRICHODESMIUM*

Gary A. Borstad[1], Jim F.R.Gower[2],
Edward J. Carpenter[3] and John G. Reuter[4]

[1]G.A.Borstad Associates Ltd., 100-9865 West Saanich Rd, Sidney, BC Canada V8L 3S1 Ph 604-656-5633 Fax 604-656-3646

[2]Institute of Ocean Sciences, PO Box 6000, Sidney, BC Canada V8L 4B2 Ph 604-356-6558 Fax 604-356-6479

[3]Marine Sciences Research Center, State University of New York, Stony Brook, New York USA 11794, Ph 516-632-8696

[4]Dept Biology, Portland State University, PO Box 751, Portland Oregon, USA 97207 Ph 503-229-3194

ABSTRACT

Trichodesmium is a major component of the global carbon cycle, but because of its sporadic occurrence it is extremely difficult to study by conventional shipboard methods, and has not often been reported in imagery from the Coastal Zone Color Scanner. We made spectral reflectance and absorption measurements on natural and concentrated populations of *Trichodesmium* as part of a research cruise in the Caribbean in November 1988. Physiological measurements including carbon and nitrogen fixation, oxygen consumption and production were also made. This paper discusses the optical data, focussing on small spectral absorption features resulting from the nitrogen containing pigments; at 495 and 567nm by Phycoerythrin, and at 625 nm by Phycocyanin; and on the relatively high scattering caused by gas vacuoles in *Trichodesmium* cells. Reflectance data from a series dilution experiment are used to evaluate the detection limits of algorithms based on these features.

INTRODUCTION

The cyanobacterium Trichodesmium spp (Oscillatoria) is the most abundant and active nitrogen fixing species in the plankton of tropical and sub-tropical seas (Carpenter, 1983). The organism plays an important role in the global flux of nitrogen and carbon. However, its appearance in infrequently travelled tropical seas far from major research laboratories, combined with the difficulty of keeping it alive in culture, makes it hard to study. The global distribution of the species, frequency of blooms and estimates of rate of N_2 fixation and primary production could be assessed more accurately by the use of remote sensing, especially when improved ocean colour sensors are launched in the next decade.

Kuchler and Jupp (1988) presented the Shuttle photograph reproduced in Figure 1 showing a massive Trichodesmium bloom off the coast of Australia. DuPuoy et al, (1988) and Kuchler and Arnold (1986) have shown examples of similar blooms imaged using Landsat and Coastal Zone Color Scanner data.

In this paper we present preliminary results from studies using a non-imaging spectrometer on board ship to obtain information on the remote sensing signature of the organism. Trichodesmium contains both chlorophyll a and the bilin pigments phycoerythrin and phycocyanin which have characteristic absorption spectra. These features should become visible in remotely observed spectra at some cell concentration. Our studies were undertaken with a view to determining the detection limits for these features and in general to preparing for the higher spectral resolution remote sensing instruments now becoming available.

A major difficulty in Trichodesmium studies is the non-uniformity of its distribution on a wide range of space scales. At the largest scales, the organism forms concentrated blooms in very clear water. Gas vacuoles cause it to float close to the surface. At intermediate scales, convergence in the surface layers of the ocean concentrates blooms into linear surface slicks as shown in Figure 1. At the smallest scales, individual chains of cells (trichomes) are concentrated into colonies of tens to thousands of individual trichomes.

On a ship, considerable luck is required to encounter a natural bloom of Trichodesmium. However, we can take advantage of the above-noted colonial behaviour to collect pure concentrated samples of Trichodesmium even from mixed phytoplankton populations. This is done by physically picking colonies out of the plankton collected by near-surface net tows.

We report here preliminary measurements on samples artificially concentrated in this way, of trichome optical properties, including spectral reflectance and absorption, relevant to remote sensing of this important species from the air or from space. All measurements were made on samples within a few minutes of their collection.

2. OPTICAL MEASUREMENTS

Remote sensing measurements were made on cruise number 18 of the University of Miami vessel RV Columbus Iselin, between November 4 and November 24 1988, along the cruise track shown in Figure 2, from Miami through the Bahamas to the Honduran Island of Roatan, and to Limon, Costa Rica, crossing a wide area where Trichodesmium is known to occur. Several interrelated experimental programs were being conducted on this cruise, providing supporting measurements of carbon and nitrogen fixation, nutrient uptake and cell counts.

2.1. ABSORPTION SPECTRA

Absorption measurements were made on Trichodesmium colonies and, for comparison, on a culture of Synechococcus, a culture of which was maintained on the ship during the first few days of the cruise. Samples were washed through filtered sea water and collected on moist GF/F filters. Absorption was measured according to the method of Mitchell and Kiefer (1988), using the IOS spectrometer (Walker et al 1975) looking up through the filter at an Osram halogen bulb.

Figure 3 shows typical relative absorption spectra for dispersed trichomes of Trichodesmium and for Synechococcus. The absorption shows a maximum peak at 460nm, with a second peak at 678nm, due to chlorophyll a. Peaks at 494nm, and 551nm are ascribed to phycoerythrin and those at 630nm to phycocyanin. Spectral peaks are sharper for the extracted samples.

2.2. REFLECTANCE SPECTRA

Reflectance spectra were collected for the concentrated Trichodesmium samples by viewing them from above under solar illumination. In order to avoid bleaching, this measurement was made within a few seconds after the cells on the filters were first exposed to sunlight, though no consistent evidence of bleaching was observed in subsequent measurements. The observed radiance spectra were normalized by the spectrum from a moistened, white filter viewed immediately afterwards in order to provide reflectance factor spectra.

Results are shown in Figure 4 for concentrations on the filters of 1300 (upper curve), 2600, 5700, 6600, and 23000 (lowest curve) trichomes.cm^{-2}. The centre curve shows some variation in spectral form, but the samples show the expected trend from the thin covering of the filter at the lowest concentration to near optical thickness for

the lowest curve.

2.3. REFLECTANCE SPECTRA OF A SERIAL DILUTION OF CONCENTRATED TRICHODESMIUM

Because a natural bloom was not encountered on this cruise, we created bloom-like conditions by putting the contents of a plankton tow into a small, clear plastic container suspended in the sea near the ship. While this artificial situation will not mimic the real optical behaviour of a Trichodesmium bloom, we feel that it a good first approximation, since in a bloom the concentrations near the surface would be very high.

At one station off Roatan Island, (Honduras) the contents of a 64um plankton tow were poured into the container, which was 10cm diameter and 25cm deep, with a rounded bottom which helped avoid reflection from the container itself. Upwelling radiance spectra were obtained by pointing the spectrometer into the container from the deck of the ship. A file of 100 or more individual spectra were obtained for each dilution, from which averages of 10 to 20 spectra from inside the container were later selected to represent that dilution. After each raw radiance data file was obtained, the contents of the container was diluted with an equal volume of sea water, and another file of spectra were measured. After the first two spectra were obtained (0 and 1 in Figure 6), the plankton was gently screened through a 64um mesh plankton net in an attempt to wash out some of the contaminating diatoms and nanoplankton. Samples at each dilution were counted using a microscope and Sedgewick-Rafter cell.

The raw upwelling radiance spectra were transformed into reflectance factor spectra by dividing by a radiance spectrum of a white card obtained at the same time. The spectra obtained are shown in Figure 5. The series of 6 samples (5 dilutions) range from 1441 trichomes/ml (top spectrum) to 46 trichomes/ml (lowest spectrum). Samples for chlorophyll, phycoerythrin and Trichodesmium filament number were also obtained.

3. IMPLICATIONS FOR REMOTE SENSING

The absorption spectrum in Figure 3 for Trichodesmium indicates sufficient similarity to other marine phytoplankton that measurements using the blue/green ratio technique, for example with CZCS satellite data, will probably be satisfactory at "sub-bloom" conditions (Lewis et al. 1988). However, specific identification of Trichodesmium will have to rely on fluorescence of phycoerythrin (Hoge and Swift 1983), or on its absorption properties.

Methods based on phycoerythrin fluorescence may be confounded by interference from Synechococcus, the other abundant blue-green alga in these waters. In Figure 3 the higher concentration of phycoerythrin in Synechococcus is indicated by the increased absorption at 546nm.

The reflectance spectra in Figure 4 show increasing saturation of the filters, with the spectra for the highest concentrations implying a cross-section area per trichome on the order of about $10^{-4}cm^2$. This agrees very well with mean cell diameters of 10u and trichome lengths of 1mm. The lowest reflectance spectrum in Figure 4 is very similar to that of terrestrial vegetation. Surface mats of Trichodesmium will show the characteristic chlorophyll "red edge" in their spectrum, a feature which should be detectable in remote sensing measurements over deep water by the radiance difference between 680 and 720nm.

We anticipate that this feature can be exploited in observing surface blooms, though present satellite sensors are not optimized for this.

This region of the spectrum has been investigated for observations of water-leaving radiance in connection with measurements of solar-stimulated phytoplankton chlorophyll a fluorescence (Neville and Gower 1977, Gower and Borstad 1981). The need to make remote sensing observations in narrow and well-defined windows that minimize contamination by atmospheric absorption features, leads to a need for specialized sensors such as the FLI (Borstad et al, 1985) and the CASI (Borstad and Hill, 1989).

Other features are also visible in the spectra of Figures 3, 4 and 5. Phycoerythrin absorption at 495 and 548nm is visible in Figures 3 and 4, and at the highest concentrations in Figure 5. Phycocyanin absorption at 630nm is similarly apparent.

The "phycoerythrin" dip in the curves in Figure 5 centred at 495nm can be quantified by measuring the difference between the reflectance at that wavelength, and the reflectance interpolated from measurements at 470 and 532nm.

Figure 6 shows a plot of this difference against the Trichodesmium concentration in trichomes/ml. The plot is reasonably linear (r^2=.92), and indicates a measurable reflectance drop at concentrations down to near 200 trichomes/ml. This is a high concentration for many tropical waters, but is relatively low for bloom conditions (Devassy et al 1978, Eleuterius et al 1981). Similar detection limits can be deduced for the other features in Figure 6.

The over-all shape of the curves in Figure 6 indicates scattering over the full range of wavelengths plotted, as would be expected for this near-surface "bloom". High scattering at wavelengths greater than 710nm should allow Trichodesmium to be visible in AVHRR

and Landsat imagery. However this scattering would be misinterpreted by the standard CZCS processing algorithm, which assumes zero water-leaving radiance at wavelengths longer than 670nm.

Our results confirm the findings of Dupuoy et al (1988) and Kuchler et al. (1986). Existing sensors will image Trichodesmium by detecting the increased brightness at all optical wavelengths. Higher spectral resolution and sensitivity with future sensors should give greater sensitivity, and permit positive identification of pigments such as phycoerythrin and phycocyanin.

4. REFERENCES

Borstad, G.A, and D.A. Hill, 1989, "Using visible range imaging spectrometers to map ocean phenomena", Paper presented at International Congress on Optical Science and Engineering, Paris, France, 24-29 April.

Borstad, G.A., H.R. Edel, J.F.R. Gower, and A.B. Höllinger, 1985, Analysis of test and flight data from the Fluorescence Line Imager, Canadian Special Publication of Fisheries and Aquatic Sciences, No. 83, 38pp.

Carpenter, E.J. 1983. "Nitrogen fixation by marine Oscillatoria (Trichodesmium) in the world's oceans", pp 65-103 In E.J.Carpenter and D.G.Capone (eds) "Nitrogen in the marine environment", Academic Press, New York, 900 p.

Devassy, V.P., P.M.A.Bhattathiri and S.Z.Quasim. 1978. "Trichodesmium phenomenon", Indian J. Mar. Sci., 7, 168-186.

Dupouy, C., M.Petit and Y.Dandonneau, 1988. "Satellite detected cyanobacteria bloom in the southwestern tropical Pacific: implication for oceanic nitrogen fixation", Int. J. Remote Sensing, 9, 389-396

Eleuterius, L., H. Perry, C. Eleuterius, J.Warren and J.Caldwell, 1981. "Causative analysis on a nearshore bloom of Oscillatoria erythrea (Trichodesmium) in the northern Gulf of Mexico", Northeast Gulf. Sci. 5, 1-12.

Gower, J.F.R. and G.A. Borstad, 1981, "Use of the in-vivo fluorescence line at 685 nm for remote sensing surveys of chlorophyll a", in J.F.R. Gower (Ed) "Oceanography from Space" Plenum Press, 329-338.

Hoge, F.E. and R.N. Swift, 1983 "Airborne dual laser excitation and mapping of phytoplankton photo-pigments in a Gulf Stream warm core ring", Appl. Optics, 22, 2272.

Kuchler, D.A. and D.B. Jupp, 1988. "Shuttle photograph captures massive phytoplankton bloom in the Great Barrier Reef", Int. J. Remote Sensing, 9, 1299-1301

Kuchler, D.A. and N.P.Arnold., 1986. "Identification and characterization of a massive phytoplankton bloom within the Capricorn Channel, Great Barrier Reef, Australia", Proc. Beijing Int. Symp. Remote Sensing, Beijing, PRC, November 18-22, 1986.

Lewis, M.R., O. Ulloa and T. Platt. 1988. "Photosynthetic action, absorption, and quantum yield spectra for a natural population of Oscillatoria in the North Atlantic", Limnol. Oceanogr., 33, 92-98.

Mitchell, B.G. and D.A. Kiefer, 1988, "Chlorophyll a specific absorption and fluorescence excitation spectra for light-limited phytoplankton", Deep Sea Res., 35, 639-663.

Neville R.A., and J.F.R. Gower, 1977, "Passive remote sensing of phytoplankton via chlorophyll a fluorescence", Journal of Geophysical Research, 82, 3487-3493.

Walker, G.A.H., V.I. Buchholz, D. Camp, B. Isherwood J. Glaspey, R. Coutts, A. Condal and J. Gower, 1974, "A compact multichannel spectrometer for field use", Rev. Sci. Instrum., 45, 1349-1352.

Figure 1 Shuttle photograph illustrating the spatial structure of a massive Trichodesmium bloom.

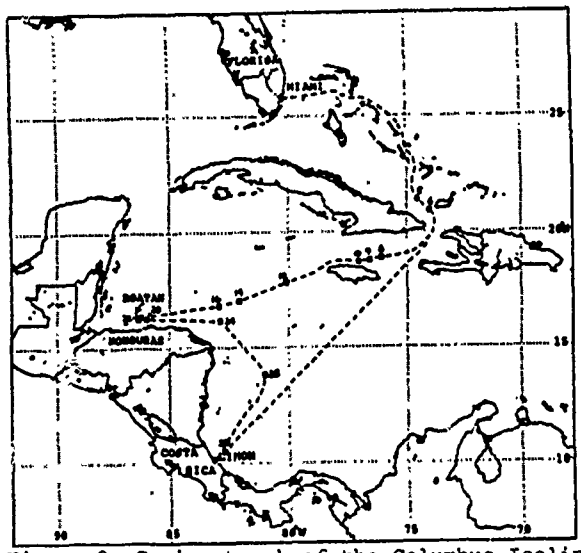


Figure 2 Cruise track of the Columbus Iselin in November 1988 showing station numbers and ports of call.

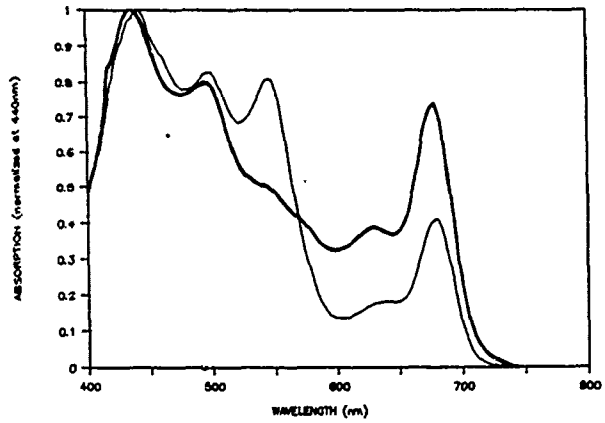


Figure 3 Spectral absorption of samples of Trichodesmium (heavy line) and Synechococcus (light line) measure on the cruise.

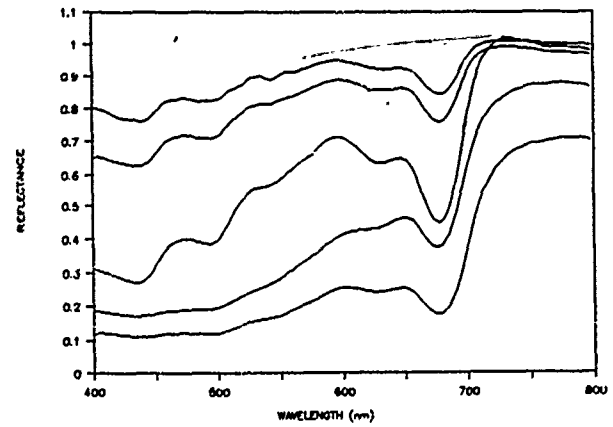


Figure 4 Spectral reflectance of different concentrations of Trichodesmium on white filter papers.

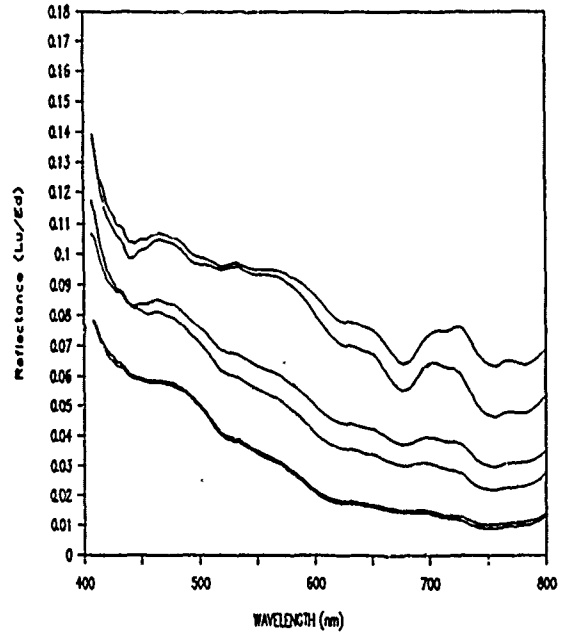


Figure 5 Spectral reflectance of different concentrations of Trichodesmium measured in a small transparent container suspended at the surface of the sea in deep water.

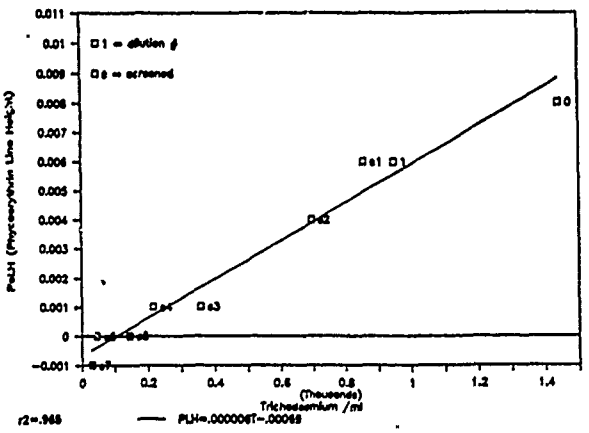


Figure 6 Correlation of the depth of the absorption minimum at 495nm shown in Figure 5, with the Trichodesmium concentration in cells/ml.

METEOROLOGICAL SATELLITE IMAGERY FOR MARINE WEATHER FORECASTING

Neil McLennan and Laurie Neil

Pacific Weather Centre
Atmospheric Environment Service
Vancouver, B. C., Canada

ABSTRACT

One of the major forecast problems facing the marine meteorologist is evaluating the potential for intensification of weather systems (cyclones). Many cyclones intensify as they move through the mid-latitudes and a sub-set of them will become extremely dangerous storms. Given the seriousness of these explosive developments, the weakness of numerical atmospheric simulation models over the eastern Pacific Ocean and the sparsity of conventional meteorological data in that same area, weather forecasters require an objective technique. This paper describes a satellite based forecast technique developed at the Pacific Weather Centre that assesses the likelihood of rapid intensification.

Key Words - weather; forecast; meteorology; cyclones and GOES.

1. Introduction

Over the years, many papers have been published that describe the dynamics and thermodynamics that occur during the formation and intensification of weather systems (cyclogenesis). However, the set of tools for solving the specific forecast problem of "How much intensification?" is still incomplete.

Of special concern to forecasters are systems that intensify very rapidly, colloquially called "bombs" (Sanders, 1980), which are a sub-set of all marine cyclones. In an average year about sixteen such storms affect the northeast Pacific Ocean. Because of the scarcity of conventional meteorological data and the questionable reliability of computer generated analysis fields, satellite information plays a crucial role in assessing the state of the atmosphere and the potential for cyclonic development.

A satellite based "bomb" forecast system has been developed at the Pacific Weather Centre and is being used by duty forecasters to assist in the prediction of explosive cyclogenesis.

2. Data

The data set used to develop the forecast system comprised a total of 111 cyclones from the winters of 1986-87 and 1987-88. Included in this set were 51 "bombs and 60 "non-bombs". The critical value for dividing the two classes is a Bergeron value of one where a Bergeron was defined by Sanders (1980) as

$$B = [\, p/24] * [\sin / \sin(\ /3)]$$

p = observed 24 hour pressure change
= latitude

The geographical area under consideration was the eastern Pacific Ocean east of 150 degrees west longitude and north of 40 degrees north latitude (figure 1).

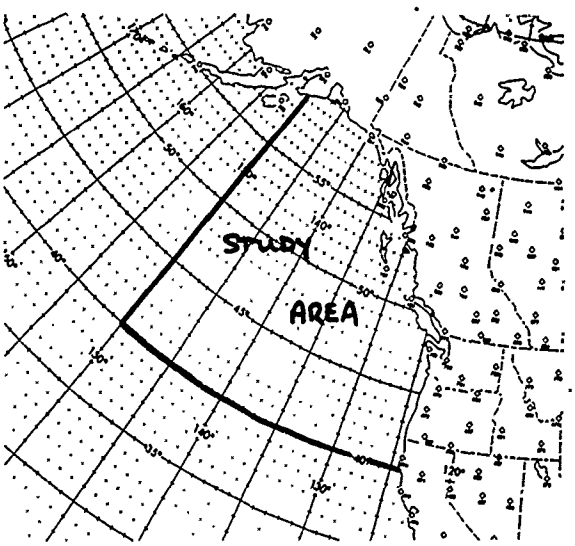


Figure 1. Location of study area over eastern Pacific Ocean. Cyclones must lie within this area at end of rapid intensification phase to be placed into data set.

For a system to be included in the data set, it had to lie within this area at the conclusion of its maximum intensification. A 24 hour time period was chosen which maximized the decline in central pressure of the cyclone.

The systems were analyzed, a posteriori, using material routinely available in the Pacific Weather Centre, namely GOES satellite imagery, data from the TIROS Operational Vertical Sounder (TOVS), 6 hourly surface weather charts, computer generated meteorological fields, and aircraft wind reports. Values were extracted for each system from answers to a series of questions that examined many facets of the intensification process. Other potential predictors suggested by MacDonald and Reiter (1988), that show promise in predicting cyclogenesis, cannot be assessed in the Pacific Weather Centre as the specialized data that is required is not available.

3. Satellite Analysis

Satellite information can be utilized in two ways when examining the potential for strong cyclogenesis. Primarily, it can indicate whether the upper air flow pattern is one that will support cyclogenesis, and secondly, it can be used to verify computer generated fields.

Weldon (1979) examined animation sequences of cyclones and was able to compile a set of 4 mid-tropospheric flow patterns that are conducive to development. By examining additional systems from the eastern Pacific, two more patterns have been added to this list. Schematic diagrams of the patterns just prior to the onset of intensification are shown in figure 2 on the following page.

These flow patterns, however, support cyclogenesis in general, not only intense cyclogenesis. To predict the latter case, an assessment of the overall dynamics of the atmosphere is required. Satellite data, both imagery and TOVS information, are utilized in this assessment to validate and, if necessary, adjust the computer objective analysis fields that are available. Using these fields and other conventional data, a subjective estimate of potential energy available for cyclogenesis can be made.

GOES satellite data also has the advantage of having a high temporal frequency, with a full disk image available every half hour. Consequently, this information is valuable in giving an up-to-date indication of the state of the atmosphere. With developing storms that can reach full intensity within an 18 hour time span, one cannot wait for the computerized forecast guidance, which appears every 12 hours, before making a forecast decision.

Aside from the standard 11.2 micrometre infra-red (IR) window channel and the visible channel, considerable use is also made of the 6.7 micrometre high-level water vapour channel. For example, strong subsidence behind an intensifying low pressure centre is an indicator that cyclogenesis has started. The speed at which this band increases in size, or darkens, is correlated with the rate of intensification. Other IR bands that have recently become available with improvements to the GOES spacecraft and the Pacific Weather Centre's satellite ground station will also be examined for clues to explosive cyclogenesis.

4. Conclusion

Satellite information plays a critical role in the assessment of potential cyclogenesis over the eastern Pacific Ocean. Infra-red, water vapour and visible band imagery can be utilized to identify atmospheric circulation patterns that support surface development. In addition, animated imagery can indicate which systems are intensifying and how quickly that process is proceeding. TOVS data can be used to validate the accuracy of numerical analysis fields and to adjust them as required.

The availability of high quality real-time satellite data allows the marine meteorologist to have more confidence in the analysis of a particular weather pattern and hence, more confidence in the accuracy of the forecast.

5. Bibliography

1. McLennan, N. and Neil, L., "Marine Bombs Project - Final Report", Pacific Weather Centre Internal Report, unpublished, 1988.

2. MacDonald, B.C. and Reiter, E., "Explosive Cyclogenesis Over the Eastern United States", Mon. Wea. Rev., 116, 1568-1586, 1988.

3. Sanders, F. and Gyakum, J., "Synoptic-Dynamic Climatology of the "Bomb"", Mon. Wea. Rev., 108,1589-1606, 1980.

4. Weldon, R., "NWS Satellite Training Course Notes, Part IV Cloud Patterns and the Upper Wind Field", Satellite Applications Laboratory, NOAA-NESDIS, unpublished, 1979.

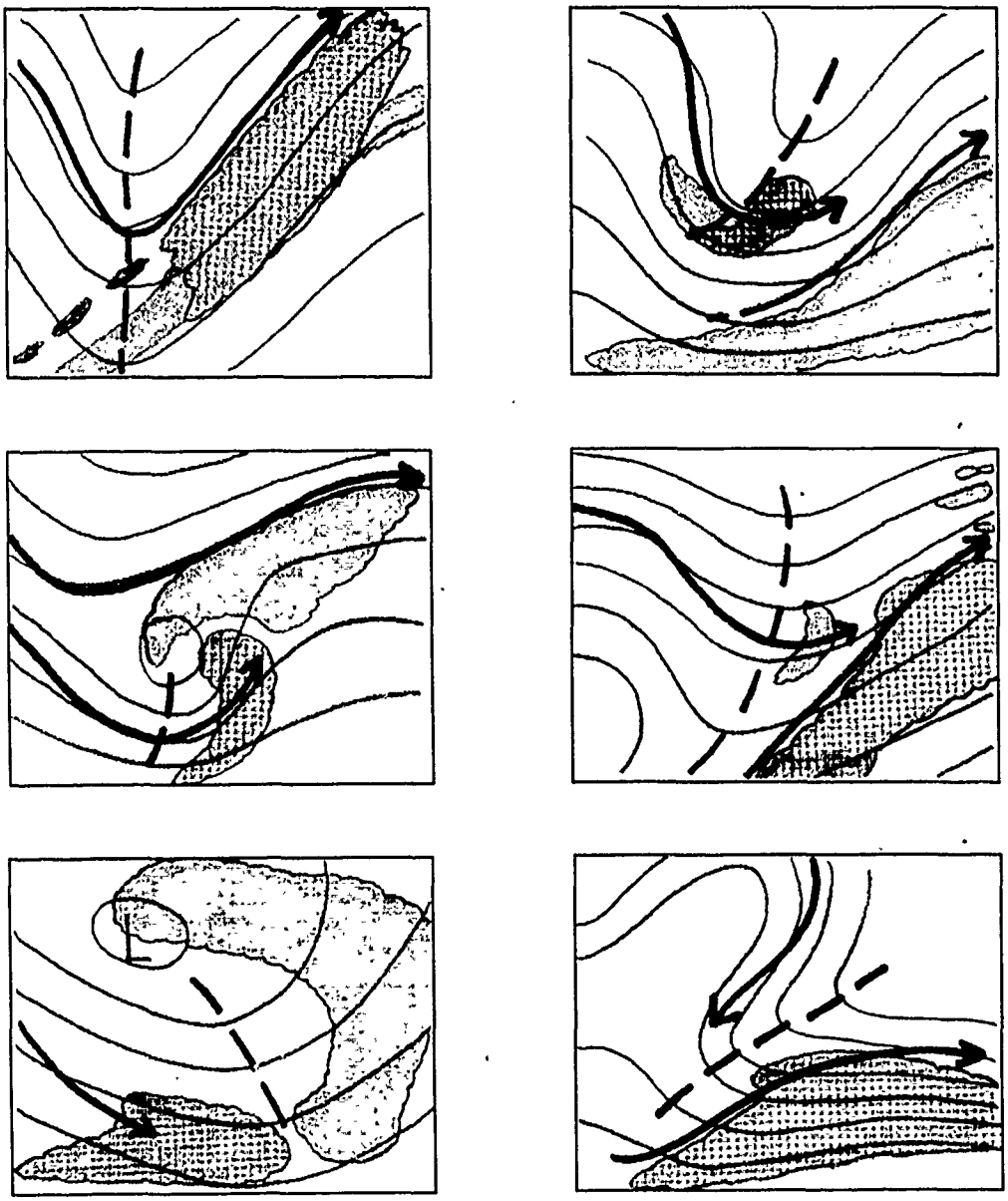


Figure 2. Six different flow patterns from satellite imagery that support cyclogenesis. Lines represent upper air streamlines with arrows marking axes of maximum winds. Cloud areas are stippled.

TÉLÉDÉTECTION DES SITES À POTENTIEL AQUICOLE DANS LA BAIE DES CHALEURS AU QUÉBEC: RÉSULTATS PRÉLIMINAIRES

Marcelle GRENIER[1], Jean-Marie M. DUBOIS[1], André LAVOIE[2] et Élizabeth LAMBERT[2]

(1) Centre d'applications et de recherches en télédétection, Université de Sherbrooke, Sherbrooke, Qc
(2) Département de géographie, Université de Montréal, Montréal, Qc

Résumé

La télédétection permet de faire un zonage de sites maricoles avec un minimum de levés au sol. Les sites identifiés peuvent, par la suite, être caractérisés en vue d'établir leur capacité de support et d'en faire le suivi.

Cette étude a pour but d'accélérer le processus d'identification et de caractérisation des sites maricoles à l'aide de capteurs aéroportés (MEIS-II, FLI) et satellitaires (Landsat-TM, SPOT-HRV, NOAA-AVHRR). Ces images sont, par la suite, combinées à des données biologiques, océanographiques, anthropiques, géomorphologiques et météorologiques en utilisant un système d'information géographique (SIG).

En démontrant le potentiel des côtes du Québec et en facilitant le choix des sites aquicoles, nous voulons aider à augmenter le nombre, la diversité et le rendement des exploitations. La méthode est basée sur un regroupement des différentes thématiques qui composent une côte et pour lesquelles il existe une expertise, soit les substrats et la géomorphologie, l'océanographie physique, les sédiments en suspension, le zoo et le phytoplancton, la végétation marine, la bathymétrie et les infrastructures.

Le traitement numérique de ces thèmes et leur intégration aux données complémentaires sur SIG permettent d'accroître la rapidité d'analyse ainsi que le nombre d'intervenants dans la sélection d'un site aquicole.

Cette étude est financée par le ministère de l'Agriculture, des Pêcheries et de l'Alimentation du Québec.

Mots clés: aquiculture, milieu biophysique, côtes, images satellitaires, images aéroportées, SIG

Introduction

Globalement, l'aquiculture n'est pas une industrie récente. La Chine la pratique depuis plusieurs milliers d'années. Au Québec, on connaît bien la culture de la truite d'eau douce qui existe depuis plusieurs années. Plus récemment, la culture du saumon en eau salée est devenue une industrie de plus en plus importante (Lafleur, 1986). On vise maintenant la production de nouvelles espèces maricoles (Poirier et *al.* 1982).

Le but de notre étude est de faire un zonage des sites maricoles dans la baie des Chaleurs. La mise au point de notre méthode vise à accélérer le processus d'identification et de caractérisation des sites maricoles en évaluant la capacité des capteurs aéroportés et satellitaires existants dans une démarche pluridisciplinaire. La méthode optimale sera celle qui opérera avec un minimum de capteurs et dont le moindre coût et la facilité de suivi seront possibles. Avec ses 13 000 km (calculés à l'échelle du 1: 250 000) de côtes, le Québec possède un potentiel important. L'expertise au sol du milieu côtier québécois par méthode traditionnelle, tant pour la capacité de support du milieu naturel que pour la recherche de sites favorables aux futures exploitations aquicoles, est freinée par les coûts et le temps d'exécution. La télédétection peut dégrossir les problèmes et cerner les zones d'expertise au sol. On peut ainsi procéder rapidement à un inventaire complet des côtes, avec l'appui d'un SIG, en vue de planifier son développement et son aménagement et ainsi permettre le suivi de son évolution (Cheney et *al.* 1984; Kapetsky et *al.* 1987).

Le site d'étude

Le site choisi représente 350 km de côtes dans la baie des Chaleurs entre la rivière Restigouche et Gaspé. Dans cette région, nous possédons des photographies aériennes de 1948 à 1980 à des échelles variant du 1: 2 500 au 1: 36 000. Nous avons aussi des survols aéroportés des capteurs MEIS-II, DAEDALUS et FLI ainsi que des images Landsat-TM, SPOT-HRV et NOAA-AVHRR (figure 1). De plus, nous possédons des données au sol prisent simultanément à l'acquisition d'images ainsi que des relevés biophysiques compilés sous forme de rapport ou de cartes qui permettrons de vérifier les résultats.

Méthodologie

La méthodologie tient compte surtout des besoins de la mytiliculture. Elle peut cependant être adaptée à d'autres types de mariculture.

Les études déjà publiées sur l'aquiculture traitent rarement ou

en faible partie l'aspect environnement du site. Nous avons donc fait la liste des paramètres biophysiques qui composent le milieu côtier et marin et qui sont nécessaires à l'évaluation de l'environnement d'un site aquicole. Ces paramètres biophysiques sont regroupés dans quatre grands thèmes:

- la géomorphologie
- la dynamique côtière
- la biologie marine
- les activités anthropiques

La méthode est basée sur la détection et l'intégration des paramètres qui composent ces thèmes dans un système d'information géographique avec un minimum de capteurs et de travaux au sol. La démarche du projet est schématisée dans l'organigramme représenté à la figure 2.

Comme nous l'avons vu, la première étape a été de détailler tous les éléments faisant partie des quatre grands thèmes avec leur niveau de précision acceptable pour le choix d'un site. Un inventaire des documents touchant de près ou de loin le traitement d'images appliqué au milieu marin, en plus des études portant sur les éléments qui composent le milieu marin, a été effectué sur support micro-informatique. Le nombre total de documents s'élève à environ 600. Nous avons analysé la documentation portant sur les pratiques aquicoles ainsi que celle sur les intrants biophysiques et les facteurs limitants des espèces maricoles.

Suite à la revue bibliographique, nous avons pu diviser, d'après l'expertise, les paramètres qui peuvent être détectés par analyse numérique des images multispectrales, ceux qui seront interprétés à partir des photographies aériennes et finalement, ceux qui seront compilés à partir de relevés au sol (déjà existants).

Une étude préliminaire de la performance de détection des paramètres biophysiques par télédétection est représentée au tableau 1. Cette analyse a été faite à partir de la documentation et, dans certains cas, à partir d'applications que nous avons nous-mêmes effectuées, toujours en tenant compte des paramètres biophysiques caractéristiques de la baie des Chaleurs. Les capteurs analysés sont ceux dont nous possédons des images.

L'interprétation comparée d'images satellites TM de Landsat et HRV de SPOT a été faite pour extraire l'information sur la végétation marine immergée et émergée, la géomorphologie, les courants côtiers et les panaches de cours d'eau. Les informations sur la température de surface proviennent des images Landsat-TM et NOAA-AVHRR. La première, de moyenne résolution spatiale (120 m X 120 m), est exploitée pour la reconnaissance des structures thermiques locales. La seconde, dont la résolution spatiale est de l'ordre de 1 km X 1 km, nous apporte une information à l'échelle régionale.

La télé-interprétation de photographies aériennes à moyenne échelle (1:15 000 ou 1: 20 000) et à petite échelle (1: 36 000) de la zone côtière a été faite pour la géomorphologie, la turbidité des eaux, des substrats, les courants côtiers et de dérive littorale, les panaches de cours d'eau et les infrastructures. Dans les zones litigieuses, il y a eu relevé au sol (étés 1987 - 1988). Par conséquent, on a utilisé les résultats de l'interprétation des photographies aériennes spécialement pour la géomorphologie.

La comparaison entre l'interprétation des photographies aériennes et celle des images satellites a été faite au niveau de la géomorphologie, des infrastructures, de la végétation marine et de la bathymétrie.

L'analyse des paramètres à partir de la documentation existante nous a permi d'établir des critères de sélection théoriques. Au niveau de l'évaluation des sites potentiels, en conjonction avec les intervenants en aquiculture, l'intégration des données anthropiques (infrastructures) et de quelques données complémentaires (biophysiques, géomorphologiques et météorologiques) sur SIG permet de commencer la sélection préliminaire des sites.

Résultats préliminaires

L'image thermique de Landsat-TM nous fournit, après seuillage de l'histogramme, une carte de la répartition des températures de surface pour le 16 juillet 1987. On observe un gradient maximal d'environ 10 degrés entre la côte du Québec et celle du Nouveau-Brunswick. Les eaux les plus froides se retrouvent sur la côte du Québec entre Carleton et Bonaventure. Bien que l'extraction d'information sur les températures de surface soit complétée, il reste à mettre en relation ces dernières avec les autres paramètres biophysiques, tels la profondeur de l'eau, les conditions météorologiques (vitesse et direction des vents) ainsi que des données existantes tirées de la documentation pour préciser l'interprétation de ces résultats. À partir de cette même carte d'indices météorologiques, nous avons déduit la direction des courants de surface. Il semble que, sous l'effet du vent, il y ait un déplacement de la masse d'eau du nord vers le sud-est pour la journée considérée. Cette hypothèse est préliminaire et une analyse additionnelle est nécessaire pour augmenter la validité des résultats. Cette image nous a également permis d'extraire le panache des rivières Bonaventure et Cascapédia. La même méthode d'analyse a été appliquée aux images AVHRR de NOAA enregistrées les 4 et 15 mai 1987. On obtient alors une carte des courants de surface pour l'ensemble de la baie pour deux autres dates. Une image du même capteur, datée du 13 avril 1987, est également disponible; l'intérêt de cette image est qu'elle nous permet de localiser les zones englacées. Il est à noter que les données sur les conditions d'englacement devront être complétées avec les cartes provenant du ministère de l'Environnement du Canada. Ces dernières sont produites sur une base régulière, ce qui nous permet de faire une évaluation plus fiable des conditions normales d'englacement.

Au niveau de la photo-interprétation, principalement pour la géomorphologie et les infrastructures, nous évaluons et tentons

présentement de regrouper les éléments qui apportent une information utile pour la sélection des sites aquicoles et ce avant leur intégration sur SIG (dérive littorale, nature du substrat).

On identifie sur l'image HRV de SPOT des courants d'eau turbide très distincts. Il est toutefois moins aisé d'identifier les peuplements de végétaux immergés, tant avec l'image HRV de SPOT que TM de Landsat. Les résolutions spectrale et spatiale des images satellites sont trop faibles pour permettre d'extraire cette information normalement accessible avec les capteurs aéroportés (MEIS-II et FLI spatial).

Les résultats obtenus montrent que les images numériques, tant aéroportées (MEIS-II et FLI) que satellitaires (SPOT-HRV, Landsat-TM et NOAA-AVHRR), combinées aux données complémentaires(biophysiques,anthropiques, géomorphologiques et météorologiques), aux mesures au sol et à la télé-interprétation de photographies aériennes, permettent une détection rapide et peu coûteuse des paramètres biophysiques.

Cependant, la précision de ces résultats diffère de celle obtenue avec les méthodes traditionnelles. La précision pour les thèmes étudiés varie selon différents facteurs. Pour la géomorphologie, la nature du substrat et la végétation emergée, la précision dépend de la résolution spatiale des images, tandis que pour les matières en suspension, la végétation immergée, les courants côtiers, la dérive littorale, la turbidité et la bathymétrie, la précision se calcule en fonction de la résolution spectrale et de la capacité du capteur à détecter à de grandes profondeurs; la limite actuelle est d'environ 10 m pour la baie des Chaleurs avec le capteur MEIS-II. Pour ce qui est des courants généraux et de la température de l'eau, la détection se limite au 1er centimètre de la surface. Néanmoins, la documentation fait état d'un bon nombre d'exemples en océanographie appliquée montrant la possibilité de relier la température de surface à certains mouvements observés en profondeur: cette relation reste à faire.

Par ailleurs, nous avons numérisé sur un SIG, pour la région comprise entre Maria et Pointe Bonaventure, le profil de la côte, les routes principales, les centres urbains et les zones urbaines, les quais et les rampes d'accès, les zones récréatives et les limites des MRC. Il reste deux éléments à numériser, soit les bancs de mollusques déjà existants et la nature du substrat. L'intégration d'une partie des données numériques aux données complémentaires numérisées a été effectuée avec succès.

Les étapes à franchir sont donc celles qui évaluent la fiabilité des outils utilisés par analyse statistique et/ou par une prise de mesure au sol pour vérification des résultats. Nous analyserons les avantages et les inconvénients de l'utilisation des méthodes de télédétection pour la sélection des sites aquicoles. Par la suite, nous pourrons faire un zonage préliminaire des sites potentiels pour la mariculture dans la baie des Chaleurs. Les résultats finaux seront représentés sur des cartes à petite échelle pour l'ensemble de la baie et à grande échelle pour les zones privilégiées. Ces cartes seront regroupées par thèmes (ex.: température, bathymétrie, etc.), par ensemble de thèmes (ex.: température et courants, profondeur et végétation marine, etc.) ou par combinaison pondérée de plusieurs thèmes, ce qui nous permettra de caractériser les sites (nul, faible, moyen, fort).

Références

1. Cheney, D.P. and Rabanal, H.R. "Remote sensing and its application to inland fisheries and aquaculture", Rome, FAO, FAO Fisheries Circular no 768, 50 p., 1984.

2. Kapetshy, J.M., McGregor, L. and Nanne E.H., "A geographical information system and satellite remote sensing to plan for aquaculture development: a FAO-UNEP/GRID cooperative study in Costa Rica", Rome FAO, FAO Fisheries Technical Paper no 287, 51 p., 1987.

3. Lafleur, P.E., "La production de salmonidés en eau salée" MAPAQ, Pêcheries, 41 p., 1986.

4. Poirier, L. et B. Myrand "Élevage de la moule bleue, Mytilus Edulis, dans les lagunes des Iles-de-la-Madeleine (Québec)", MAPAQ, Direction générale des pêches maritimes. Travaux sur les pêcheries du Québec, n° 49, 64 p., 1982.

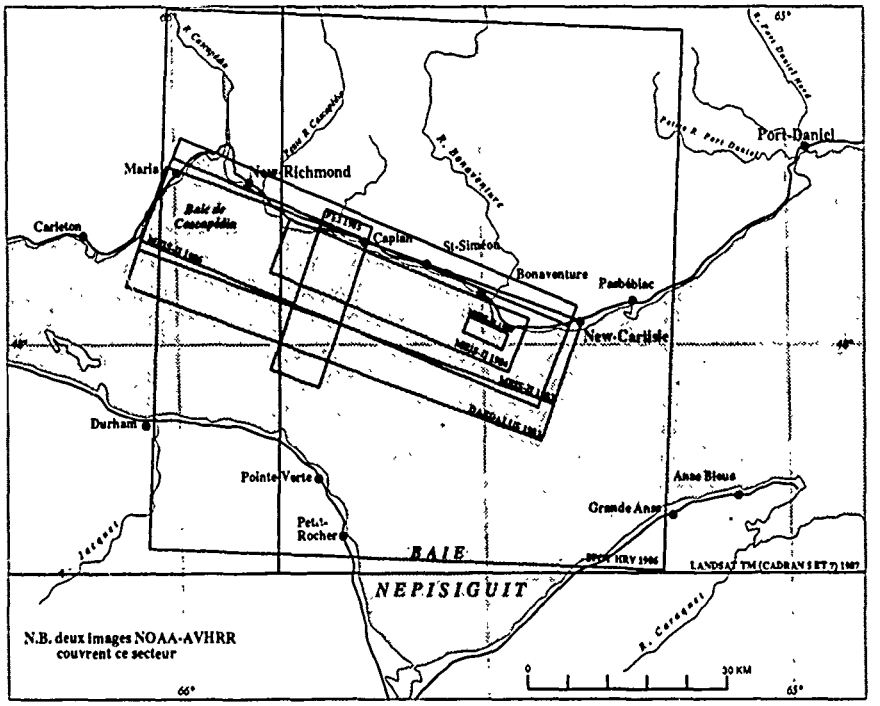


Figure 1: baie des Chaleurs: données de télédétection disponibles

Tableau 1 : Evaluation de la possibilité de détection des paramètres biophysiques pour divers capteurs en fonction de l'aquiculture

1. bon
2. moyen
3. passable
- non applicable

	paramètres biophysiques	aéroportés						satellitaires		
		photo noir et blanc	photo couleur	photo infra-rouge	balayeur multispectral DAEDALUS	balayeur électro-optique à barrette MEIS-II	capteur à fluorescence FLI	balayeur multispectral à haute résolution de Landsat-TM	capteur à barrette à haute résolution de SPOT-HRV	radiomètre à très haute résolution de Tiros-NOAA - AVHRR
Géomorphologie	trait de côte et évolution littorale	2	2	1	1	1	1	1	1	3
	géomorphologie émergée	1	1	1	2	2	-	3	3	-
	géomorphologie immergée	3	3	-	2	2	1	3	3	-
	marais	1	1	1	1	1	-	2	2	-
Dynamique côtière	température et courants	3	3	-	1	2	-	1	-	1
	matières en suspension	2	2	-	1	1	2	1	1	2
	glaces et dérive des glaces	1	1	1	1	1	-	2	2	3
	vagues	3	3	-	2	2	-	2	2	-
	panaches de rivières	3	2	2	1	1	1	1	1	3
Biologie marine	chlorophylle et plancton	3	3	-	1	1	2	1	1	2
	végétation émergée	2	2	1	1	1	1	2	2	-
	végétation immergée	3	3	-	1	1	2	3	3	-
	mammifères marins et poissons	1	1	-	2	2	2	3	3	-
Activités anthrop.	structures et aménagements	1	1	1	2	2	3	2	2	-
	déversement de polluants	3	2	2	1	1	1	1	1	3

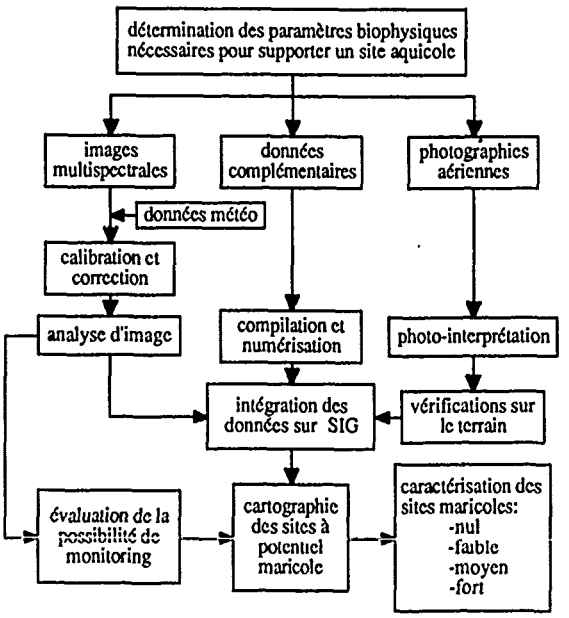


Figure 2 : Organigramme méthodologique

A precise APT receiver for cloud and sea surface temperature imaging.

J F R Gower,
Institute of Ocean Sciences, P.O. Box 6000, Sidney, BC, V8L 4B2, Canada
Tel: (604) 356-6558, Fax: (604) 356-6479, Telex: 049-7281

T.B. Lowe,
Galaxy Consulting, 630 Speed Ave., Victoria, BC, V8Z 1A4, Canada
Tel: (604) 381-2208

B. Wannamaker,
Sea Scan, 16065 Fifth Line, Albion, R.R. 3, Caledon East, Ont., L0N 1E0, Canada
Tel: (416) 880-0528

ABSTRACT

Use of the APT, reduced resolution thermal and near-infrared AVHRR data from the NOAA weather satellites has been demonstrated for a variety of purposes, including the real-time mapping of weather patterns and sea surface temperatures. Lower cost systems have been used for producing cloud images, but have not preserved the data quality needed for water temperature mapping. Higher quality commercial receiving systems have in the past cost over $50,000. We present preliminary examples of imagery collected from a simplified receiver, digitizer and display system, which uses custom electronics to input image data a to PC-based display system. The temperature sensitivity is designed to be sufficient to map coastal sea surface temperature patterns (on the order of 0.5C), though this has not yet been achieved with our system. The PC-based display software supports a wide variety of image corrections and enhancements.

APT IMAGERY

The AVHRR imagery from the NOAA series of weather satellites provides 1.1km nominal ground (sub-track) resolution, in 5 spectral channels in the HRPT (High Resolution Picture Transmission) mode. These provide cloud, land and water visible imagery, and cloud-top, land and water surface thermal imagery. Imaging of weather patterns is certainly the most important application, but the data are also used for mapping and tracking Arctic and Antarctic ice, monitoring land use, and mapping global sea surface temperatures.

Reception of the data requires an S-band receiver with a directional, tracking antenna of about 3 - 5m diameter, though recent developments in low-noise receivers are reducing the required antenna diameter. An alternative data stream of APT (Automatic Picture Transmission) imagery is also sent from the satellite at a reduced bit rate, requiring a reduced data bandwidth, and hence less antenna gain. It is designed to be received using an omni-directional (non-tracking) antenna, making for a considerable cost reduction in this item. The lower data rate and number of elements making up an image allows for processing and storage using simplified equipment.

Starting with TIROS-N in the NOAA weather satellite series, APT data was upgraded to a digitally processed subset of the AVHRR imagery. One near-infrared and one thermal infrared band is transmitted during the day, and two thermal infrared bands are transmitted at night. Every third line of the two channels is selectively averaged to maintain a nominal ground resolution of 4km. The averaging is varied (less averaging away from the scan centre) to compensate the effect of the non-linear scan geometry.

The APT system was designed to allow simple reception of cloud images. LaViolette (1975) reported on the use of the data from an APT receiver installed on board a research vessel. For oceanographic surveys that involve tracking of features with strong surface thermal signatures the data was useful both as an operational guide during the survey, and as a data source in the post-survey analysis. This application has since been exploited commercially with several installations now being available in the $40,000 to $80,000 price range.

More recently, Wannamaker (1984) reported on analysis of the present APT data, giving examples of imagery and showing a comparison of the cloud and temperature sensing capability of the APT and HRPT products. He showed the example in Figure 1 of the same line of near and thermal infrared data recorded from the APT and the HRPT channels of the AVHRR. The line crosses a sun-glint pattern at 170 to 120km along track, and a narrow cloud streak at 430km that is partially smoothed in the lower resolution APT data. Within this resolution capability, the two results agree extremely well. Comparison after calibration showed differences of 0.3 degrees in sea surface temperature, and of 0.1% in albedo.

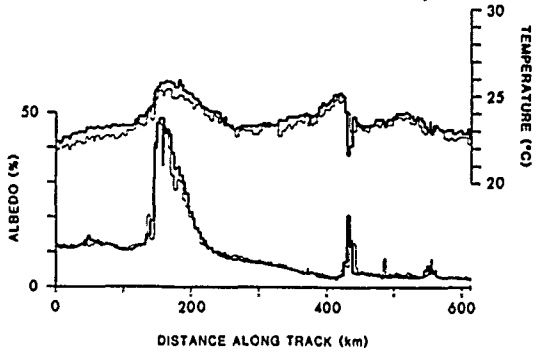


Figure 1. A comparison of a line of HRPT (heavy line) and APT (light line) data, showing good agreement between results from the two data streams in both near (albedo) and thermal (temperature) infrared bands.

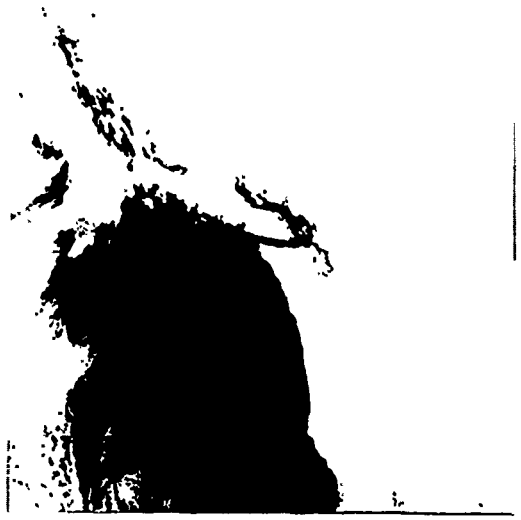


Figure 2. Thermal APT image collected on February 6 1989, showing a 2000km square area of the west coast of Canada and the USA.

APT DATA RECEPTION AND PROCESSING ON A PERSONAL COMPUTER

The APT signal is presently received through a circularly polarized spiral antenna using a simple commercial preamplifier and receiver (Hamtronics GaAsFET preamplifier model LMG144, AOR communications receiver model AR2001). This type of receiver provides a 100KHz FM bandwidth at the receiving frequency of 137MHz. This bandwidth is wider than the 30KHz optimum required to cover the APT signal FM deviation of ±15KHz, and increases the sensitivity of the system to interfering signals which are a common problem at these (VHF) frequencies. The narrower bandwidth setting of this receiver (20KHz) is insufficient to cover the FM deviation.

The receiver has the advantage of digital local oscillator synthesis and a scanning capability which allows it to automatically search among the several frequencies currently in use by US and USSR satellites. Only US satellite data have been used so far.

The output from the receiver is an amplitude modulated 2.4KHz sub-carrier, which is demodulated and digitized by a video decoder produced by Sea Scan in Toronto, Canada. This digitizes the signal to 10-bit resolution with a sampling rate of nominally 4.8KHz. The output is made available through an IEEE-488 standard interface. Doppler shifts due to the relative motion of the satellite are eliminated from the data stream. The system is designed to continue data output at a constant rate during interfering noise bursts.

Image data from the decoder are stored and processed on a standard PC/AT computer equipped with a 70Mbyte hard disk and VGA colour and black and white display. Each image scan-line is 1200 pixels long by 10 bits deep in two bands. A typical satellite pass gives about 750 scan-lines of data, making a total of 3.7 Mbytes per image.

Custom software, written in the compilable "C" language, has been written to reduce the data volume by extracting smaller sub-areas of 512 by 512 pixels. The software also reduces the 10 bits of data to 8 using a piece-wise linear compression which maintains full precision over a limited subset of the data range. A variety of image brightness enhancements are provided in black and white and colour.

Calibration of the data is possible using analog calibration step wedges provided in both thermal and near infrared bands. Capability for this is currently being added to the software.

EXAMPLES OF IMAGE DATA

Figures 2 and 3 show examples of 512 by 512 pixel images collected with the present system and extracted fro[illegible]e full APT images. Figure 2 shows a cloud-free area of [illegible]c Ocean off Vancouver Island on Februa[illegible] 6 198[illegible] covered by the image is about 2000km [illegible] along-scan averaging applied to the A[illegible] da[illegible] re[illegible] distortion characteristic of uncorrected AVHRR

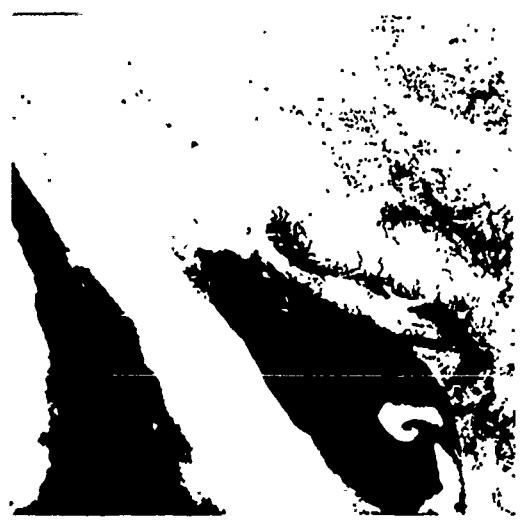


Figure 3a. Near infrared band APT image collected on April 10 1989.

Figure 3b. Thermal infrared band APT image collected on April 10 1989, at the same time as Figure 3a.

Figure 3 shows images in both near and thermal infrared taken on April 10 1989. Faint periodic patterns visible in both bands are due to interfering signals which were not present on the February image. The thermal contrast between Juan de Fuca Strait and the open ocean is just visible in this image. This is expected to be 3-7C at this time of year, indicating that with the present installation, only large thermal contrast features are being detected.

On all these images the rms digital signal variation in uniform areas is about 20rms in the full 0 to 1023 range of the digitized signal. The interfering signal causes additional oscillations also of 20 digital units peak to peak range. Since the temperature sensitivity of AVHRR data is about 0.2C per digital unit, this noise level is consistent with the apparent sensitivity noted above.

APT data is capable of considerably greater temperature sensitivity. We are presently working on improving our system to achieve the expected level of 0.5C.

CONCLUSIONS

Although the present installation is still in the development stage, the APT imagery from the NOAA satellite series has proved useful for monitoring clear sky conditions for later ordering of more sophisticated satellite image data, and for demonstration of satellite imaging techniques. The system is certainly compact and simple enough for ship-board use.

REFERENCES

Wannamaker, B., 1984, "An evaluation of digitized APT data from the TIROS-N/NOAA A-J series of meteorological satellites," Int. J. Remote Sensing, 5, 133-144.

LaViolette, P.E., L. Stuart and C. Vermillion, 1975 "Use of APT satellite infrared data in oceanographic survey operations" EOS, Trans. Am. Geophys. Union, 56, 276-282.

TOPOGRAPHIC MAPPING FROM GPS-SUPPLEMENTED AIRBORNE RADAR

by

J. BRYAN MERCER AND PETER BUTTON
INTERA TECHNOLOGIES LTD.
#2500, 101-6th Avenue S.W.
Calgary, Alberta, Canada T2P 3P4
Telephone: (403) 266-0900
Fax No: (403) 265-0499

M. MILLOT, M. KARSPECK AND F. LEBERL
VEXCEL CORPORATION
2905 Wilderness Place
Boulder, Colorado, U.S.A. 80301
Telephone: (303) 444-0094
Fax No: (303) 444-0470

ABSTRACT

In parts of the world which are often cloud or haze covered, airborne radar offers an attractive alternative to conventional aerial photography or satellite solutions to the problem of creating topographic mapping products at scales up to 1:50,000. The capability to derive Digital Terrain Models (DTMs) and rectified images using digital data from STAR-1, an airborne SAR, has been previously demonstrated. In that work, however, a relatively large number of ground control points (GCPs) were incorporated into the process in order to counter the effects created by aircraft motion with respect to the mean flightpath.

In regions where this new technique is most needed, GCP's usually do not exist, and are very expensive to create. The solution to this dilemma has been to incorporate a more precise navigation system into the acquisition and subsequent processing phase. In this paper, we report on the general methods and results of this approach.

A test was performed in the Brazeau area of Alberta, using INTERA's STAR-2 airborne SAR. A Global Positioning System (GPS) was used in differential mode to provide three-dimensional positioning of the aircraft at 1.2 second intervals. GPS data were logged onto a lap-top computer, with timing markers used to approximately synchronize the GPS and radar data streams. Algorithms were developed and coded to enable the radar image pixels to be geo-referenced, based upon the GPS derived aircraft location, the radar geometry, and the radar processor response characteristics.

Internal consistency checks of the stereo pairs gave RMS errors in the horizontal of less than 20 meters (RMS). On the other hand, vertical accuracy was about 60 meters (RMS), which is double the error achieved in earlier STAR-1 tests. This was due to sub-optimal geometry for the test configuration, and limitations in the GPS accuracy. It is noted that these limiting factors will not preclude considerably better vertical results being achieved in future work. It is concluded that, with certain qualifications, GPS supplemented airborne SAR, flown in stereo mode is a viable method for obtaining topographic map products at scales of 1:50,000 and smaller.

COHERENT SAR SIMULATION

V.H. Kaupp, N.D. Blackford,
W.P. Waite, and H.C. MacDonald

University of Arkansas, Fayetteville

ABSTRACT

A coherent SAR (synthetic aperture radar) simulation model using DTMs (digital terrain models) has been developed and implemented on a digital computer. In this approach, a set of discrete scattering centers are draped over a DTM. The DTM provides both the topography and access to a scattering model. The scattering centers provide the phase coherence and lead to a speckle model. They also provide a mechanism for exploring different scattering models.

The inputs to the SAR simulation computer program are 1) a DTM, 2) a scattering model, and 3) a sensor model. The program provides a coherent SAR signal data stream as it's output.

The simulation approach treats the incorporation of various scattering models, from purely specular to purely diffuse. The work reported illustrates that such a simulation approach is valuable for understanding and insight into SAR terrain imaging.

AUTOMATED STEREO MATCHING USING INTELLIGENT INTERPOLATION IN DIFFICULT TO MATCH REGIONS

By

James P. Strong[1]
H. K. Ramapriyan [1]
and
William P. Haynes[2]

The Shuttle Imaging Radar-B (SIR-B) experiment flown in October, 1984 provided the first opportunity to examine synthetic aperture radar (SAR) images of a particular area taken at multiple incidence angles. It also provided the first opportunity to use such images as stereo pairs from which elevation information could be derived. Analysis of various stereo pairs from this experiment have shown that based on the accuracy of the knowledge of the orbit parameters, elevations can be determined to within 80 meters assuming a perfect match of corresponding pixels in the pair.

To make use of such imagery on an operational basis, however, an automated technique for the matching of corresponding pixels (stereo matching) must be developed. One of the most promising techniques for stereo matching is the Hierarchical Warp Stereo algorithm which starts with low resolution versions of the images and performs the matching at increasing resolutions. At each step, one image is "warped" corresponding to the disparity measured so far and used as input for the next step. This algorithm is computationally expensive but well matched to massively parallel architectures. A similar algorithm has been implemented on the Massively Parallel Processor (MPP) at the Goddard Space Flight Center which performs this task on every pixel in a 512 x 512 image in less than one minute. The major problem with applying this technique to SAR stereo pairs is the fact that they are typically noisy, often have high relative distortion between one another, and have areas of low contrast. All of these effects lead to a low probability of obtaining an accurate match.

This paper describes the hierarchical stereo matching algorithm implemented on the MPP and an interpolation process which "intelligently" interpolates the disparity in areas where the match is likely to be inaccurate. The stereo matching algorithm outputs a "match score" function as well as the disparity function at every pixel in the image. The match score function is based on the input image contrast and the shape of a correlation function computed during the matching process. (A high peak indicates a good match. Multiple high peaks or low valued peaks indicate a poor match.) The match score function is input along with the disparity function to a multi-grid relaxation interpolation algorithm. This algorithm produces a new disparity function which progresses from the original disparity values in regions where the match score is high to interpolated values (with constrained change in the local gradient) in areas where the match score is low. The match score acts as a weighting function to control the degree of interpolation. Thus this algorithm produces the disparity function over the entire image with high confidence at regions where matching is easy and an "intelligent" approximation of the disparity where the matching is difficult.

[1]Code 636, Goddard Space Flight Center, Greenbelt, MD 20771
[2]ST Systems Corporation, 4400 Forbes Blvd., Lanham, MD 20706

EFFECTS OF SPECKLE AND ADDITIVE NOISE ON THE ALTIMETRIC RESOLUTION OF INTERFEROMETRIC SAR (ISAR) SURVEYS

C. Prati*, F. Rocca and A. Monti Guarnieri

* CSTS-CNR - Dip. di Elettronica - Politecnico di Milano
Dip. di Elettronica - Politecnico di Milano - Piazza L. da Vinci 32 - 20133 Milano - Italy
Fax n.: (+39)2 – 23993587

Keywords: SAR, interferometry, digital elevation map, motion detection.

Abstract

In Synthetic Aperture Radar interferometry, the altimetric information of the terrain can be obtained from the phase difference of two focused complex images. These images can be gathered by the same sensor in two passes along different orbits or by two sensors mounted on the same platform. The altimetric resolution of such a system improves when the sensors displacement is increased in the cross-track direction. The maximum allowed displacement, limited by speckle noise, increases with the spatial resolution of the SAR image. Excluding the additive noise, we show that the achievable vertical resolution is better than the slant range resolution by a factor of $\cos\alpha/2\sqrt{2\pi}$, where α is the off-nadir angle. As an example, an altimetric map of the Panamint Valley area is calculated using repeated passes of the SEASAT satellite. The effect of the additive noise is visible only when the cross- track distance of the two orbits is low. If airborne ISAR images are considered, the obtained terrain elevation map are strongly affected by altitude and attitude variations of the platform. Thus, if a priori information on the terrain elevation is available, interference fringes can be exploited to estimate the platform motion. Interferometric SAR data collected by the NASA DC-8 Airplane (courtesy of JPL) have been used to get experimental results.

1 Introduction

Single SAR images do not carry information on the terrain elevation [1]. SAR interferometry, on the contrary, allows to derive altimetric information exploiting the phase of focused images gathered in multiple satellite passes or in a single pass of a platform with two antennas (i.e. airborne interferometry) [2,3,4]. The construction of an altitude map with this technique can be carried out in three steps. First, accurate focusing and registration of the complex SAR images is performed [6]. An incorrect alignment as well as an image defocusing have a great influence on the quality of the final result. The second step consists of the evaluation of the phases of each picture elements. A map of the phase differences between the same complex pixel in the two images (interference fringes), which are related to the picture elements elevation, is constructed. Its values, however, are limited to the range $-\pi$ and $+\pi$ and they must be unwrapped. The last step has a great influence on the choice of the system parameters. In fact the parallax angle between the two passes must be small enough so that the interference fringes frequency does not exceed half the range sampling frequency. Within the above said limit, however, the greater is the parallax angle the greater the altimetric resolution. This depends both on the presence of speckle and additive noise. The first one is almost independent from the sensors baseline. The additive noise, on the contrary, has more influence on the closer sensors combination. Experimental results are in good agreement with the theory. These have been carried out by using both SAR raw data of three orbits of the SEASAT satellite over the Panamint Valley area and JPL DC-8 airborne data recorded over Ventura -Los Angeles- (courtesy of JPL).

2 Interference fringes generation

Interferometric phase images can be generated in three steps: data focusing, images registration and phase unwrapping.

2.1 Raw data focusing

SAR data have been focused through wave migration techniques [5,6]. Main advantages of the 'frequency-wave number' algorithm with respect to conventional ones are:

i - high precision: the 'f-k' algorithm is substantially exact.

ii - High computational efficiency and implementation simplicity. the operations to perform are a 2-D FFT, two complex multiplies per sample and a 2-D inverse FFT.

iii - No parameters updating with range is needed.

2.2 Images registration

Interference fringes result from the multiplication of pixels of the first focused image times the complex conjugate of the correspondent ones in the second image. Thus, focused complex images have to be accurately aligned to maximize the signal to noise ratio of the interference fringes [7]. To do that, the two complex images are oversampled by an FFT interpolation (8 : 1). Then, 5 areas 100 by 100 pixels wide (located on the four corners of the image and one near the center) are used to identify the relative shift between the two images (in range and azimuth), in order to maximize the contrast of the interference fringes of these areas. Finally, the proper relative shift of each pixel is computed using a bi-linear interpolation of the shift values obtained in the previous step

2.3 Phase unwrapping

Interference fringes images represent the principal values (in the range $-\pi$, π) of the phase of the ISAR image. Thus, to recover the terrain elevation, these phase images must be unwrapped. To do this, we must assume that the variation of the terrain elevation between adjacent picture elements, added to the unavoidable noise, is not so high to produce phase variations greater than π. The phase unwrapping,

then, can be easily carried out by integrating, along an arbitrary path, the phase differences between adjacent pixels. Whenever the phase difference is lower than $-\pi$ or greater than π then 2π is added or subtracted. With real data, the phase difference between adjacent points can be greater than π because of two distinct effects: loss of coherence or phase aliasing. Thus, a more sophisticated technique must be adopted. First, the areas that do not satisfy the above said condition are identified. Then, the standard procedure is applied along an integration path that does not cross these areas [8].

3 Effects of the finite range resolution

The finite slant range resolution, L_r, (approx. $7m$ for the SEASAT SAR) is a fundamental parameter in ISAR images In fact, the two measured fields would be coherent everywhere (isotropic emission from the scatterer), only if the elementary scatterers were really "point scatterers". In practice this is not the case, since the elementary cell is characterized by a finite extension dependent on the range resolution. This cell acts as an antenna with a finite width [7] and therefore coherence between the two images is obtained only if the two satellites are within the beam of the antenna corresponding to the cell, that is, within an angle β proportional to the slant range resolution L_r. Further, the ground resolution, L, or distance between the closest distinguishable points, also depends upon the terrain slope, θ, and the off-nadir angle, α, as:

$$L = \frac{L_r}{\sin(\alpha - \theta)} \tag{1}$$

The ground resolution L can be then regarded as providing the width of an equivalent antenna with an aperture angle $\beta = \lambda/L$ (λ is the transmitted wave length). This antenna generates a "celestial footprint" [7] at the satellite's distance, h, whose linear aperture is βh. In the case of SEASAT SAR images, for example, the angular aperture of the "celestial footprint" for a flat terrain is about 75 degrees and its linear aperture about $10Km$. For the flat terrain in this example, the normal displacement of the two satellites n_s must be smaller than $5Km$. In the case of an earth surface dipping towards the satellite, the ground antenna length increases, and the celestial footprint shrinks. The opposite happens when the surface is dipping in the opposite direction.

4 Limits to the altitude resolution

The determination of the altitude resolution depends both on the presence of speckle [1] and additive noise. The first one is almost independent from the sensors baseline. The additive noise, on the contrary has more influence on the closer sensors combination.

4.1 Effect of the speckle noise

The reflectivity of a pixel of the focused complex images measured by the two satellites at distance h from the target can be represented as: $\vec{A}(\alpha+\frac{\gamma}{2})$, respectively. In these expressions the nominal off-nadir angle has been indicated with α and its variation with the distance of the orbits, n_s, as $\gamma = n_s/h$. The phase variations, $\Delta\psi$ of the ISAR image with respect to the scatterer displacement normal to the slant range direction n_p can be expressed as a function of the angle γ and of the wavelength λ: $\Delta\psi = 4\pi\gamma n_p/\lambda$.

Thus, neglecting the random additive noise , the two complex values of the same focused pixel in the two images are:

$$\vec{L_1} = \vec{A}\left(\alpha - \frac{\gamma}{2}\right) e^{-j2\pi\gamma n_p/\lambda} \tag{2}$$

and

$$\vec{L_2} = \vec{A}\left(\alpha + \frac{\gamma}{2}\right) e^{j2\pi\gamma n_p/\lambda} \tag{3}$$

Our goal is the measurement of n_p. Assuming that the reflectivity is uniformly distributed within the ground resolution cell, the expression of the Root Mean Square value of the error on the measure of n_p holds if the approximation is made that $\gamma \ll \lambda/(2\pi n_p)$.

$$\sqrt{\overline{n_p^2}} = \left[\frac{L\cos\alpha}{2\sqrt{2\pi}}\right] \tag{4}$$

To get this expression, consistently with the previous assumption, we have estimated the ratio between the reflectivity and its derivative with respect to α as

$$\overline{\left|\frac{dA}{d\alpha}\right|^2} = \overline{|A|^2}\left(\pi\frac{L}{\lambda}\cos\alpha\right)^2 \tag{5}$$

Finally, the following expression of the altitude noise due to speckle holds:

$$\sqrt{\overline{n_h^2}} = \sqrt{\overline{n_p^2}}\sin\alpha = \left[\frac{L\sin\alpha\cos\alpha}{2\sqrt{2\pi}}\right] \simeq \frac{L_r\cos\alpha}{5.01} \tag{6}$$

Note that the value of $\sqrt{\overline{n_h^2}}$ depends only on the slant range resolution and on the cosine of the off-nadir angle: the greater the slant range resolution and the off-nadir angle the greater the altitude resolution. For the SEASAT the altitude resolution limit is about 1.2 meters. This could vastly improve if the reflectivity is concentrated to a single point of the cell (corner reflector).

4.2 Effect of the additive noise

The second noise term is due to additive noises on the measurements of $\vec{L_1}$ and $\vec{L_2}$. Then, by defining the "celestial footprint" of the ground resolution cell as:

$$F = h_s\frac{\lambda}{L\cos\alpha} \tag{7}$$

it can be shown that the following expression of the RMS value of altitude noise due to additive noise holds:

$$\sqrt{\overline{n_{h2}^2}} = \sqrt{\rho}\frac{F\ L\ \sin\alpha\cos\alpha}{\sqrt{2\pi}\ n_s}\sqrt{\frac{\pi}{2}} \tag{8}$$

where ρ is the noise to signal power ratio This term depends on the ratio between the celestial footprint and the sensors' baseline n_s.

Taking advantage of equations 8 and 6, we find that the ratio between the effects of the additive and speckle noise is:

$$\sqrt{\frac{\overline{n_{h2}^2}}{\overline{n_h^2}}} = \sqrt{2\rho}\,\frac{F}{n_s} \tag{9}$$

For the SEASAT cases examined in this paper, it can be noted that the second noise term equals the speckle one if the signal to noise ratio is 22.7dB and 43dB respectively for the far and close orbits.

4.3 Experimental results

The effect of the additive noise on the altitude resolution is clearly visible in the following experimental results. Three repeated passes of the SEASAT satellite over the Panamint Valley area have been exploited to generate two ISAR images whose baselines where respectively 99 and 1030 meters. A small area (3 × 6 Km) of Panamint Valley has been taken as a test area. The altitude elevation map of that area has been constructed by exploiting the interference fringes obtained both from the closer and the far orbits. The results are shown in figures 1 and 2 respectively. A comparison between figure 1 and 2 shows that the first map is much more affected by the additive noise than the second one. The altitude resolution can be estimated by comparing the two images (no Digital Terrain Map of that area was available to us). A low-pass filtered replica of the higher altitude resolution image, 1, served as the estimate of the true altitude map. Then, the difference between this map and the low altitude resolution

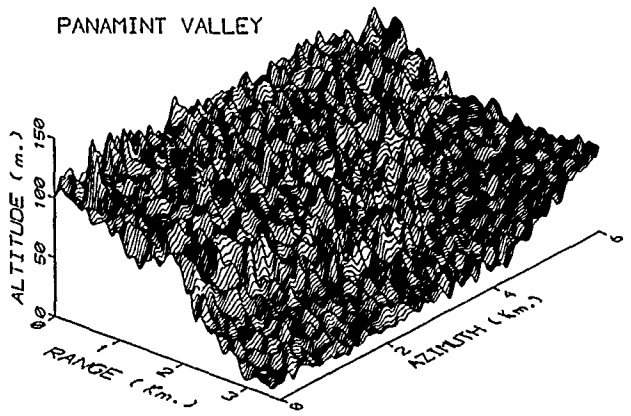


Figure 1: Tridimensional plot of the altitude map derived from the closer orbits combination.

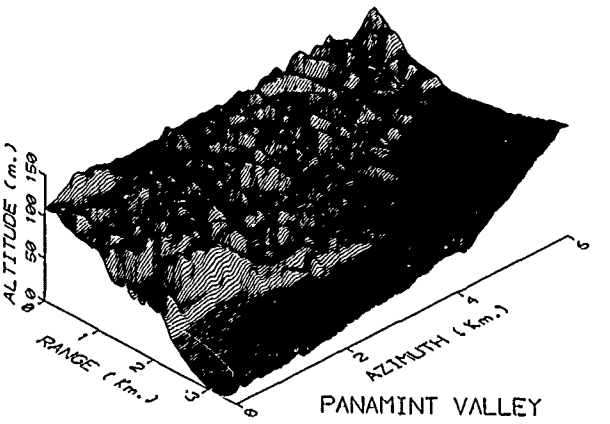


Figure 2: Tridimensional plot of the altitude map derived from the far orbits combination.

one (figure 2), represents the estimate of the noise. For the altitude map of figure 1 the estimated RMS value is about 32 meters and for that of figure 2 it is about 3.1 meters. Exploiting equation 9 it is easy to verify that these values imply a Signal to Noise Ratio (SNR) $1/\rho$ of about 14dB. Thus, from equation 9, we can compute the baseline for which the RMS value of the additive noise equals that of the speckle. Its length is about 2820 meters. This value can be regarded as the largest baseline useful to increase the altitude resolution. In fact, over this value, we would see a greater quantity of speckle noise, while the additive noise contribution would become more and more negligible.

5 NASA DC-8 ISAR images

ISAR data collected by the NASA DC-8 Airplane have been used to derive experimental results (courtesy of JPL). Raw Airplane SAR data have been focused and, after a proper images registration, interference phase images have been obtained (see figure 3). In this case, however, the obtained terrain elevation map was clearly in contrast with the real situation due to the altitude and attitude variation of the airplane. Since the observed area was almost flat, interference fringes images have been exploited to estimate the platform motion. Thus a model of the sensing system has been identified, and four parameters that allowed the best fit with the experimental data have been estimated range line by range line. The set of the estimated parameters consists of two angles which characterize the airplane attitude (pitch, roll), the airplane altitude and the antennas off-nadir angle. The obtained results are in good agreement with the on-board measured parameters. However, the noise superimposed to these Interferometric SAR data was too high to allow results better than conventional techniques using one range line at a time. Thus, a simple low-pass filter (with a cut-off frequency of 1/3 of Nyquist) has been applied to the unwrapped phases in the azimuth direction. Figure 4 shows the unwrapped phases at a fixed azimuth superimposed to the best fitting theoretical curve. The RMS value of the differences is about .1 radians. Better results could be got using a dynamic model for the airplane and then a Kalman filter. Nonetheless, it has been estimated that if a larger antennas baseline was adopted (the actual baseline was of only 50cm), and corner reflectors were exploited, the measured accuracy would have a great improvement. As an example, the airplane altitude could be measured with a precision better than 1 meter (over 8500) instead of 30 as it is now. On the other hand, we have already measured the roll angle with a precision better than .05 degrees. As an example figure 5 shows the measured roll angle compared with data recorded on the CCT header (note the delay between the two curves).

Figure 3. Interference fringes of the area of Ventura (Los Angeles) as seen from the NASA DC-8 (C Band). Sampling: 6 by 6 meters approx. Dimensions: 3 by 3 Km approx.

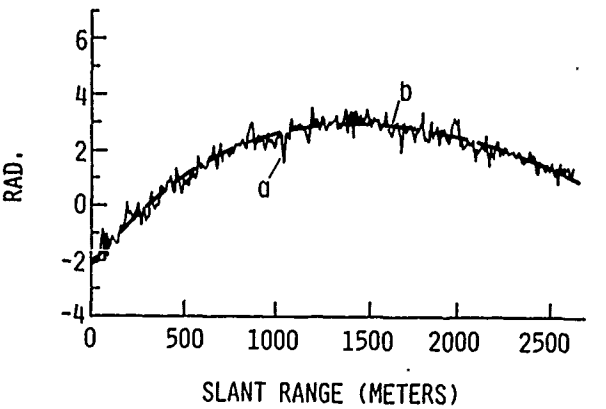


Figure 4. a - Unwrapped phases at a constant azimuth after 18m averaging. b - Best fitting theoretical curve.

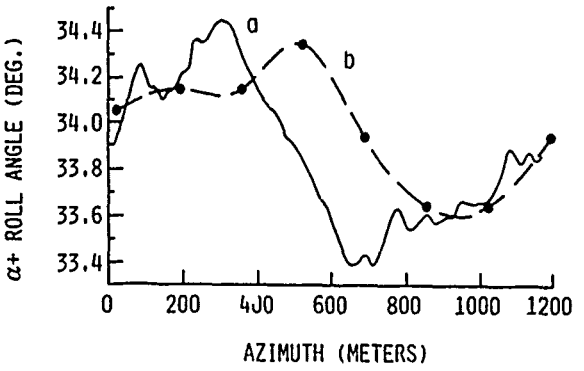


Figure 5: a - ISAR measured roll angle versus azimuth. Each point of this plot derives from the parameters set which produces the best fitting curve of figure 4. b Roll angles recorded on the CCT header

6 SAR geodesy using corner reflectors? A discussion

The previous analyses show that, if we had corner reflectors on the earth's surface, SAR interferometry would vastly improve for two reasons:

a) the aperture angle of the antenna corresponding to the corner reflector can be much wider than that of the onboard transmitter/receiver, so that we could guarantee that the image of the reflector on the surface is phase coherent, whenever in sight of the platform. In the processing phase, it would be possible to correct for the phase shifts induced by different incidence angles of the radiation. Then, it would be possible to achieve complete stigmaticity, i.e. to simulate a pointwise reflector, notwithstanding the finite reflectivity of the terrain surrounding the reflector and the phase rotations due to the antenna.

b) the signal to noise ratio for the reflection would be much higher, and therefore the dispersion of the measured phases would be smaller. This dispersion would depend on the dimension of the corner reflectors and could decrease further in the case of active transponders. In that case, however, the improvements in the SNR could be more than offset by the incertitude of the delay time. Suppose now that we had a network of corner reflectors whose relative positions had been measured once by means of laser ranging, GPS, etc. with such a precision that aliases due to wavelength errors were avoided. This network could be used for two (three) purposes:

1) to measure accurately the motion parameters. The network would act as a very large antenna;

2) to measure the relative motion of the reflectors in the radial direction;

(3) to try to measure the relative motion in the tangential direction by locating the positions of the maxima of the return and not their phases only.

At each pass, the platform would monitor the variations of the phases of returns from the corner reflectors. After proper phase unwrapping, regressions could be made to estimate the motion parameters, to compensate for weather conditions and drifts of the local oscillator, and then to evaluate global motion parameters like motion of landslides, tectonic plates, etc. The density of the corner reflectors could increase in the areas of interest, say across a fault, or to monitor local earth's motions. One satellite like ERS-1 could monitor the variations of these distances with time at each pass, say once every 19 days. It is important to remark again that using corner reflectors there is no need for orbit coherence, even if it would be a good thing to have.

7 Conclusions

In this paper the limits to the resolution of elevation maps from SAR interferometry have been discussed both from theoretical and experimental points of view. It is reported that the altitude resolution depends both on speckle and additive noise. Speckle noise is, in a first approximation, almost independent from the orbits distance: it depends only on the slant range resolution. Additive noise, on the contrary, has more influence when the orbits are closer. By taking advantage of these theoretical results, the altitude resolution of the elevation maps is estimated. The results of this work could be useful to define the system parameters of future interferometric SAR missions. The use of corner reflectors could vastly improve the precision of the results. Another possible use of ISAR that has been discussed is the estimation of the motion of the platform.

8 Acknowledgments

The authors would like to thank Dr. J. Curlander and Dr. R. Goldstein of JPL for the SEASAT and airborne SAR data. They wish to express their appreciation to E. Damonti and E. Tote' for their invaluable support [9]. This work has been carried out with the sponsoring of the Italian National Research Council and grants from the Ministero della Pubblica Istruzione 40% and PSN.

References

[1] Curran, P.J, 1985, Principles of Remote Sensing, Longman, London.

[2] Li F., Goldstein R.M., 1987, Studies of multibaseline spaceborne interferometric Synthetic Aperture Radar, Proc. of IGARSS'87, Ann Arbor, 1545-1549.

[3] Zebker, H.A., Goldstein, R.M., 1985, Topographic mapping from interferometric Synthetic Aperture Radar observations, Proc. of the IGARSS'85, Amherst, Massachusetts.

[4] Gabriel A., Goldstein R., 1988, Crossed orbits interferometry: theory and experimental results from SIR-B, Int. Jou. on Remote sensing, 9, No.5, 857-872.

[5] Cafforio C., Prati C., Rocca F., 1988, Full resolution focusing of SEASAT SAR images in the frequency-wave number domain, Proc. of the 8th EARSeL Symposium, Capri(NA), pp.336-355.

[6] Cafforio C., Prati C., Rocca, F., Synthetic aperture radar: a new application for wave equation techniques, Geophysical Prospecting, to be published.

[7] Prati C., Rocca F., 1988, 3-D Synthetic Aperture Radar Surveys, Stanford Exploration Project, SEP-57, 463-477.

[8] Prati C., 1988, Bidimensional phase unwrapping applied to SAR interferometry, internal report, Dip. di Elettronica, Politecnico di Milano.

[9] Damonti E., Tote' E., 1989, Identificazione del moto di una piattaforma da dati SAR interferometrici, Tesi di laurea, Dipartimento di Elettronica, Politecnico di Milano.

GENERATION OF GEOCODED SPACEBORNE SAR IMAGE PRODUCTS

E. Meier, Ch. Graf and D. Nüesch
Remote Sensing Laboratories
University of Zurich
Winterthurerstr. 190
CH - 8057 Zürich, Switzerland

ABSTRACT

This paper discusses a procedure for high precision geocoding of spaceborne SAR images. The algorithm described is based on a range-doppler approach, considering each individual image pixel location and its height dependent doppler frequency shift, user selected map projections and their particular geodetic datum. The flexibility of the procedure easily permits the generation of layover and shadow masks and the computation of auxiliary geometric data such as local resolution or local incidence angle in order to support the user in the complex procedure of SAR image analysis.

With the achieved accuracy between 20 m and 30 m for Seasat and SIR-B image data the possibility to generate multiframe mosaics is given, no geometric boundary discrepancies do exist. Therefore, accurate radar image maps of extended areas can be produced.

Keywords: SAR-Geocoding, Range-Doppler Approach, Map Projections, Geodetic Datums, Value Added Products, Multiframe Mosaics

1. INTRODUCTION

To improve the usefulness of Synthetic Aperture Radar (SAR) images their inherent geometric distortions must be eliminated in order to achieve a radar ortho-image that corresponds to a well defined map projection as it is used by the national surveys. It is quite obvious that a successful removal of these distortions requires an approach that interfaces directly with the image processor and that treats sensor and processor as a system, i.e. a range-doppler approach (Curlander, 1984). Kwok et al. (1987) developed a rectification procedure based on that approach for computing pixel positions on the geoid. Afterwards the height dependent location errors in range direction have been eliminated by means of DEMs. Accordingly, relief inherent azimuth shifts are not considered. Intending to increase the throughput of geocoded images several simplifications have been made (Curlander 1986). Unfortunately, a better throughput decreases the accuracy.

The rectification procedure described in this paper is based on the range-doppler approach. Because the shape of the earth (i.e. relief) can not be described precisely enough with ellipsoid or geoid equations, DEMs exclusively have been used to describe the earth's surface. The range-doppler approach was applied to each DEM cell in order to be able to consider the relief's influence on both pulse transit time as well as individual doppler frequency shift. A pixel-wise processing consumes a considerable amount of computing time but it guarantees a constant and high accuracy over the entire image.

2. GEOCODING WITH TERRAIN CORRECTION

A precise geocoding of SAR images requires an approach that reconstructs the sensor-processor system. For such rectification procedures the so called range-doppler approach is used. Because the relief has a great influence upon the doppler this range-doppler approach has to be applied to each backscatter element. This chapter shall demonstrate the requirement of a rigorous doppler approach. Further, the different coordinate systems that have to be taken into account when geocoding and the rectification procedure itself will be outlined.

2.1 Rigorous Doppler Approach

Variations in topography primarily cause distortions in range direction such foreshortening and layover. But they also affect the radial velocity $v_\perp$ between sensor and backscatter element and therefore have an influence upon the geometry in azimuth direction.

Fig. 1 shows an extreme foreshortening situation. In spite of a great horizontal and vertical distance between the backscatter elements P_1 and P_2 on earth in the slant range image the points of P_1 and P_2 are very close to each other because their range is fairly identical. They do lie in the same range segment and are therefore processed with the same azimuth reference function. The velocities v_{p1} and v_{p2} of the two points do hardly differ but the difference in their radial velocity $v_\perp$ is considerable. With ΔH = 1000 m on 46° N lat. a difference in

the radial velocities of about 1m/s results. For Seasat SAR on an ascending orbit this is equivalent to a doppler shift of about 8Hz. The azimuth reference function used in that particular range segment for the compression of both backscatter elements can not be trimmed for both frequencies. At least one of the points will be compressed with a doppler reference that does not correspond to the ideal input signal of the antenna's center. Thus the particular backscatter element will be out of focus and a reduction of the signal to noise ratio will result.

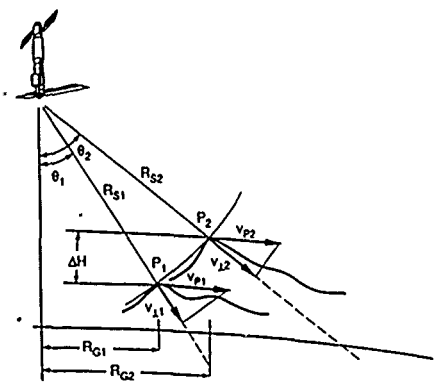


Figure 1. Effect of height differences on the radial velocity $v_\perp$

A further and far more important consequence is that due to the compression of a signal, which is backscattered from a point beside the antenna axis, the corresponding image point will lie at a wrong azimuth coordinate, an azimuth shift results. Tab. 1 shows a numerical example. The values were computed with fictitious backscatter elements. The equations used are explained in section 2.3. For one point on sea level in near range and one in far range the doppler frequency shift f_D for a squint angle of 0° was computed. If f_D is used as doppler centroid f_d for the azimuth compression, backscatter elements above sea level are shifted in azimuth direction. The azimuth shift has its maximum at near range and slightly decreases towards far range. These considerations make it obvious that for high precision geocoding the doppler approach has to be applied rigorously, i.e. for every single backscatter element.

	H (m)	θ (°)	$v_\perp$ (m/s)	f (Hz)	ΔA (m)
R_S = 851855 m	0	19.47	-110 85	-942.80	0.00
	1000	19.65	-111.83	-951.02	-111.95
	2000	19.83	-112.80	-959.27	-222.76
	3000	20.00	-113.74	-967.33	-330.15
	4000	20.18	-114 69	-975.41	-438.67
R_S = 877875 m	0	23.29	-131.42	-1117.64	0 00
	1000	23.43	-132.20	-1124.27	-91.83
	2000	23.58	-132.97	-1130.86	-182.48
	3000	23.73	-133.73	-1137.32	-271.95
	4000	23.87	-134.49	-1143.78	-361.42

Table 1. Azimuth shift ΔA dependent on the elevation H (Seasat SAR, ascending orbit, 46° lat. N, $v_{s/c}$ = 7456 m/s, f_0 = 1.274 GHz)

2.2 Geodetic Reference Systems

For an unsupervised geocoding, i.e. a rectification technique that does not require ground control points (Meier, 1989), it is necessary to precisely formulate the spatial relationship between the sensor position during data acquisition and the backscatter element. Therefore information about the used reference systems is required.

The satellite's ephemeris data are given in an earth-centered inertial coordinate system that has reference to a satellite datum such as the WGS 72 or WGS 84 (World Geodetic System) or GRS 80 (Global Reference System). The coordinate origin of these systems is intended to be the earth's center of mass, the geocenter.

The position coordinates of the backscatter elements e.g. DEM cells are usally based on topographic maps. For that reason and for the final representation of the geocoded image the mapping equations and constants must be available. Basically there are only four projection types among topographic maps (Transverse and Oblique Mercator, Lambert Conformal Conic and Polar Stereographic) (Graf, 1988). It therefore seems realistic to present the geocoded image in the map projection selected by the user. The reference system for positions on topographic maps is an oblate ellipsoid. Unfortunately different countries use different reference ellipsoids. They have their own local geodetic datum which defines the size and shape of the ellipsoid used and its location in space. The purpose of a geodetic datum is to minimize the deviations between ellipsoid and geoid in the area of interest. The origins of the ellipsoids do not coincide with the geocenter. The shift between geocenter and center of ellipsoid may amount to several hundreds of meters. (Graf et al., 1988). It therefore must not be neglected in a high precision geocoding process.

The transition between local datum and satellite datum is most conveniently carried out using cartesian coordinates rather than geographic coordinates because they are independent of ellipsoids. The complete three-dimensional Helmert transformation (Eq. 1) involves three translation parameters (Δx, Δy, Δz) to relate the origins of the systems, three rotation parameters around the coordinate axis to relate the orientation of the two systems and one scale factor (M) to account for any difference in scale between the two systems. Thus in order to precisely formulate the relationship between sensor and backscatter element not only the map projection formulae have to be known but also the geodetic datum which the maps have reference to.

$$\vec{P_T} = \Delta\vec{p} + M \cdot \mathbf{R} \cdot \vec{P_A} \qquad (1)$$

where: $\vec{P_T}$ ... transformed point
$\vec{P_A}$ ... point to be transformed
$\Delta\vec{p}$ ... transformation vector with coordinates Δx, Δy, Δz
M ... scale factor
R ... orthogonal rotation matrix

A datum transformation has also to be carried out in case DEMs based on different local datums are used for geocoding. For Fig. 2 a DEM with Swiss oblique Mercator coordinates was added to a DEM in the French Lambert conformal conic projection. In Fig. 4 a French DEM has been combined with a DEM in German transverse Mercator coordinates (see section 2.4).

2.3 Rectification Procedure

A backward transformation is used to convert the positions of the backscatter elements into the slant range image coordinates. This transformation of the three-dimensional object coordinates given in a cartographic reference system into the two-dimensional row and column coordinates of the slant range image requires position and velocity vectors of both sensor and backscatter elements as well as doppler frequencies and pulse transit times used for processing.

Out of the ancillary data sensor position and velocity have been computed for each azimuth pixel of the slant range image. The data have been stored in an earth-centered and earth-fixed Cartesian coordinate system. To enable the computation of the doppler, position and velocity of the backscatter elements have to be transformed into the same coordinate system: First, map coordinates are transformed into geographic coordinates, these are converted to Cartesian coordinates, afterwards the datum transformation is applied.

With the knowledge of the doppler centroid f_d used as azimuth reference the sensor's position can be determined for any backscatter element: For each backscatter element with a corresponding estimated sensor position the slant range R_S and the doppler frequency f_D is computed (Eqs 2 and 3).

$$R_S = \sqrt{(\vec{S} - \vec{P}) \cdot (\vec{S} - \vec{P})} \tag{2}$$

$$f_D = \frac{2 f_0}{c} \frac{(\vec{v_P} - \vec{v_S}) \vec{R_S}}{|R_S|} \tag{3}$$

where: R_S ... slant range
$\vec{S}, \vec{P}$ ... spacecraft and backscatter element position
$\vec{v_S}, \vec{v_P}$ ... spacecraft and backscatter element velocity
f_0 ... carrier frequency
c ... speed of light

The calculated frequency shift f_D is compared with the doppler centroid f_d which was used during the azimuth compression. The case $f_D > f_d$ means that the closing rate is too high and that the estimated sensor position related to the flight direction lies behind the correct position. Thus the position number will be moved one step forward and the calculation will be repeated. A program code of this iterative position search algorithm might be written like:

```
DO WHILE (fD > fd)
   N = N + 1
   ...
END

DO WHILE (fD < fd)
   N = N – 1
   ...
END
```

N means the orbit position number. The exact position, where $f_D = f_d$, lies between two sensor positions. It will be interpolated linearly because the doppler frequency rate between two state vectors is basically constant.

After the determination of the sensor position that provides the slant range image pixel coordinate in azimuth direction the line coordinate in range direction can be computed after Eq. 4.

$$n_L = (R_S - R_0) / \Delta_{PR} \tag{4}$$

where: n_L ... line number
R_0 ... slant range to first pixel (near range border)
Δ_{PR} ... pixel spacing in range

Finally a resampling based on the determined pixel and column coordinate in the slant range image gives the gray value for the processed pixel in the geocoded image.

2.4 Results

With the procedure as outlined above Seasat SAR and SIR-B images have been geocoded. The lack of the doppler frequencies that served as azimuth reference function prevented an unsupervised geocoding. Thus they first of all had to be reconstructed by means of control points. A noise reduction in the original slant range images was achieved by a Frost filter.

The position accuracy of the Seasat images amounts to about one to two slant range pixels in both range and azimuth direction (i.e. about 20m). Containing a lot of noise the SIR-B images did not permit a reliable determination of the control points needed for the reconstruction of the azimuth reference function. An accuracy of two to three pixels in both directions has been achieved (i.e. 25 to 30m). A better accuracy could not be verified because of the poor image quality of SIR-B data.

Fig. 2 shows a mosaic of two Seasat SAR scenes of the orbit 762 covering an area of 95 x 110 km. The image is represented in the projection system of the French ordnance survey. Lambert zone II, étendu. Lines of the geographic graticule help to localize the recorded area. Geneva in the center is surrounded by the Jura in the north

west, Lake of Geneva in the east and the Savoy alps in the south east. The partially fairly curved near and far range edges of the imaged area demonstrate vividly the influence of the relief. In Fig. 3 two Seasat SAR scenes near Cologne (FRG) at the Rhine can be seen (orbit 891 and 1493). The grid shows geographic coordinates. The adjacent areas, where no SAR image information was avaible, are shown with a shaded DEM. The processed area covers about 140 x 110 km. Fig. 4 is a mosaic of two SIR-B scenes imaging about 150 km² (ascending orbit 45.2 and descending orbit 97.4). The image is represented in the Gauss-Krüger system of the FRG (stripe 3). It shows the Rhine river valley with mount Kaiserstuhl near the city of Freiburg.

The mosaicing was performed in both cases entirely unsupervised. Nevertheless, an edge where the images overlap is not visible. A representation of the geocoded image in the same projection that is used for the topographic maps of the area in question is the most convenient solution for the user. A combination with other data sets, for example geographic information systems, can therefore be achieved easily.

3. VALUE ADDED PRODUCTS

For a detailed thematic or radiometric interpretation of SAR images it is very important to have knowledge about the quality of the radiometric information as shown in the geocoded image.

3.1 Layover and Shadow Mask

Fig. 5 a) shows a layover situation between the points F and C. In this area a definite spatial assignment is not possible because echos from two or more points reach the sensor simultaneously. An information overlay results which makes the interpretation questionable. Fig. 6 shows the layover mask of the north eastern part of Fig. 2. Depending on the grid size of the DEM used for the layover computation the imaged area contains between 12 and 14 percent of layover. As it can be seen from Fig. 5 a), layover is a function of the slant range R_S. Shadow on the other hand can be calculated by means of the off-nadir angle θ (Fig. 5 b)). Shadow areas represent information gaps, they do not furnish any signal for the generation of image pixels.

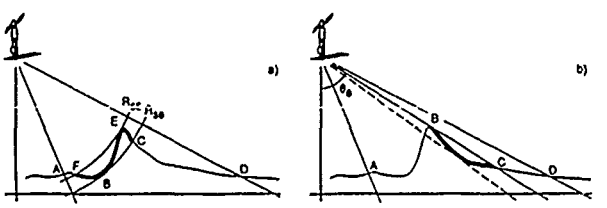


Figure 5. Layover and shadow areas

3.2 Local Geometric Values

Whereas layover and shadow areas should be excluded from interpretation, the radiometric investigation of all other areas with useful or valid information will be very much supported in case additional information about the pixels is available (Holecz, 1989 et al.). Interesting values are the local spatial resolution and the local incidence angle. The local spatial resolution means the area on the earth that is equivalent to a processed pixel in the slant range image. Because of the properties of SAR systems it must be distinguished between local azimuth resolution and local range resolution. The local incidence angle is measured between the normal of the backscatter element and the incoming radiation.

4. CONCLUSIONS

An accurate geocoding procedure has to consider all parameters that affects the image geometry. The outlined rectification procedure based on the range-doppler approach meets this requirement. It is neither sensor specific nor bound to a specified orbit configuration or particular map projection. With a rigorous application of the procedure to each pixel a single processing mode can be realized and the processing of subimages with approximative solutions can be avoided.

The accuracy is independent of the size of the scene, its geographic location and the used geodetic reference system. The lack of the doppler centroids used during processing as azimuth reference prevented an unsupervised geocoding. The image quality affects the accuracy because the azimuth reference functions had to be reconstructed out of the image. It therefore results the urgent need of the knowledge of these values used during the correlation process. The ancillary data must contain precise orbit data, the processed pulse transit times and doppler frequencies.

Value added products such as layover and shadow mask, local incidence angle or local spatial resolution map will most certainly improve the value of geocoded SAR products and therefore efficiently support image classifications and radiometric investigations.

5. ACKNOWLEDGEMENT

The work reported here was sponsored by the European Space Agency (ESA) for ERS-1 geocoding software development within the German Processing and Archiving Facility (D-PAF) at DLR, Oberpfaffenhofen.

6. REFERENCES

1. CURLANDER, J. C., "Utilization of Spaceborne SAR Data for Mapping", IEEE Transactions on Geoscience and Remote Sensing, Vol. GE-22, No. 2, pp 106 - 112, 1984

2. CURLANDER, J. C., "Performance of the SIR-B Digital Image Processing Subsystem", IEEE Transactions on Geoscience and Remote Sensing, Vol. GE-24, No. 4, pp 649 - 652, 1986

3. GRAF, Ch., "Verwendung geodätischer Abbildungen bei der Geocodierung von Satellitenbildem", Remote Sensing Series Vol. 13, Remote Sensing Laboratories, University of Zurich, 1988

4. GRAF, Ch. et al., "Map Projections for SAR - Geocoding", ERS-D-TN-22910-A/9/88, Remote Sensing Laboratories, University of Zurich, 1988

5. KWOK, R. et al., "Rectification of Terrain Induced Distortions in Radar Imagery", Photogrammetric Engineering Vol. 53, No. 5, pp 507 - 513, 1987

6. MEIER, E., "Geometrische Korrektur von Bildern orbitgestützter SAR-Systeme", Remote Sensing Series Vol. 15, Remote Sensing Laboratories, University of Zurich 1989

7. HOLECZ, F., NUESCH, D., "Study on Texture Measurements applied to SAR Images", Report to the JRC of the CEC in Ispra, University of Zurich, 1989

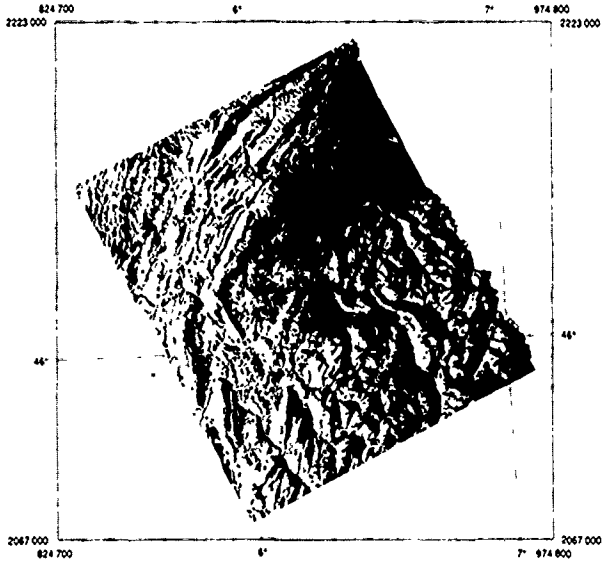


Figure 2. Geocoded Seasat SAR mosaic "Geneva"

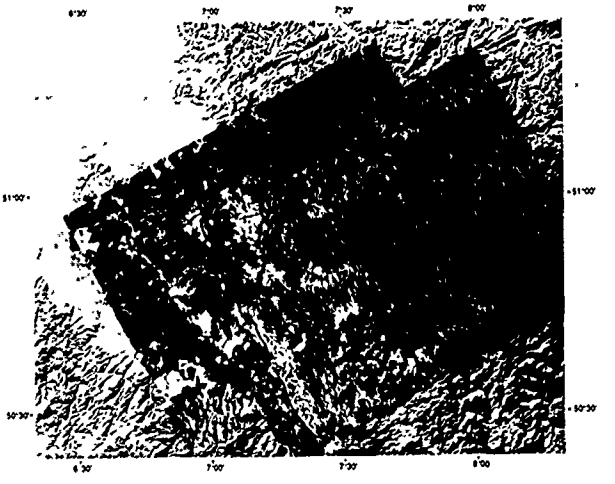


Figure 3. Geocoded Seasat SAR mosaic "Cologne"

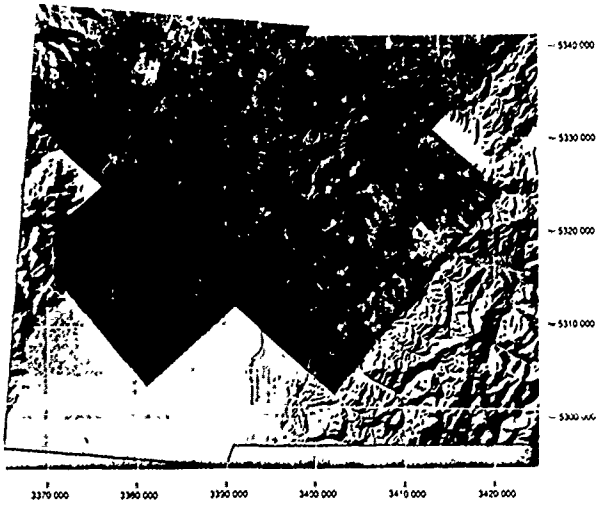


Figure 4. Geocoded SIR-B mosaic "Kaiserstuhl"

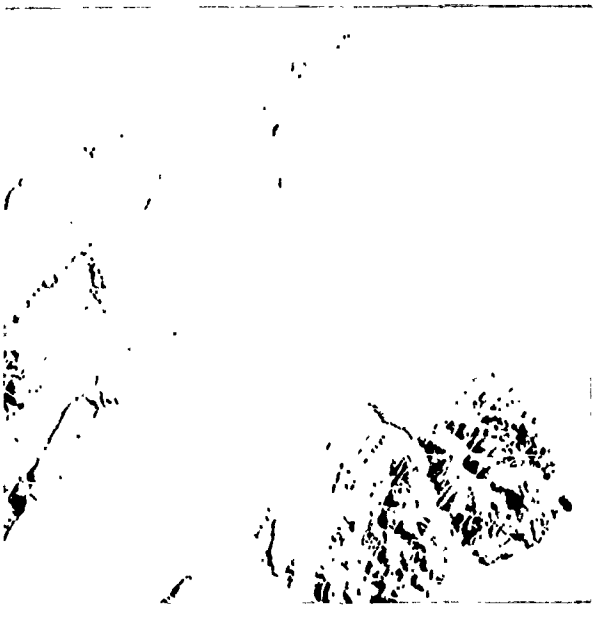


Figure 6. Layover mask of the north eastern part of Fig. 2

CLOSED FORM SOLUTIONS TO THE SINGLE-IMAGE AND STEREO SYNTHETIC APERTURE RADAR MAPPING EQUATIONS

by
Michael Kobrick
Jet Propulsion Laboratory
Mail Stop 300-233
4800 Oak Grove Drive
Pasadena, CA 91109 USA
Telephone (818) 354-4631
TELEX 67-5429

Synthetic aperture radar (SAR) imagers hold great potential for cartographic mapping because the images they produce do not suffer from the geometric distortions inherent in optical systems, and the derivation of object-space coordinates is independent of the pointing of the instrument. A minor disadvantage has been that, although the principles and mathematics governing SAR imaging are well understood, the equations are more complicated than equivalent optical systems and closed-form solutions for either single image or stereo SAR pairs do not exist. Until now.

While instruments that use lenses to form images basically are mapping brightness as a function of two orthogonal angles measured from the optical axis, SAR maps brightness as a function of two-way radar echo time delay and relative phase history. This phase history is modified by the Doppler shift caused by the relative motion of the target and sensor, and although this shift is not necessary for the SAR principle (and not even particularly useful) the common use of the terminology "Doppler" for this component will be maintained here.

Thus the location of a SAR image point is determined by the intersection of a spherical surface of constant time delay, the "range sphere", with a conical surface of constant Doppler, the "Doppler cone." This intersection defines a circle in space centered on and orthogonal to the instantaneous velocity vector of the platform. To determine where on this circle the point lies requires an additional datum, the elevation of the point. For an orbital SAR this can be expressed as a spherical earth. The latitude and longitude of the point are then determined by the intersections of the circle with this sphere, which for the non-degenerate case occurs at two points (this final ambiguity is resolved by the antenna pointing, which is either to the left or right).

Although this geometric picture is straightforward enough the three applicable equations are usually solved by iterative or least square methods since no closed form solution has appeared in the literature. Such a solution will be presented here, accomplished through the use of a not necessarily orthogonal coordinate system centered on the sensor, with axes along the radius vector to the platform, the platform velocity, and the direction defined by their cross product. The solution vector is then easily transformed back to earth-centered or any other coordinate system.

A similar solution will be presented for the stereo case. Here there are two platform locations involved and thus two range spheres and two Doppler cones. Any three of these four data may be used for the solution and the point elevation is not required, indeed this is the important quantity to be measured in stereo.

Accuracy in positioning with stereoscopic SLAR imagery for topographic mapping

J. Wu and D.C. Lin

Center for Space and Remote Sensing Research (CSRSR)
National Central University, Chung-Li,
Taiwan, Republic of China

Abstract

For side-looking airborne radar (SLAR), two radargrammetric conditions are formulated. The one equation expresses slant ranges between kinematic positions of SLR and ground points of interest. The other equation states that the scanning direction of radar pulses is oriented perpendicular to flight heading, if allowance for a systematic squint is made. For time-dependent position and attitude parameters along a flight path, only those at the SLR (pseudo-) stations separated by given intervals are really determined. This requires that use of parameter modeling be made for the orientation parameters at stations lying between the pre-selected stations.
Tests are performed on a pair of film strips with real-aperture SLAR at an average scale of 1 : 250 000, covering stereoscopically on the same side a hilly area of 24km*15km. 82 conjugate points in the test area are manually identified and their coordinates measured on the images and on 1 : 50 000 topographic line maps. Accuracy variations are determined with respect to different numbers of ground control points and of SLR (pseudo-)stations. Accuracy in three-dimensional positioning with SLAR stereo images meets the requirements on making (image) maps at a scale of 1 : 100 000 or smaller.

Keywords: Active side-looking radar, Radargrammetric equations, Parameter modeling, Real-aperture SLAR stereo images, Accuracy in planimetry and in height

1. Introdution

Side-Looking (Airborne) Radar, SL(A)R, is an active sensor having a salient feature that it can operate at almost any time day and night. From the time that elpses for a radar pulse to travel from an antenna to objects and back, ranges can be calculated. Strength of return signals depends on the aspect of and on the kinds of objects in an area illuminated and resolved by radar. The strength received may be recorded on photographic films in terms of gray shades. Knowledge about radiometric and geometric properties of side-looking radar images is essential for good qualitative interpretations and quantitative measurements.
Analytic formulations that describe the image-forming of side-looking radar relate points in image space to points and to SLR stations in object space. Analogue or digital images from side-looking radar can be applied to three-dimensional positioning of points on the earth's surface, if they are veiwed stereoscopically by radar at least twice.

2. Radargrammetric equations

When microwaves are transmitted sidewards, they look in a direction which is nominally perpendicular to the flight heading of a SLR carrier. Ranges between a radar antenna and objects on the ground are used to control synchronously positions of a light spot before a roll-film recording device. Thus, a line image is defined by range projection. A two-dimensional radar image is a record of consecutive line images synchronized with the relative speed of a carrier to the ground.
In a reference system of coordinates, positions X_i, Y_i, Z_i of a target are denoted by $\mathbf{p}$. At the moment t_j, positions X_{o_j}, Y_{o_j}, Z_{o_j} of SLR scanning the target are represented by the vector $\mathbf{s}$. With the slant range $R = |\mathbf{r}|$, the range equation is (see Fig.1) :

$$R = |\mathbf{p} - \mathbf{s}| = \sqrt{(X_i - X_{o_j})^2 + (Y_i - Y_{o_j})^2 + (Z_i - Z_{o_j})^2} \quad . \quad (1)$$

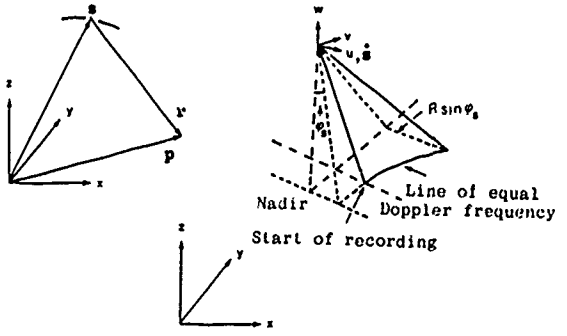


Fig.1: Range measurement with side-looking radar having a squint φ_s

Geometrically, Equation (1) stands for a range sphere with its center at $\mathbf{s}$ and of the radius R .
For side-looking radar that has a real-aperture antenna, rotational parameters φ_j (pitch), κ_j (yaw) around the axes of a coordinate system at, for instance, the center of the radar antenna have an effect on the direction of microwaves propagated to the side; ω_j (roll) may be disregarded, because it

influences ranging insignificantly. If a squint σ_s exists, ranging takes place on a cone with its apex at s and its axis coinciding with the longitudinal axis of the real antenna. The x-component of a range in the reference system, when transformed into the antenna system, should be identical to the u-component of the range. This gives the geometric condition for the direction looking to the side. Letting the u-component in Fig.1 be equal to the first equation in three-dimensional photogrammetric transformation from object to image space, one gets

$$R \sin \varphi_s = \cos \varphi_j \cos \kappa_j (X_i - X_{o_j}) + \sin \kappa_j (Y_i - Y_{o_j}) - \sin \varphi_j \cos \kappa_j (Z_i - Z_{o_j}) \quad . \qquad (2)$$

If ground range projection is preferred to slant range projection, R must be modified to $(R^2 - H^2)^{1/2}$ where H means an average flying altitude. When line images are recorded, synchronism in range direction and in azimuth (flight) direction are not necessarily unified; nevertheless, image scales in both directions are assumed to be the same and constant.

3. Models for time-dependent parameters

Parameters of exterior orientation X_{o_j}, Y_{o_j}, Z_{o_j}, φ_j, κ_j are time-dependent in that the coordinate system at a radar antenna varies when side-looking radar is carried onward station by station. In order to reduce the total number of orientation parameters to the extent that does not exceed the number of measurements made on radar imagery, a realistic parameter modeling based on linear interpolation can be set up as follows:

$$\begin{aligned} X_{o_j} &= (1 - s/d)\, X_{o_k} + (s/d)\, X_{o_{k+1}} \\ Y_{o_j} &= (1 - s/d)\, Y_{o_k} + (s/d)\, Y_{o_{k+1}} \\ Z_{o_j} &= (1 - s/d)\, Z_{o_k} + (s/d)\, Z_{o_{k+1}} \\ \varphi_j &= (1 - s/d)\, \varphi_k + (s/d)\, \varphi_{k+1} \\ \kappa_j &= (1 - s/d)\, \kappa_k + (s/d)\, \kappa_{k+1} \quad . \end{aligned} \qquad (3)$$

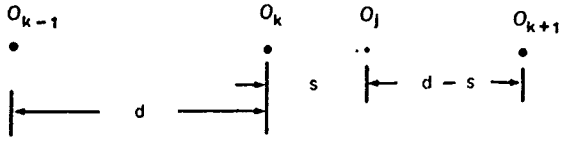


Fig.2: Side-looking radar station O_j lying between O_k and O_{k+1}

O_{k-1}, O_k and O_{k+1} in Fig.2 denote SLR (pseudo-) stations equally-separated and limited in number. Separation d, s can be expressed in units of time or strip coordinates on images. After integration of Equations (1), (2) and (3), a pair of error equations can be derived from linearization according to Taylor. Only unknown position and attitude parameters at O_k's are then determined in iterative least squares adjustments.

4. Stereoscopic SLAR imagery

In 1981, Taiwan was imaged with a real-aperture side-looking airborne radar having a nominal wavelength of 3 cm in X-band; ground ranges were recorded on analogue film strips. Range resolution is 32 m at a look angle of 70° off-nadir and azimuth resolution amounts to 57 m which is worse because of a fairly long slant range. There are 15 km or 60 % sidelap between two neighboring film strips.

Between stereo image pairs at an average scale of 1 : 250 000 as well as between the images and topographic maps at a scale of 1 : 50 000, 82 conjugate points are manually identified. Coordinates at image points are measured on WILD Analytical Plotter BC-1; coordinates at points on the maps read from a digitizer. The following standard deviations are given for calculation of a priori weights: $\pm 50\mu$m for time and range coordinates in image space, ±30 m for coordinates at GCPs (ground control points); ±500 m for kinematic positions and ±0.106 rad for attitude parameters at SLR (pseudo-) stations. No correlation is assumed between stochastic observations. Iterative least squares adjustments using CDC 180/840 with 64-bit words end, when $5*10^{-6}$ is accepted as a numerical zero for non-linear Equations (1) and (2). A pair of same-side stereoscopic images of a hilly area to the west of Taipei is shown in Fig.3.

Fig.3: Stereoscopic SLAR images (1 : 250 000) covering an area of 24 km * 15 km having a maximum height difference of 611 m

When three SLR (pseudo-)stations or orientation points are made available as O_{k-1}, O_k and O_{k+1} in Fig.2, there are after 7 iterations at 23 GCPs root mean square errors in x ±42.7 m, in y ±50.9 m and in z ±12.3 m; at 59 independent check point σ_x = ±74.7 m, σ_y = ±69.7 m and σ_z = ±111.2 m. Discrepancies between the known and computed coordinates in planimetry and in height are plotted as residual vectors in Fig.4.

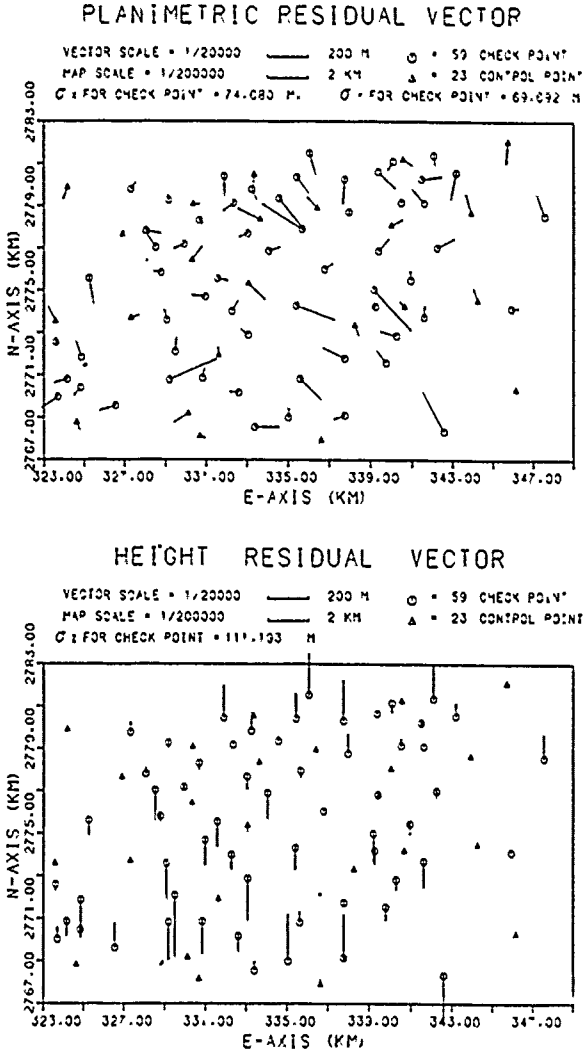


Fig.4: Coordinate differences at ground control/check points in the hilly area

How accuracy measures σ_x, σ_y and σ_z vary as a function of the number of GCPs can be examined in Fig.5 where one notices that the number is reducible to 12.
In Fig.6, the relation between variations of accuracy and numbers of SLR stations is illustrated, while 23 GCPs are in use. Variations of position and attitude parameters after convergence of adjustments in cases of 2, 3, 5, 6 SLR stations (orientation points) are superimposed in Fig.7 for each of the two flight paths.
The reason why accuracy in height deteriorates as the number of SLR station increases to six is in part that the number of conjugate points in each segment

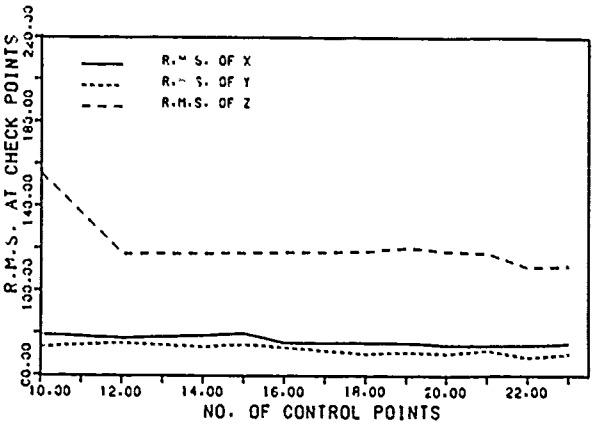


Fig.5: Accuracy variations vs. numbers of GCPs

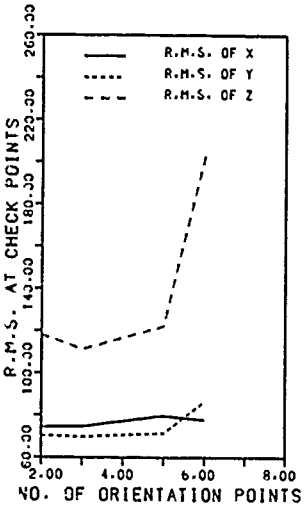


Fig.6: Accuracy variations vs. numbers of orientation points along flight paths

decreases approximately from 40 (three stations) to 16. In particular, one notes that fling altitudes are estimated rather dissimilarly. While they are lower than 5 km along the left flight path, they are near to 6 km along the right path. If standards for making (image) maps at a scale of 1 : 100 000 are drawn to as references : ±50 m in planimetry and ±17 m (one third of a 50 m contour interval) in height, the SLAR stereo images analyzed with radargrammetric Equations (1), (2) are qualified only for planimetric mapping at scales not larger than 1 : 100 000.

5. Concluding remarks

The reason for slightly better accuracy in y (cross-track) is evident, for the ground strip is scanned twice in stereo by side-looking radar. The accuracy in x (along-track) is not so good as that in y, partially because parameter modeling by linear interpolation is insufficient. Height accuracy being the worst can be attributed to acute stereo intersection angles reaching

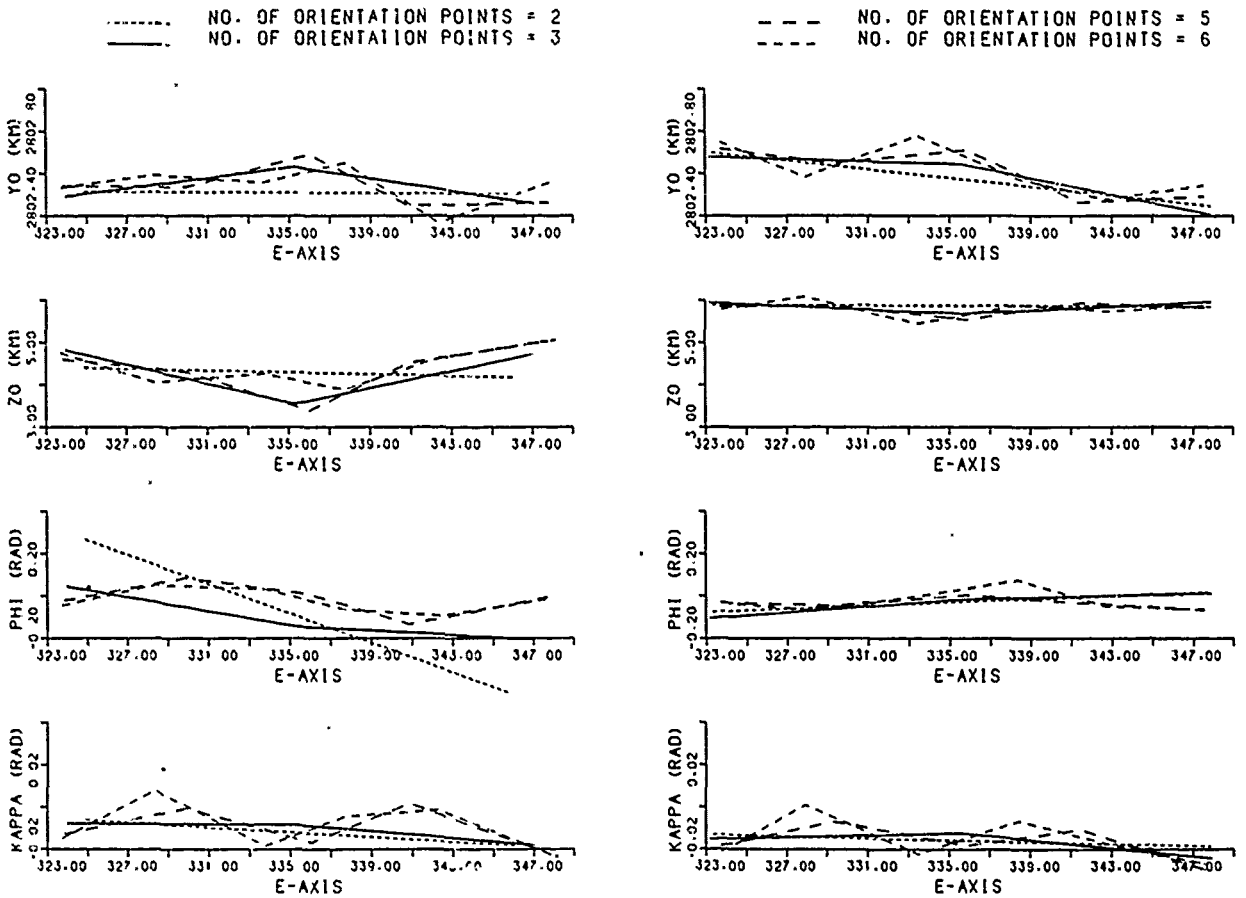


Fig.7: Four sets of parameters estimated at SLR stations

from 2°.5 to 10° for same-side geometry. There are often non-linear variations of scales and deformations on photographic records. It is better to correct SLR images for systematic errors calibrated in advance; otherwise radargrammetric equations should be extended functionally to allow for additional self-calibrating parameters.

With time-referred parameters of exterior orientation being known beforehand, coordinates at points on the earth's surface can be determined more accurately, because three-dimensional positioning has a combined characteristic. Therefore, it is recommended that navigation data of carriers (their position and attitude or velocity variations) be stored simultaneously for later use with side-looking radar imagery.

6. References

Curlander, J.C., Kwok, R. and Pang, S.S.,1987 : a post-processing system for automated rectification and registration of spaceborne SAR imagery, International Journal of Remote Sensing 8(4), pp. 621-638

Konecny, G. and Lehmann, G., 1984 : Photogrammetrie, Walter de Gruyter, Berlin, ISBN 3-11-007358-7

Leberl, F., 1976 : Imaging radar applications to mapping and charting, Photogrammetria 32, pp. 75-100

Leberl, F.W., 1987 : Some mapping applications for digital SAR imagery, proceedings of the Willi Nordberg Symposium in Graz, pp. 109-114

Light, D.L., 1980 : Satellite photogrammetry, Chapter XVII of the Manual of Photogrammetry by C.C. Slama (ed.), ISBN 0-937294-01-2

Mars, Inc., 1981 : SLAR Survey of the Republic of China, Final Report

Moore, R.K., 1976 : Active microwave systems, Chapter 7 of the Remote Sensing of Environment by J. Lintz, Jr. and D.S. Simonett (eds.), ISBN 0-201-04245-2, pp. 234-290

Norvelle, F.R., 1972 : AS-11A radar program, Photogrammetric Engineering 38(1), pp. 77-82

Wu, J., 1988 : A sound approach for stereo three-line scanner's images, International Archives of Photogrammetry and Remote Sensing 27(B8), pp. III-45 to III-51

Ziemann, H., 1985 : System calibration and self-calibration with fully-controlled vertical aerial photography, Technical Papers of the American Society of Photogrammetry Annual Convention in Washington, D.C., pp. 736-745

EFFECTS OF RADAR IMAGE DISTORSIONS IN A RADARGRAMMETRIC APPLICATION OF VARAN-S DATA

Laurent POLIDORI* DEA de Méthodes Physiques en Télédétection.
Université Paris-7, 2 place Jussieu, 75252 PARIS CEDEX 05 (FRANCE).

Paul KAMOUN. AEROSPATIALE. 100, boulevard du Midi, CANNES LA BOCCA (FRANCE)

Hannes RAGGAM. Institute of Image Processing and Computer Graphics, Wastiangasse 6, A-8010 GRAZ (AUSTRIA).

Abstract

A radar stereopair over Montagne-Sainte-Victoire (Southern France) was chosen in order to test the suitability of the Varan-S data for a radargrammetric application. A radar stereomapping software was used at the Institute of Image Processing and Computer Graphics (Graz, Austria) to provide a DEM (Digital Elevation Model) of the test-site. The study showed that the Varan-S data are fit for a radargrammetric application unless they display undue distorsions.

Keywords: SAR, distorsions, radargrammetry.

1. INTRODUCTION

Radargrammetry dates back to the early 50s, but no accurate results could be obtained before the 70s, when the combined use of SARs (synthetic-aperture-radars) and analytical stereoplotters allowed the extraction of topographic information from radar stereopairs with a reasonable accuracy ([1] , [2]).
The purpose of our experiment was to evaluate the Varan-S data from a radargrammetric point of view. As will be shown later, although the inherent characteristics of those data are suitable for stereomapping, some accidental distorsions may prevent the set-up of a stereo model.
The aim of radar stereomapping is to obtain a DEM (Digital Elevation Model) using two overlapping radar images, the known flight parameters and some GCPs (Ground Control Points).
Mathematically the problem consists in positionning a target within an XYZ coordinate system using its radar coordinates r, t, τ on both images. Each image provides, for every ground point, a slant-range r, a time t and a squint angle τ. This point can then be located on a circle defined by its radius r, its center (the antenna position at t) and its axis, which makes an angle τ with the platform velocity vector. The imaged point is located at the intersection of the two circles corresponding to the two images.
In the usual case of parallel flight paths, we can consider two different configurations: same-side looking and opposite-side looking. The Varan-S data over the Montagne-Sainte-Victoire test-site (fig. 2) were obtained in a same-side looking configuration (which allows an easier stereoscopy focusing) and with a zero Doppler shift, i.e. with no squint angle.

2. THE VARAN-S DATA

Varan-S is a French SAR developped by CNES (Toulouse) since 1984 ([10]). Its main characteristics are:

- wavelength : 3.2 cm
- frequency : 9.375 GHz
- polarisation : HH, VV, VH or HV
- sampling window : 60 µs
- peak power : 6 kW
- noise figure : 7 dB
- antenna length : 1 m

The pixel size is 1.60 meter. Each image is a 10.080 km side square, i.e. 6300 x 6300 pixels. The images of the stereopair must exhibit:
- a good radiometric quality, for an easy stereoscopic observation and an accurate plotting of GCPs and contour-lines;
- a good geometric quality because distorsions cause parallax and can hinder the model set-up.

Geometric and radiometric quality are closely dependent. Indeed a geometric feature is only a contrast in the radiometry, and geometric rectification includes a recomputation of the radiometric values. The Montagne-Sainte-Victoire images have a good geometric quality, but the site has very few planimetric features and a very homogeneous landcover, so that the GCPs were identified with difficulty.
The same experiment has been run with Varan-S data acquired over the Cévennes region (fig. 1); but because of its excessive radiometric and geometric distorsions, the stereopair used did not allow good stereo viewing and measurement (fig.7).

3. STEREOSCOPIC OBSERVATION OF RADAR IMAGES

The Varan-S images were used in an analog form, on negative slides which were put up on the plotter plate.
Stereoscopic observation does not involve the same rules whether it concerns radar images or aerial photographs. The differences are of two main kinds:
- From a geometric point of view, the observation is different because the radar images are built with an unusual perspective, whereas the conical perspective of photography is very similar to human vision. The curvatures and angles of some ground features can vary from one radar image to another. Besides the layover effect and range projection generate a relief inversion, so that a mountain which is seen from the South seems to be seen from the North (fig.3). Another consequence of this effect is that some pixels represent two terrain points at least, in which no stereo measurement is possible, for one point does not define a beam.
- From a radiometric point of view the differences lie in the textural aspect of radar images. Indeed the active character of the sensor and the microwave physical properties give the radar data an unusual aspect for a human eye used to optical images of sun-illuminated objects.

But the main difficulty of radar images stereoscopic observation lies in the image distorsions. An acceleration of the platform causes defocusing (alteration of azimuth resolution comparable with an real-aperture-radar image); besides it compresses or dilates the image and almost prevent the affected areas from being stereoscopically observable because of the resulting azimuth parallax. A wavy pattern effect, due to interferences and called wriggle, adds an extra difficulty to near-range stereoscopic observation although it does not involve any geometric deformation, because it creates a false relief impression. The near-range area (generally the most

* now at IGAC (Bogotá, COLOMBIA).

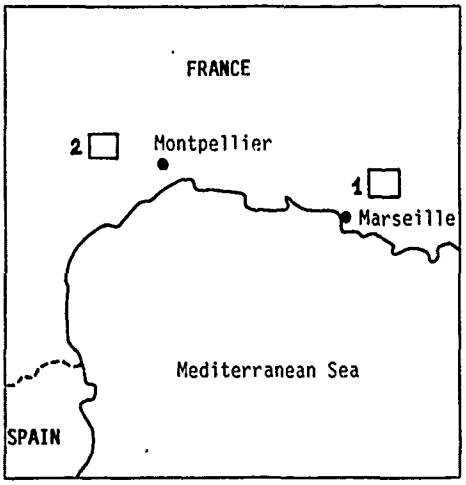


Fig.1: the Montagne-Sainte-Victoire (1) and Cévennes (2) test-sites.

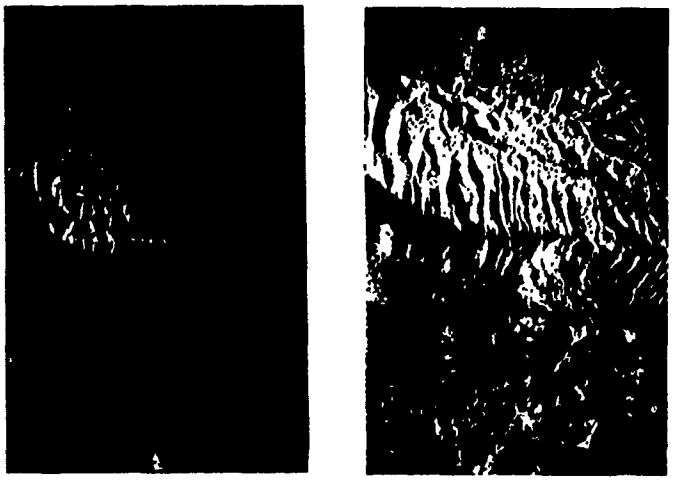


Fig.2: the Varan-S stereopair over Montagne-Sainte-Victoire.

distorted one) is affected by a saturation effect too. With a synthetic antenna the energy received from distance r is proportional to $1/r^3$. Hence in the Varan-S images, the geometry of which is shown in fig.4, the antenna receives ten times more energy from near-range than from far-range. Consequently an r^3 correction is needed. But the very near-range saturation cannot be corrected, so that the grey-values systematically remain very low. On the plotter the brillance can be regulated independently for each image: the saturated one can be lightened more than the other one. Nevertheless, the radiometric saturation it suffers from cannot be compensated. A self calibration could allow a correction of this effect on digital images.

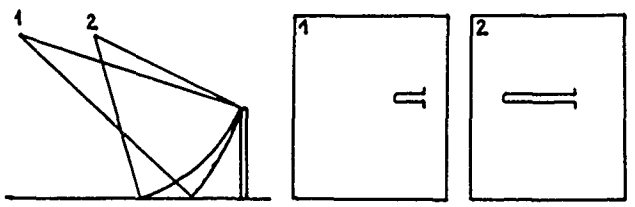


Fig.3: the layover effect

Fig.4: Geometry of Varan-S images

4. EXTRACTION OF THE DIGITAL ELEVATION MODEL

4.1 Basic equations: The Z-coordinate computation relies on two basic equations ([3]):

- squint condition:

$$\frac{ds}{dt}.(p - s) - \sin(\tau).\left|\frac{ds}{dt}\right|.| p - s | = 0 ;$$

- range condition:

$$r - | p - s | = 0 ;$$

where r is the target slant-range, p and s the target and the sensor position vectors and τ the radar squint angle.
They define the circle on which the imaged point is located (see §.1). The former defines the orientation of the circle axis, while the latter defines its radius. The target height is computed by intersecting the two circles corresponding to the two images.

4.2 Hardware and software: The analytic stereoplotter used for that experiment was a Kern DSR-1 assisted by a host computer PDP-11. The SMART software is described in [3]. Some 20 GCPs were digitized from a 1/25000 scale map on a digitizing table. Once the stereo-model has been set-up using the basic equations, the pair can be stereoscopically observed and the contouring process can start. The contour-lines were recorded on a file for the interpolation of the DEM.

4.3. Results: The stereopair was stereoscopically observable in spite of the effects of layover, foreshortening, saturation, shadowing,... As it was said in §.2, the lack of planimetric features reduced the number of GCPs and their coordinates accuracy. Nevertheless the stereo-model was set-up with a reasonable accuracy and a topographic map was drawn with a 50 meter equidistance (fig.5). The contour lines were interpolated to obtain a 20 m resolution DEM (fig. 6). The RMS errors are:

Ex = 51.3 m
Ey = 36.7 m
Ez = 38.1 m

Fig.5: radar derived topographic map

Fig.6: radar derived DEM

They provide a good evaluation of the DEM's error. 8.5 contains an interpretation of that error.

5. INTERPRETATION OF THE ERRORS

There are several types of errors. We saw that their combination amounts to some tens of meters.
Some errors are almost unavoidable, and we could not have corrected them. The site itself was unpropitious to an accurate mapping. Its lack of planimetric features jeopardized the accuracy of the GCPs' coordinates, and its mountainous character required the GCPs to be chosen in steep areas, so that any planimetric plotting error caused a corresponding height error, which became more important for steeper areas. Besides relief causes layover, foreshortening and shadowing, which help the morphological interpretation but make the stereoscopic observation less accurate. We could note that the range RMS error is far greater than the azimuth one, which is a logical result, for the relief contribution to parallax is only in the range direction.
Some errors lie in the data and could not have been corrected easily, unless in some cases with a digital processing (filtering, self-calibration, geometric rectification,...). Variations of the platform attitude are an important source of error and they are compensated in real-time: while the rolling effect is of no consequence for a zero Doppler shift, yaw and pitch offsets, due to heading rectifications, produce azimuth first-order errors and range second-order errors. Accelerations lead to geometrical over- or sub-sampling and to defocusing, and they do not affect the two images in the same way: they make the observation very difficult and degrade the results' accuracy in the azimuth direction. Interferences are a source of error as well: speckle, due to the reflexion of coherent radar waves on rough surfaces, alter the radiometric resolution ([12]) and the wriggle fringes create a false relief impression. The image of metallic features, for the strong backscattering they produce and more generally every bright image feature, are often saturated and oversized, so that the GCPs could not have been chosen among them with accuracy.

Some errors occurred during the experiment and could have been avoided. The GCPs were digitized on a 1/25000 map: the combination of the digitizing errors (due to both the operator and the digitizing table) and the map inherent uncertainty, may have induced a planimetric error of several meters (2.5 m for 0.1 mm). More over the altitudes were visually interpolated between the 10 m increment contour-lines: that can have led to errors of a few meters too.
Some of the flight parameters were not available for this study, and the platform motion had to be considered as linear, horizontal and uniform in a first approximation, and as a second-order curve after the resection in space. We unfortunately had to use bad quality negative images, extensions of positive papers: this was one of the most critical sources of error.
On the contrary some errors are absolutely negligible. The underground penetration is generally inferior to the wavelength (3.2 cm) and the stop-start approximation (see [11]) is justifiable within 1 cm. Even the measuring and sampling uncertainties can be disregarded because they are inferior to the resolution.

6. CONCLUSION

This experiment showed that the Varan-S data are fit for a radargrammetric utilization with a 50 m accuracy at least. But the quality of the results could have been improved under certain conditions:
- During the data acquisition process the platform should have drifted towards an approximate heading, so that the images would have been angularly shifted but almost not distorted;
- The GCPs' coordinates should have been directly measured on the field using geodetical points;
- The platform motion should have been interpolated between numerous and accurate position, attitude and velocity values;
- The digital images should have been directly converted into a transparent analog form, with a Vizircolor or similar system .

This being so, the resulting accuracy could have been greatly improved.
From now on additional experiments to those described

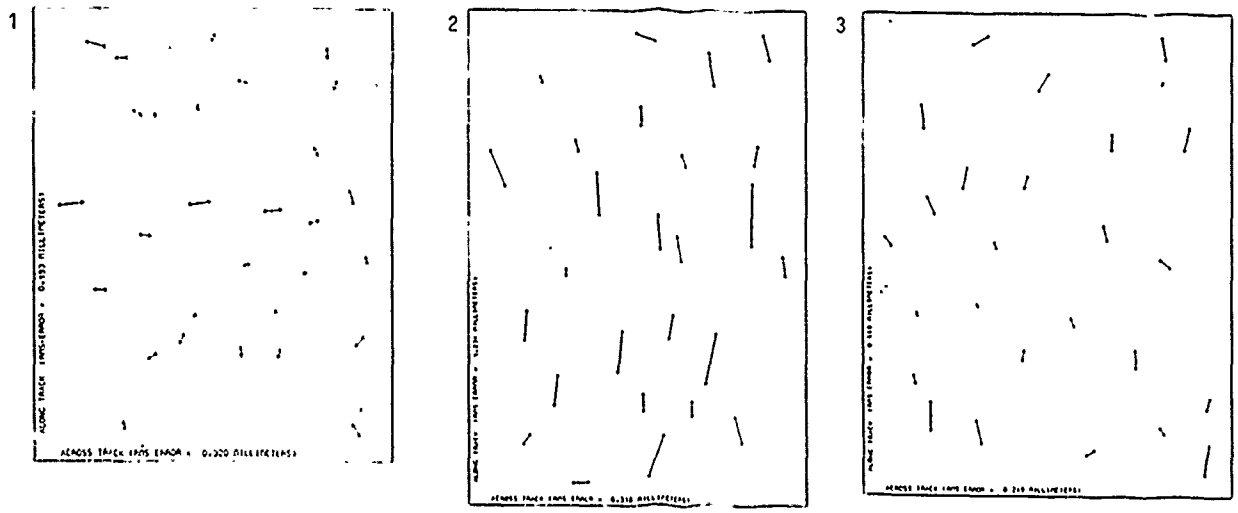


Fig.7: RMS errors of the Montagne-Sainte-Victoire (1) and Cévennes (2 and 3) stereopairs (3-parameter fitting of set of homologous points).

in this paper will be carried out as a preparation for the use of the ERS-1 data when they become available. Indeed it is expected that in a very near future, the ERS-1 stereo images will provide topographic information in countries where it is needed.

ACKNOWLEDGMENT

We are grateful to D.Massonet and J.P.Rudant for the data set and for their help. The Institute of Image Processing and Computer Graphics (GRAZ) and SODETEG (VALBONNE) provided hardware and software. They are gratefully acknowledged.

REFERENCES

1. Norvelle F.R., "AS-11A Radar Program" Photogrametric Eng., jan.1972, vol.38, Nº1, pp.77-82.
2. Leberl F., "Imaging Radar Applications to Mapping and Charting", Photogrammetria, nov.1976, vol.32 Nº3, pp.75-100.
3. Raggam J., "Radar Stereomodel Set-up on the Analytical Plotter Kern DSR-1", Technical University and Graz Research Center, Austria, 1984.
4. Derenyi E.E., "Topographic Accuracy of SLAR Imagery" Bildmessung unt Luftbildwesen, jan.1975, vol.1, pp.17-22.
5. Graham L.C., "Flight Planning for Stereoradar Mapping", Goodyear Aerospace Corporation, Litchfield Arizona 85340.
6. Leberl F. & Domik G. & Kobrick M., "Mapping with Aircraft and Satellite Radar Images", Photogram. Rec., 1985, 11 (66), pp.647-665.
7. Bear G.L. & Carlson G.E., "Height Measurement with Stereoradar", Photogram. Eng. & Rem. Sens., feb. 1975, vol.41, Nº2, pp.167-176.
8. Leberl F. & Domik G. & Raggam J. & Kobrick M., "Radar Stereomapping Techniques and Applications to SIR-B Images of Mt Shasta", IEEE Tr. Geosc. & Rem. Sens., vol.GE-24, Nº4, jul.1986, pp.473-481.
9. Raggam J. & Triebnig G. & Buchroithner M.F. & Domik G. & Leberl F., "Radargrammetric Aspects of SAR Data Evaluation", Proc. Workshop on Thematic Applications of SAR Data, Frascati, Italy, 9-11 sept. 1985 (ESA SP-257), pp.57-64.
10. Vaillant D., "Varan-S: An Airborne Synthetic Aperture Radar for Research in Microwave Remote Sensing", Earsel/ESA, Symposium on European Remote Sensing Opportunities, Strasbourg, France, 1985, pp.167-174.
11. Barber B.C., "Theory of Digital Imaging from Orbital Synthetic Aperture Radar", Int. Journal of Rem. Sens., 1985, vol.6, Nº7, pp.1009-1057.
12. Barber B.C., "Some Properties of SAR Speckle", in Satellite Microwave Remote Sensing, edited by T.D.Allan, Ellis Horwood, chap.8, 1983.

SAR Image - Terrain Database Registration

James Curlander, Wolfgang Kober, Franz Leberl

VEXCEL Corp.
2477 55th Street Suite 201
Boulder, CO 80301
Tel: (303) 444-0094
Fax: (303) 444-0470

Abstract

The capability of estimating sensor platform position using SAR imagery, and a Digital Elevation Model (DEM) or a feature database has been the object of extensive research. Inertial Navigation System (INS) data is subject to drift, and it is assumed that no Global Positionu g System (GPS) data is available. Two approaches to image-database matching are discussed in depth, along with some results using actual data. The first approach involves the use of a DEM to form simulated SAR images which are compared to the actual SAR image.

A second approach involves segmentation of candidate features from a SAR image and matching with possible corresponding features in a feature database. A particularly efficient data representation is used to eliminate features that are unlikely to be matched from the process. A logical extension of this problem lies in the decision of which features to choose from a feature database, given a particular SAR scenario. A rule based expert assistant for such a task is discussed.

Range differences in the images are used to estimate a correction to the estimated sensor position. A novel method of combining range differences to linearize the resection problem [Smith, 87] is used to simplify the estimation of the sensor position.

Introduction

Image to map matching or geocoding, is the process of assigning a geographic coordinate to each pixel of a sensed image. This task can be achieved in a number of ways. One approach involves computng an estimate of the sensor platform position and possibly the sensor attitude. This problem is called "resection in space" in photogrammetric terminology. In the past, resection in space has been achieved with computational procedures which relate individual pixel locations to ground control points with known geographical coordinates. These procedures are based on point to point correspondences. If enough individual control points are imaged and recognized, then such procedures are sufficient. However, often an imaged scene does not contain the signatures of sufficiently many individually recognizable control points, but rather contours of known objects.

Within this scenario, the flight log for actual SAR missions is based upon INS, and GPS data is not available. We use the flight log data as the original estimate of platform position. Since, INS is subject to drift, our original estimate of platform position is subject to a finite, but random error. Hence, the match and resection operations must account for and correct this error.

The first approach described below is based upon exploitation of contour to contour correspondences. We selected terrain induced shadows in SAR images because they are definitive and relatively static contours for use in matching with Digital Elevation Model (DEM) based simulated SAR imagery. The method presented consists of a two stage matching algorithm. The first stage consists of a classical area correlation used to provide a better approximation of matching locations and reduce error from the original platform position estimate provided by flight data. The second stage consists of an exact point to point matching based upon corresponding shadow contours in the actual and simulated SAR images.

The second approach is based upon matching of feature data. Clearly, the extraction of features in the sensed imagery should not proceed independently of knowledge represented in feature databases, and any approximate knowledge of sensor platform position. Unguided segmentation based solely on image luminance values has long been known to lead to problems because the object and intensity boundaries are not always the same. The approaches generally taken can be classified into two areas, namely region-based, and boundary-based. The method presented utilizes a combination of classically region-based correlation and boundary-based matching on preprocessed SAR and photographic imagery, and produces a rank for comparison of the goodness of matches. The particular representation of the feature boundaries allows fast and efficient feature matching. This method was demonstrated using National High Altitude Photography (NHAP), and the preprocessing required for SAR imagery was investigated.

Particular features must be chosen from a feature database for matching to the image. A rule-based system for choosing data from a feature database provides a logical extension for automation of the feature matching process. The immediate goals of such a system would include choosing features that would be likely candidates for matching in an image, as well as weighting chosen features according to several preselected criteria.

Once a point to point correspondence is determined, either by shadow matching or feature matching, the resection in space can establish the platform position. Since the resection generally consists of a non-linear system of equations, a method of employing range differences to linearize the problem is applied, and the system simplifies to one that can be solved by the generalized inverse solution.

Matching Using Simulated SAR Imagery

In the particular case when GPS data is not available, and not enough control points are available for a resection in space to be successfully computed, additional data must be found in order to accurately determine platform position with the resection procedure. In this section, we describe a method that employs contour to contour matching using terrain induced shadows in radar imagery to augment or replace control point data to improve accuracy of the subsequent resection in space.

The method requires the availability of a DEM from which a simulated radar image is created using the flight log parameters from the radar mission. These parameters which generally consist of an approximate straight line flight path will

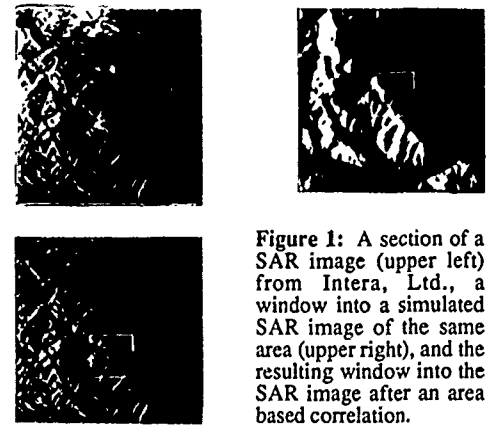


Figure 1: A section of a SAR image (upper left) from Intera, Ltd., a window into a simulated SAR image of the same area (upper right), and the resulting window into the SAR image after an area based correlation.

act as the initial estimate of the platform position. Also, it is assumed that the actual flight direction is known to within one degree, so that rotational effects are negligible. The range and azimuthal resolution of the DEM and the radar image must be comparable. A two stage matching algorithm is then employed between the simulated and actual radar images.

The first stage of the algorithm involves an area based correlation between a subsection of the actual SAR image that is known to contain terrain induced shadows, and a section of the simulated image. Figure (1) illustrates this step using STAR-1 SAR imagery from Intera Ltd. of Calgary, Canada, and a simulated SAR image created at VEXCEL. The subsection of the SAR image can be determined from analysis of the DEM using the radar flight parameters to determine reasonable areas in which to find terrain induced shadows. The subsection of the simulated image is determined by the amount of error expected in the initial estimate.

In our experiments, we used the entire simulated image for correlation, assuming that the error could be as much as half the image size. A hierarchical form of correlation was used that employed a sequence of matches on a pyramid of reduced resolution versions of the real and synthetic images taken along the approximate straight line flight path described by the initial estimate. Coarse registration is performed at the coarsest resolution, leading to finer adjustments on higher resolution versions of the imagery [Rosenfeld, 84].

Both normalized correlation, and sum of absolute differences correlation [Barnea, Silverman, 72] were tested for the area based correlation scheme, with similar results. The sum of absolute differences correlation was possible because the radar simulation incorporates a radiometric equalization with the actual SAR image histogram. Also, the sum of absolute differences metric proved to be more computationally efficient than normalized correlation and thus would be the method of choice. Since rotational inaccuracies are very small, the area based correlation can generally bring the actual and simulated images to within a few pixels of each other, but not close enough for precise pixel to pixel matching to be successful.

The next step uses contour to contour matching in order to achieve a point to point correspondence between the simulated and real images. Since the simulated image was created from the DEM by a known process, once the actual SAR image is matched to it, the pixels of the SAR image are, in essence, known in DEM coordinates. The resection can then be used to determine the sensor platform position in the DEM coordinate frame. Such a resection carries out at sufficiently small steps along the flightpath can provide an incremental correction to the INS data such that a more exact flightpath can be determined.

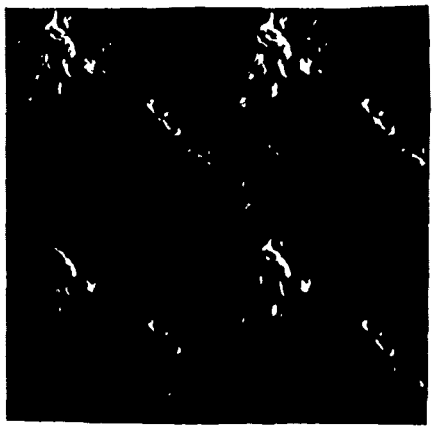


Figure 2: A comparison is sigma filter images. The original SAR image (upper left) and increasing values of sigma for smoothing: σ=0.15 (upper right), σ=0.30 (lower left), σ=0.45 (lower right).

The second stage of the algorithm involves matching the outlines of the shadows within the correlated areas in the actual and simulated images. First, the shadow outlines must be extracted from each image. No speckle noise is added to the simulated radar images, therefore, the shadow outlines can be extracted by thresholding the image at zero, and employing an edge following algorithm to vectorize the shadow outlines.

Shadow segmentation in the real SAR image is based upon the inherent bimodality of the local histogram in the correlated shadow region. When noise and speckle content is high, however, this bimodality can be corrupted. In such cases, the modality is restored using nonlinear preprocessing to preserve shadow edges and smooth noise. An example of such processing is nonlinear speckle suppression using sigma filtering [Lee, 83]. See Figure (2). This can also be viewed as an edge interest operator.

The SAR image is then thresholded at the gray value corresponding to the minimum in the bimodal histogram, and the shadow edges are vectorized with an edge following algorithm. Next, the vectorized edges in each image are encoded in $\Psi(s)$ representation. The $\Psi(s)$ representation encodes a two dimensional polygon into a one-dimensional function by taking the local tangent angle as a function of the arc length [Ambler et al. 75]. Making the arc length sufficiently large results in a smoothed $\Psi(s)$ function. This provides a simply way to remove pathological small perturbations in the shadow contours.

Clearly, the points of highest curvature along any given polygon will be given by the zero crossings of the second derivative of the $\Psi(s)$ function. We found

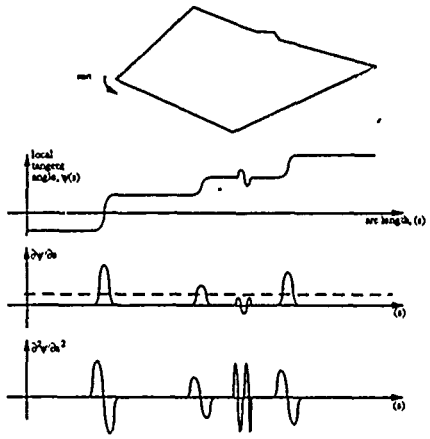


Figure 3: Thresholding for the maxima of the first derivitive of the $\Psi(s)$ function allows for extraction of points of highest curvature, and avoids small scale roughness in shadow contours.

however, that for the case of terrain induced shadows, the actual vectorized contours are somewhat ragged. Hence, the zeros of the second derivative pinpoint every inflection in curvature around the contour. In order to extract only the points of sharp tangent angle changes, the first derivative of the $\Psi(s)$ curve was thresholded for sections of maximal curvature fluctuation. The median of each thresholded section was taken to be the point of highest curvature. In this way, only points along the contour at which the tangent changed sharply would be extracted. See Figure (3).

This $\Psi(s)$ function represents the angular measure of the local tangent described by the derivative with respect to the arc length of the vector function defining the curve. Since curvature is associated with the second derivative, we see that the maxima of curvature occurs where $\Psi'(s)$ is maximum.

This process was carried out for both the actual and the simulated SAR images. Finally, points in the simulated and actual SAR images are matched based upon their locations within the corresponding correlation windows, as well as their angular locations with respect to one another. See Figure (4).

Next the simulation procedure which led to a ground range radar projection is reversed, and the matched pixels in the SAR image are assigned coordinate locations in the DEM reference frame. A resection in space can be calculated to yield and estimate of the platform position. A small error will occur because the association of SAR image and DEM points is dependent on occluding points for shadows. If the sensor position changes, the occluding points change slightly. However, this is a second order effect which may be successively reduced if the estimation process is repeated by iteration on the new estimated sensor position.

Matching Using Feature Database

Another method for determining ground control points involves the use of imaged feature data. This method is particularly useful when a DEM is not available, or not sufficiently accurate, but flight data from the imaging mission is still suitable for an initial estimate. For large scale implementation of a feature matching method, a digitized feature database such as the DMA Digital Landmass Database, is required. For the purpose of experimentation, we utilized DMA Digital Feature Analysis Data (DFAD) in a region approximating the actual imagery we tested.

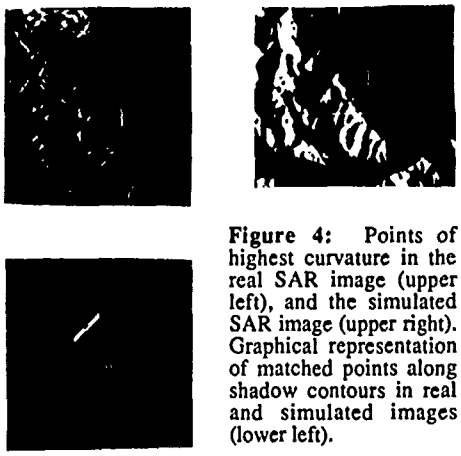


Figure 4: Points of highest curvature in the real SAR image (upper left), and the simulated SAR image (upper right). Graphical representation of matched points along shadow contours in real and simulated images (lower left).

Classically, feature matching algorithms are either region based, or boundary based. The method presented here is a combination of both types. First a boundary extraction and segmentation is applied, then a template oriented correlation matching metric is used to find the best boundary to match the imaged feature.

DMA Digital Feature Analysis Data consists of four types of features: point data, collections of points, lineal data, and areal data. All four types consist of lists of points in geographic coordinates which, for the latter two types, when linked together by straight lines describe a feature by a polygonal approximation.

The appearance in an optical image of such a polygonal approximation will be modified by a perspective transformation. A ground range SAR image will similarly have some distortions if the feature lies in terrain relief.

The feature matching procedure explored, used NHAP data which exhibited a negligibly small amount of perspective distortion. The method itself involved an efficient search for one particular arbitrarily chosen lineal or areal feature. In a fully automated system, the initial feature used in the matching would be chosen by a rule-based expert assistant. This idea is discussed more fully in the last section.

Once the initial feature was chosen, the optical image was preprocessed using the Sobel edge operator [Pratt, 78], and thresholded at two standard deviations above the mean gray value to produce a binary image containing the major edges. We found that images processed with the Sobel edge operator tended to exhibit a log-normal histogram. The micro-edges tended to occupy the lower "less bright" gray values, while the major edges generally due to outlines of major features, were brighter. The thresholding step removed the micro-edges, leaving only the bright major boundaries in the binary image. Isolated pixels were then deleted, completing the preprocessing of the imagery.

Next, the DFAD was converted to image coordinates from geographic by a linear transformation given the resolution of the image. The selected feature was stored as a list of pixel coordinates extending between the transformed DFAD coordinates that described it. The raster coordinates along the feature were then subsampled to produce a mask, which was subsequently correlated with the binary edge image using sum of absolute differences correlation [Barnea, Silverman, 72]. The correlation was carried out only at locations in the edge image where the first pixel of the feature mask fell on a bright pixel, thus increasing computational efficiency of the operation. The top three locations of correlation maxima were retained, and the full feature pattern was then correlated at those locations to produce the best estimate. Localized edge detection was then employed to produce better estimates of the original DFAD tie points and the image. Once one feature was matched, a localized method could be applied to features throughout the image to gain multiple tie points, leading to a resection operation.

Presently, a human must intervene in the above operations to choose the initial feature and to resolve conflicts during the correlation steps. Therefore, the development of a rule-based expert assistant is a logical step toward automation of this process. The groundwork for such a system is described in the last section.

Linearization of Resection Problem Using Range Differences

A novel method of combining range differences borrowed from a method used in sonar [Smith, 87] is used to simplify the estimation of the sensor position. The technique was originally developed to localize a single source given a set of noisy range difference measurements. This is ideally suited to our problem in that it should be assumed that some error is associated with all matched ground locations, as well as range data. The method is derived from linear least squares estimation theory, and approximates the maximum likelihood estimate without requiring the more rigorous solution to the system of nonlinear equations provided by a classical resection in space.

The scenario assumes that the direction of the flight path is known to sufficient accuracy, that the INS positional errors for the imaging period give good relative positional estimates, and that what is required is an offset estimate in crossrange and altitude for each azimuthal section of an image. Thus, if an exact flight path were required, an offset would have to be computed for each range line of an image. This offset, of course, must be based upon data from range lines around the particular line of interest.

Given that N ground locations were matched, we set up the problem as follows:

$$\begin{aligned} \bar{x}_S &= \text{sensor position} \qquad (1)\\ \bar{x}_i &= \text{row vector, known ground locations, where } i = 1, \ldots, N\\ c_i &= |\bar{x}_S - \bar{x}_i|\\ \Delta c_{ij} &= c_i - c_j, \text{ where } j = 2, \ldots, N\\ M &= 3 \times N \text{ matrix of } \bar{x}_i \text{ vectors} \end{aligned}$$

Letting $\bar{x}_1$ arbitrarily be the origin point for range difference computations, the following scalar and vector values are computed.

$$\begin{aligned} r_S &= |\bar{x}_S|\\ \bar{v} &= (r_2^2 - \Delta c_{21}^2, \ldots, r_N - \Delta c_{N1}^2)^T\\ \overline{\Delta c} &= (\Delta c_{21}, \ldots, \Delta c_{N1})^T \end{aligned}$$

The approximate least squares esumate according to Smith is given by:

$$\bar{x}_S = \frac{1}{2}(M^T M)^{-1} M^T (\bar{v} - 2 r_S \overline{\Delta c}) \qquad (2)$$

This is simply the generalized inverse solution to the minimum least squares value for $\bar{\varepsilon}^T \bar{\varepsilon}$, where $\bar{\varepsilon}$ is given by:

$$\begin{aligned} \varepsilon_i &= r_{i2} - \Delta c_{i1}^2 - 2 r_S \Delta c_{i1} - 2 x_i^T x_S, \; i = 2, \ldots, N\\ \bar{\varepsilon} &= (\varepsilon_2, \ldots, \varepsilon_N)^T \end{aligned}$$

This solution closely approximates the maximum likelihood solution for sensor location, and could be used as a starting estimate to iteratively compute the maximum likelihood solution to the nonlinear resection equations. Such a procedure is virtually certain to converge quickly with a good starting estimate given by the linear procedure above.

We can generalize this result to incorporate INS measurements, again assuming that an estimate in crossrange and altitude is required. Using the INS measurements, we compute the following values:

$$\begin{aligned} t_i &= \text{the time points corresponding to a range measurement with respect to an identified DEM position}\\ p(t_i) &= \text{the set of INS measurements for the time points } t_i, \text{ where } i = 1, \ldots, N\\ \bar{x}_o &= \text{the desired INS offset} \end{aligned}$$

Then modify one of the definitions of Eqn. (1) as follows:

$$\begin{aligned} c_i &= |p(t_i) - (\bar{x}_o - \bar{x}_i)|\\ &= |\bar{x}_o - (\bar{x}_i - p(t_i))| \end{aligned}$$

Therefore, one can use the generalized inverse solution (Eqn. (2)) to obtain the INS offset but with $\bar{x}_i$ replaced by $(\bar{x}_i - p(t_i))$.

The accuracy of this resection operation is limited by the resolution of the imagery, and the resolution of the terrain or feature database. The INS compensations to the SAR phase history are also assumed to be good such that the INS can

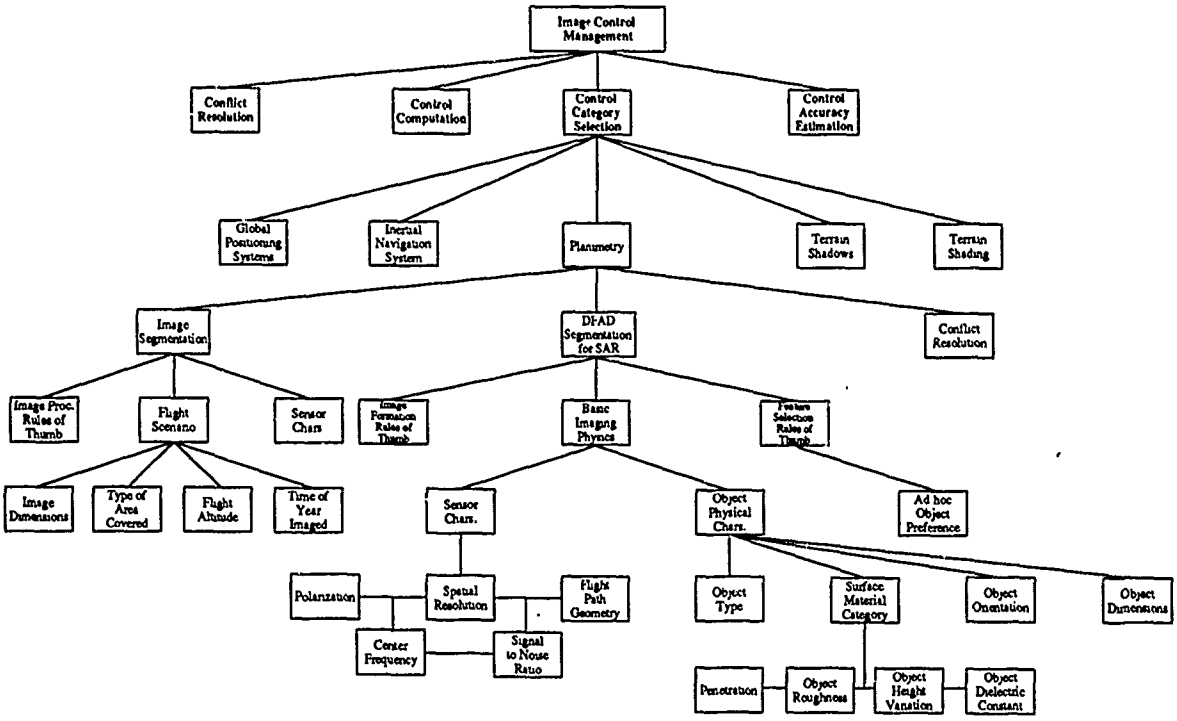


Figure 5: Hierarchical organization for rule base for feature segmentation.

be used as a basis for a more precise description of the platform flightpath. Also, in order for a resection operation to be convergent, the ground control points must be located in a relatively wide area in the image That is, a small cluster of points will lead to an ill-conditioned resection, and poor results.

The least squares solution given in Eqn. (2) above can be generalized to include a weighting matrix W, by replacing M^T by $M^T W$.

Expert Assistant for Feature Data Segmentation

A rule-based feature segmentation system would be only a part of an overall strategy for utilizing INS, GPS, and terrain induced shadow, and shading data, to determine platform position. Such a system would be somewhat misnamed as an "expert system", since an argument may be given that whenever one is operating strictly according to rules, one is actually operating at the level of an advanced beginner. A real expert, on the other hand, is able to create new rules based on experience, modify and find appropriate exceptions to existing rules, and make novel hypotheses. However, a reasonable and achievable rule-based system could be developed to carry out simple tasks and make routine choices. The output from such a system is potentially very helpful in average situations that do not require any extraordinary interpretation by a human expert in radargrammetry. We shall term such a system an "expert assistant".

In the case of feature segmentation for matching to a given radar image, the assistant would be required to choose initial features, choose appropriate image processing, and resolve conflicts throughout the process. It would also decide which procedural modules to execute, and would evaluate their outputs using its knowledge base. The rules making up the knowledge base must be based on the physics of the imaging scenario, as well as heuristics, and "rules of thumb" commonly employed by practicing experts. It is important to organize the rules in a hierarchical fashion. Figure (5) shows the hierarchical structure by which we have developed an initial rule base.

The subgoals of such a system would include weighting chosen features according to several criteria. These may include the expected visibility of the feature, and its distinguishability from others that are similar to it, as well as from its background. The predictability of the features appearance, and the complexity of the matching process should also figure into a composite score for the feature. Once all the criteria have been evaluated for each candidate feature, a rank list can be created of the most likely and productive features for a matching process.

Conclusions

We have demonstrated that digital terrain and feature data can be combined with automated image matching techniques to provide ground control information suitable for determining SAR sensor platform position using resection in space. This method finds local corrections for INS data which is subject to drift, and is applicable when GPS is not available.

The limitations of the resection accuracy are due to the image resolution, the faithfulness of the INS compensations into the SAR phase history processing, the accuracies of the features in the database, and the numerical stability of the resection geometry.

One of these methods uses image processing techniques to automatically extract and match radar shadows from a SAR image with the predicted shadows that are consistent with and estimated sensor location relative to a terrain model.

Efficient methods were also developed for locating arbitrary planimetric features in a SAR image for the purpose of matching with a digital feature database. These matched features were also used for resection. Present efforts now focus on the development of a rule-based expert assistant to aid in the selection of appropriate planimetric features for such matching.

References

1. J.O. Smith, and J.S. Abel, "Closed-form least-squares source location from range-difference measurements", *IEEE Transactions on Acoustics, Speech and Signal Processing, Vol. ASSP-35*, 12, 1987, 1661-1669.
2. Rosenfeld A., ed. *Multiresolution Image Processing and Analysis*, Springer Verlag, 1984.
3. D.I. Barnea, and H.F. Silverman, "A class of algorithms for fast digital image rectification", *IEEE Transactions on Computers, Vol. C-2*, 2, 1972, 179-186.
4. J.S. Lee, "Digital image smoothing and the sigma filter", *Computer Vision, Graphics, and Image Processing, Vol 24*, 1983, 255-269.
5. A.P. Ambler, H.G. Barrow, C.M. Brown, R.M. Burstall, and R.J. Popplestone, "A versatile system for computer controlled assembly", *Artificial Intelligence,Vol. 6.* 2, 1975, 129-156.
6. Pratt W., *Digital Image Processing*, John Wiley & Sons,1978.

USING MIMICS TO MODEL MICROWAVE BACKSCATTER FROM TREE CANOPIES

K. C. McDonald, M. C. Dobson and F. T. Ulaby

The Radiation Laboratory
Department of Electrical Engineering and Computer Science
The University of Michigan
Ann Arbor, MI 48109-2122, U.S.A.
Phone: (313) 764-0500
FAX: 313/936-3492
Telex: 432-0815 UOFM-UI

ABSTRACT

The Michigan Microwave Canopy Scattering Model (MIMICS) is used to model microwave scatterometer data that were obtained during the August 1987 EOS Simultaneity Experiment. During this experiment, L-, C- and X-band truck-based scatterometers were used to measure radar backscatter from a walnut orchard in Fresno County, California. Multi-polarization data were recorded for orchard plots of varying irrigation levels.

MIMICS, a scattering model based on radiative transfer theory, is applied to model various data sets recorded during the experiment. Groundtruth data are used as inputs to MIMICS and the resulting modeled data are compared to the measured backscatter.

Data examined in this study include a series of diurnal measurments in which a single orchard plot was observed continuously over several 24 hour periods, a multi-angle data set for which this same plot was observed at varying incidence angles, and a series of data recorded on two plots of different irrigation levels.

In modeling the canopy backscatter, MIMICS accounts for scattering contributions directly from the trees themselves, direct backscatter contributions from the underlying ground surface, and contributions resulting from interactions between the trees and the ground. The model is shown to account for variations in canopy backscatter that are driven by diurnal processes as well as by the differing irrigation levels. The distribution of branch sizes and orientations within the tree crowns is shown to be an important parameter for modeling canopy backscatter.

GLOBAL ALTIMETER MEASUREMENTS OF EXTREMES OF WIND SPEEDS AND WAVEHEIGHTS AND THEIR IMPACT ON FORECASTING

Ella B. Dobson
The Johns Hopkins University
Applied Physics Laboratory
Johns Hopkins Road
Laurel, Maryland 20707-6099
Telephone: (301) 953-5000, ext. 8645
FAX: 301-953-1093

Meteorologists have long known that weather patterns are cyclical in nature and in many cases these tracks can be accurately forecast within statistical accuracy. There are, however, many cases when accurate forecasting is not possible. More investigation of the spatial and temporal scales of intense weather patterns may lead to a better forecast method or at least augment the present systems.

This paper presents a study of extreme wind speeds and wave heights associated with intense weather patterns. Measurements of these parameters by the Geosat altimeter over a three year period are analyzed to determine spatial and temporal scales. Scales and movement in the southern oceans are compared to northern patterns. A discussion is presented on the potential usefulness of these statistics on forecasting. For example, in a region where Geosat data reveals a 80 percent probability of storm formation within a certain area, this region could be routinely monitored using the altimeter data. The question arises whether this would be valuable to the meteorological community.

RADIATIVE TRANSFER THEORY FOR ACTIVE REMOTE SENSING OF A FORESTED CANOPY

M A KARAM & A K FUNG

Wave Scattering Research Center
Dept. of Electrical Engineering , Box 19016
Unversity of Texas at Arlington
Arlington , Texas , 76019 . USA

ABSTRACT

A canopy is modeled as a two layered medium above a rough interface. The upper layer stands for the forest crown which is composed of leaves and branches. The leaves are modeled as randomly oriented and distributed discs and needles for deciduous and coniferous forest respectively. The branches are modeled as randomly oriented finite length dielectric cylinders. The lower layer contains the tree trunks which are modeled as randomly positioned vertical cylinders above the rough soil.

The radiative transfer theory is applied to calculate electromagnetic scattering from such a canopy. The total scattering coefficient is expressed in terms of the scattering amplitude tensors of leaves, branches, trunks and soil. For leaves the generalized Rayleigh Gans approximation is applied to obtain the scattering amplitude tensor. The branch and the trunk scattering amplitude tensors are obtained by estimating the inner field by fields inside a similar cylinder of infinite length. The Kirchhoff method is applied to calculate the soil scattering amplitude tensor.

For a plane wave exciting the canopy the radiative transfer equations are solved by iteration to the first order in albedo of the leaves and the branches. Numerical results are illustrated as a function of the incidence angle.

Keywords: Forest, leaves , Branches, Trunks, Soil

1. INTRODUCTION

A complete phenomenological model including the leaves ,the branches and the trunks was developed for forest stand (Richards et al, 1987). In this model only one polarization was considered (HH) and empirical formula for the backscattering coefficients of each forest component were used.

This study is aimed at developing a complete simple physical model for a forested canopy. In such a model the canopy is modeled by a two-layer medium above a rough interface as shown in Fig.1 The upper layer stands for the crown which is composed of leaves and branches. The lower layer stands for the trunks. The leaves will be modeled by discs and needles for deciduous and coniferous forest respectively. The branches will be modeled by finite length cylinders. The trunks will be modeled by vertical cylinders.

The backscattering coefficient from such a model is obtained by applying the first order solution of the radiative transfer equations (Karam and Fung, 1982). Then numerical calculations will be presented to illustrate the backscattering coefficient as a function of leaf type, orientation and volume fraction.

2 - PROBLEM FORMULATION

consider a plane wave exciting the forested canopy in $\hat{i}(\theta_i,\phi_i)$ direction and with $\hat{q}$ polarization ($q=\hat{v}_i,\hat{h}_i$). Upon solving the radiative transfer equations (Karam and Fung,1982), the first order backscattering coefficient from the canopy can be written as

$$\sigma_{pq}(\theta_i) = \sigma_{pq}(\theta_i, s)+\sigma_{pq}(\theta_i, c) +\sigma_{pq}(\theta_i, c\rightleftharpoons s)+\sigma_{pq}(\theta_i, t\rightleftharpoons s) \qquad (1)$$

Where the first term is due to surface scattering; the second term to the crown scattering ;the other two terms to the crown-surface and the trunk-surface scattering respectively.

3- SURFACE SCATTERING

The backscattering coefficient due to the surface scattering can be written as,

$$\sigma_{pq}(\theta_i,s) = L_{1p}(\theta_i).L_{2p}(\theta_i)R_{pq}L_{2q}(\theta_i).L_{1q}(\theta_i) \quad (2)$$

Where R_{pq} is the surface backscattering coefficient (Ulaby et al,1982). $L_{1q}(\theta_i)$ and $L_{2q}(\theta_i)$ are the crown and the trunk loss factors

$$L_{1p}(\theta_i) = e^{-[k_{1p}(\theta_i).H_1 \sec\theta_i]}$$

$$L_{2p}(\theta_i) = e^{-[k_{2p}(\theta_i).H_2 \sec\theta_i]} \quad (3)$$

and $k_{1p}(\theta_i)$ and $k_{2p}(\theta_i)$ are the crown and trunk extinction coefficients respectively (Karam and Fung,1989b)

4 - CROWN SCATTERING

The crown backscattering coefficient has the value

$$\sigma_{pq}(\theta_i) = 4\pi\cos\theta_i\left\{n_l\left\langle\left|F_{lpq}\right|^2\right\rangle + n_b\left\langle\left|F_{bpq}\right|^2\right\rangle\right\} \cdot\left\{\frac{1-L_{1p}(\theta_i).L_{1q}(\theta_i)}{k_{1p}(\theta_i)+k_{1q}(\theta_i)}\right\} \quad (4)$$

Where n_l and n_b are the leaf and the branch number densities. F_{lpq} and F_{bpq} are the scattering amplitude tensors for the leaves and the branches in the backward direction. The ensemble average< > is taken over the leaf and branch orientations.

5- THE CROWN - SURFACE INTERACTION

The crown-surface backscattering coefficient is

$$\sigma_{pq}(\theta_i, c\rightleftarrows s) = \sigma_{pq}(\theta_i, c\rightarrow s) + \sigma_{pq}(\theta_i, s\rightarrow c) \quad (5)$$

Where the first term is due to scattering from the crown followed by scattering from the surface. The second term is due to scattering from surface followed by scattering from the crown. The explicit contents of these terms are

$$\sigma_{pq}(\theta_i, c\rightarrow s) = 4\pi\cos\theta_i .L_{1p}(\theta_i).L_{2p}(\theta_i).R'_{pp}L_{2p}(\theta_i) \cdot\left\{n_l\left\langle\left|F'_{lpq}\right|^2\right\rangle + n_b\left\langle\left|F'_{bpq}\right|^2\right\rangle\right\} \cdot\left\{\frac{L_{1p}(\theta_i)-L_{1q}(\theta_i)}{k_{1q}(\theta_i)-k_{1p}(\theta_i)}\right\}$$

$$\sigma_{pq}(\theta_i, s\rightarrow c) = 4\pi\cos\theta_i\left\{n_l\left\langle\left|F'_{lpq}\right|^2\right\rangle + n_b\left\langle\left|F'_{bpq}\right|^2\right\rangle\right\} \cdot\left\{\frac{L_{1p}(\theta_i)-L_{1q}(\theta_i)}{k_{1q}(\theta_i)-k_{1p}(\theta_i)}\right\}L_{2q}(\theta_i) \cdot R'_{qq}.L_{1q}(\theta_i).L_{2q}(\theta_i) \quad (6)$$

Where R'_{pp} is Fresnel reflectivity and F'_{1pq} is the scattering amplitude in the specular direction (Karam and Fung, 1988 , 1989b) .

6- TRUNK - SURFACE INTERACTION

The trunk-surface backscattering coefficient is equal to

$$\sigma_{pq}(\theta_i, t\rightleftarrows s) = \sigma_{pq}(\theta_i, t\rightarrow s) + \sigma_{pq}(\theta_i, s\rightarrow t) \quad (7)$$

The first term is due to scattering from the trunk followed by scattering from the surface. The second term is due scattering from the surface followed by scattering from the the trunk. Since the trunks have dimensions larger than the incident wavelength the cross polarized component due to trunk - surface interaction will vanish and the like cross polarized components will have the value

$$\sigma_{pp}(\theta_i, t\rightarrow s) = \sigma_{pp}(\theta_i, s\rightarrow t) = 4\pi H_2 n_t\left[L_{1p}(\theta_i).L_{2p}(\theta_i)\right]^2 \cdot\left|F'_{tpp}\right|^2 . R'_{pp} \quad (8)$$

Where n_t is the trunk number density. F'_{tpp} is the trunk scattering amplitude in the specular direction.

7 - NUMERICAL RESULTS & DISCUSSION

This section is aimed at investigating the effect of the leaf-type, volume fraction and orientation on the backscattering coefficient. The Eulerian angles (α,β,γ) are used to describe the branch and the leaf orientations. The cylinders modeling the branches are identical and have radius "a" and length "2h". The branch dielectric constant is taken to be

$\varepsilon_b(\varepsilon_b=12.4-j4.9)$. The leaves are modeled by circular discs or needles with radius "a" and thickness "2t".The leaf dielectric constant has been given by Ulaby and El-Rayes(1988) in terms of the leaf moisture content($m_g=0.46$). In all figures the ground truth of the branches, the surface and the trunks are kept fixed (the branches : $v_{fb}=0.1x10^{-2}$, $0<\alpha<2\pi$, $0<\beta<\pi/6$, $\gamma=0$, a=0.025m, h=0.4m. the trunks : $n_t=0.11m^{-2}$, a=0.12m, the surface : $k\sigma=0.96$, kl=1.9. = H1=H2=10m, F=9Ghz). Figs 2 and 3 illustrate the backscattering coefficient dependence on the leaf orientation. The leaves are modeled by circular discs (a=0.02m,t=0.125mm, $v_{fl}=0.1x10^{-2}$). Figs. 4 and 5 illustrate the backscattering coefficient dependence on the leaf volume fraction. In these figures the needle-shaped leaves are used (a=1.25 mm, t=0.1m). The volume fraction used in Fig. 4 and Fig. 5 are $0.1x10^{-2}$ and $0.1x10^{-3}$ respectively. The difference in backscattering coefficient for leaves of different types can be recovered from Figs. 3 and 4.

8. CONCLUSION

The backscattering coefficient of a forested canopy is illustrated. It is seen that the backscattering coefficient is sensitive to the leaf-type, orientation and volume fraction.

ACKNOWLEDGMENT

This work is supported in part by the National Aeronautics and Space Administration under Grant NAG5-486 and in part by the Texas Advance Research Program under Grant No. 1861.

9- REFERENCES

1. Karam,M.A.,and A.K.Fung," Propagation and scattering in multi-layered random media with rough interface," Electromagnetic, Vol.2, pp 239-256, 1982.

2. Karam, M.A., and A.K. Fung,"Electromagnetic scattering from a layer of finite length, randomly oriented, dielectric,circular cylider over a rough interface with application to vegetation", Int. J. Remote Sensing,Vol.9, No.6, pp 1109-1134, 1988.

3. Karam, M.A. and A.K.Fung,"Leaf Shape effects in electromagnetic scattering from vegetation", IEEE, Trans. Geosci. Remote Sensing, Sent for Publication, 1989b.

4. Karam, M.A. and A.K.Fung, "Electromagnetic wave extinction within a forested canopy", Digest of Intr. Remote Sensing Symp. (IGARSS'89), This Issue 1989a.

5. Richard, J.A., G.Q. Sun, and D. S. Simonett," L-band radar backscatter modeling of forest stands", IEEE Trans. Geosci. Remote Sensing, Vol. GE-25, No.4, pp 487-498, 1987.

6. Ulaby, F.T., R.K.Moore and A.K.Fung, Microwave Remote Sensing, Active and Passive, Vol. II Addison-Wesley Publishing Company, Reading, MA. pp 936-942, 1982.

7. Ulaby, F.T. and M.A. El-Rayes," Microwave Dielectric spectrum of vegetation part II: Dual dispersion model", IEEE Trans. Geosci Remote Sensing, Vol.25, No.5, pp 550-557, 1987.

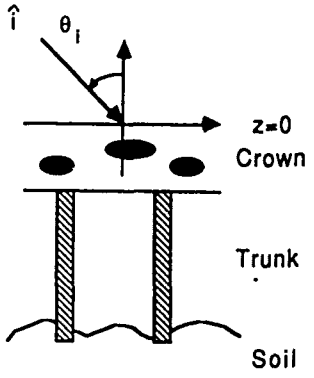


Fig.1 The problem configuration

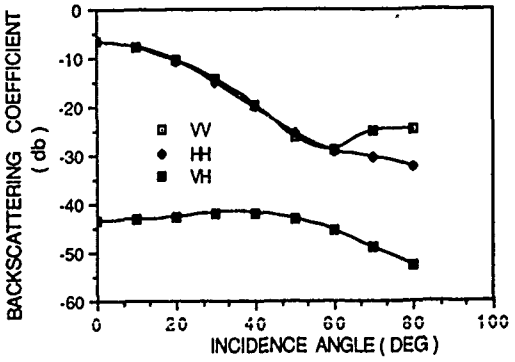


Fig.2 Scattering from a canopy with disc-shaped leaves vertically oriented

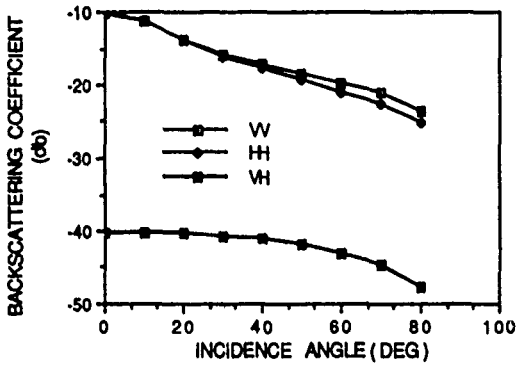


Fig.3 Scattering from a canopy with disc-shaped leaves uniformly oriented

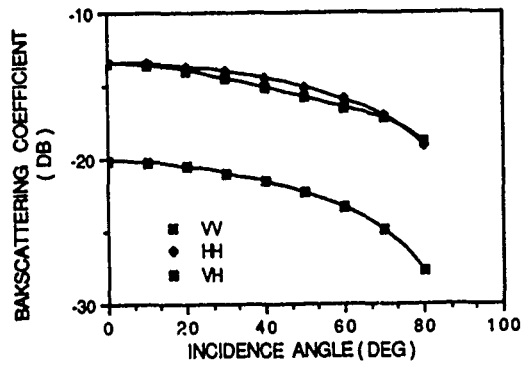


Fig.4 Scattering from a canopy with needle-shaped leaves uniformly oriented

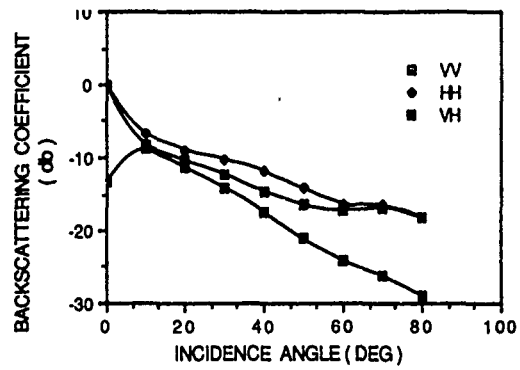


Fig.5 Scattering from a canopy with needle-shaped leaves uniformly oriented

RADAR MODELING OF TROPICAL MANGAL FOREST STANDS

Yong Wang*, Marc L. Imhoff**, and David S. Simonett*

*Department of Geography, University of California, Santa Barbara, CA 93106, U. S. A.
Tel: (805) 961 - 2013; FAX (805) 961 - 8016.
**NASA/Goddard Space Flight Center, Greenbelt, MD, 20771, U. S. A.
Tel: (301) 927 - 4975.

Abstract

Radar modeling of mangal forest stands, in the Sundarbans area of Southern Bangladesh, has been developed. The modeling employs radar system parameters with forest data on tree height, spacing, biomass, species combinations, and water (including slightly conductive water) content both in leaves and trunks of the mangal. For Sundri and Gewa tropical mangal forests, six model components are proposed, which are required to explain the contributions of various forest species combinations in the attenuation and scattering of mangal vegetated nonflooded or flooded surfaces. Statistical data of simulated images have been compared with those of SIR-B images both to refine the modeling procedures and to appropriately characterize the model output. The possibility of delineation of flooded or non-flooded boundaries is discussed.

Key words: radar modeling, mangal forest stands.

Essential Characteristics of Gewa and Sundri Forests

Gewa *(Excoecaria agallocha)* and Sundri *(Heritiera minor Syn H. formes)* are mangrove species. The shoots of both are mostly orthotropic with infrequent diffuse branches, which sometimes may branch from the base in a shrub-like form. Leaves are simple in shape, and broad in size (Tomlinson, 1986). Gewa lacks any elaborated aerial parts (pneumatophores), which function to supply oxygen to the root system at high tide. Sundri does have pneumatophores.

A mature Gewa tree: grows intermittently and irregularly, up to 20 meters high for a mature tree. Branching is diffuse, irregular, and by prolepsis. Canopy is dense (FAO, 1981).

A mature Sundri tree. is about 10 to 25 meters tall, with low-branches which are thick and crooked. Canopy is very dense (FAO, 1981). Illustrations of Gewa and Sundri (pp. 63 - 64 in Tomlinson, 1986) are of very similar form.

Gewa and Sundri forest stands occur essentially and dominantly at the tidal mouth of the Baleswar River in southern Bangladesh, and belong to closed broadleaved woody tropical rain forests (FAO, 1981). Most of the mouth areas are frequently flooded by fresh water during the monsoon rains (May - June to October - November, Chapman, 1977); with salt water intrusion during the rest of the year. Gewa and Sundri forests are very difficult of access with soft footing, and extremely high canopy and stem densities. These mangrove forests (estimated at about 590,000 hectares, FAO, 1982) contribute very important forest resources and environmental protection for Bangladesh: they provide marketable products, and also play roles as buffers for storm surges, and tidal waves coming from the Bay of Bengal (Imhoff et al., 1986).

Ground surfaces: the soil is a clay loam lying over alternating layers of clay and sand, highly saturated with exchangeable bases (C_a^{++}, K^+, N_a^+, etc.). The salinity of the soil varies yearly as a result of the changing balance of fresh and salt water. Chapman (1977) pointed out that there is a very high soil salinity at the beginning of the dry season, and this factor considerably limits the development, maintenance, or regeneration of the less halophilous species such as Sundri, which is on the way to extinction because of excessive salt water intrusion. During the rainy season (when the SIR-B image was obtained), the salinity will be lower.

Analysis of Ground Truth Data

Ground truth data on the Sundri and Gewa forest in this southern Bangladesh site were originally collected by a NASA and Bangladesh science team concurrently with the SIR-B Mission in October, 1984, and subsequently by Chaffey (1985). All these data were prepared by averaging measurements made for

inventory plots each 100 * 100 meters (one hectare) in size, consisting of: 1) a digitized forest stand map; 2) DBH count distributions by diameter segments within one hectare area; 3) tree heights and canopy depths by their DBHs; 4) canopy biomass in each DBH segment; and 5) gravimetric moisture content of canopies and trunks.

A digitized forest stand map containing different Sundri and Gewa stands was geo-registered with the SIR-B images, as shown in Table 1. These stands have different tree densities and combinations of species for each DBH segment. The ground surface is either flooded or non-flooded.

DBH count distributions by diameter segments per hectare are shown in Table 2. These are used to generate random tree trunk distributions with varying DBHs in a simulated stand area such as twenty hectares (assumed in our research). Based on both the combination of species and random locations of the tree trunks, tree counts for each DBH segment of Sundri and Gewa in each pixel in the stand area are created.

Tree heights and canopy depths by their DBHs are derived by regression, based on DBH, tree height, and canopy depth, as summarized in Table 3. These models have been employed to calculate tree heights and canopy depths in each simulated image pixel for the whole forest stand.

Mean canopy biomass per tree in each DBH segment is calculated by dividing the total biomass (kg) by the tree numbers in each DBH segment, as shown in Table 4. The biomass data is further converted to volume ratio (m^3 / m^3) (Ulaby, et al., 1986), which is an input parameter to compute attenuation coefficients for each simulated pixel.

Gravimetric moisture content of canopies and trunks are derived from dry biomass data and wet biomass data. Because of the lack of explicit data to separate Gewa and Sundri, we assume the water content in trunks and canopies to be the same for each species. The water contents are: canopy, μ = 0.402668, σ = 0.038955; and trunk, μ = 0.453775, σ = 0.057921. These data are used to compute the dielectric constants of the canopies and the trunks (Ulaby and El-Rayes, 1987).

Modeling of Sundri and Gewa Forest Stands*

Modeling components of Sundri and Gewa forest stands. six model components are proposed (see Figure 1):

1 Direct backscattering from ground surface (short form σ^o_{dbs}).

2 Direct volume scattering from upper canopy (short form σ^o_{vsu}),

3 Direct volume scattering from lower canopy (σ^o_{vsl}),

4 Interaction of trunk/ground forward reflection (σ^o_{itg}),

5 Interaction of Ground/upper canopy forward scattering (σ^o_{guc}), and

6 Interaction of ground/lower canopy forward scattering (σ^o_{glc}).

The total radar return is

$$\sigma^o_{tot} = \sigma^o_{dbs} + \sigma^o_{vsu} + \sigma^o_{vsl} + \sigma^o_{itg} + \sigma^o_{guc} + \sigma^o_{glc}$$

σ^o_{dbs}: the small perturbation model (Dobson and Ulaby, 1986) is used for both flooded and nonflooded surfaces, whose roughness parameters (roughness heights and correlation lengths) are assumed (no measurement data available) as h_{sea} = 0.01m, l_{sea} = 0.15m; h_{sur} = 0.02m, and l_{sur} = 0.075m, respectively. The dielectric constants of the flooded surface is assumed to be equivalent to 1 % salinity of water, because of extensive [illegible]ation during the monsoonal rainy season. The dielectric constant of the ground surface is also derived from the mixed model of Ulaby, et al., (1986).

σ^o_{vsu}: the water cloud model (Attema and Ulaby, 1978) is chosen. this model is simple and functional, and no measurement data on orientations of branches or leaves are available. The extinction coefficient is calculated for each pixel. The ratio of the volume scattering coefficient to the extinction coefficient is assumed as 0.3, and 0.4 for the upper and the lower canopy layers respectively, because the lower canopy is denser than the upper canopy. These values are higher than those used by Sun and Simonett (1988): the mangal forest canopy is much denser than the pine forests of the northern California; and Gewa especially may also have a slight salinity inside leaves and branches (Snedaker, personal communication, April, 1989). In addition, the sensitivity of the model to change in the ratio is low as shown by Richards, et al. (1987).

σ^o_{vsl}: basically is the same model as σ^o_{vsu}, but with different densities.

σ^o_{itg}: is the specular scattering coefficient of dielectric cylinders with finite lengths (Ruck, et al., 1970) multiplied by the reflection coefficient of the surfaces (flooded or nonflooded). The dielectric constants of the Gewa and Sundri trunks are about (22.0 - j 7.0); (18.5 - j 5.0), which are derived from the model developed by Ulaby and El-Rayes (1987) with gravimetric moisture content, equivalent salinity, and temperature as input. A total trunk height is calculated from the regression model, and the effective trunk length or the dielectric cylinder length is the total length minus two thirds of its canopy depth. Also, for each σ^o_{itg}, coherent summation of each is performed. The phase produced by range difference is taken into account.

σ^o_{guc}: the same model as used by Sun and Simonett (1988).

σ^o_{glc}: a similar model as above, except for different densi-

*: Upper and lower canopies are defined as canopies of trees whose DBH are respectively >= 5.0 cm or < 5.0 cm, the latter is also termed regrowth.

ties, mean canopy depth, and equivalent dielectric constants.

All these model components may or may not be attenuated by the upper and/or the lower canopies, depending on their presence or not.

Discussion

The subareas of three forest stands (Table 1) are extracted from SIR-B images (DT 120 incidence angle 26°; and DT 104 incidence angle 46°; see also Imhoff, et al., 1986), and the DNs of the images converted to relative radar backscattering coefficients (dB) by

$$\sigma^o = 10 * log(DN^o - C_{1i}) - C_{2i}$$

where i = 1, 2, corresponding to DT 120 and DT 104. C_{1i} are the noise levels, C_{2i} constants for calibration. Both C_{1i} and C_{2i} were provided by JPL.

The boxplots of SIR-B images are shown in Figures 2 (DT 120) and 3 (DT 104), with the modeling results plotted to the right. From these Figures, we conclude that:

The fit of the models to the SIR-B images: the distributions of radar returns for each stand and two incidence angles are almost identical. The fit of the model to the image data is good. There are differences in the absolute values (the returns from our model is about 5 to 8 dB higher than the returns of SIR-B images). The reason may be that the models are theoretical solutions and each model component is incoherently added, which could produce higher values.

Radar returns of nonflooded and flooded areas: at 26° and 46° degree incidence angles, the returns (for both SIR-B images and modeling results) from flooded surfaces are slightly higher than those from nonflooded surfaces. However, the difference appears to be slight because of strong attenuation by the dense canopies. The return difference from large to small of each stand is in the order of GS (highest), G, and SG (lowest) stands. There is insufficient data in Table 2 to account for these differences.

Differences in radar returns between different stands: there are no obvious differences. We reason as follows: a) each stand is a mix of Sundri and Gewa in varying proportions; b) the number of stems is very large, which is of importance in both enhancing the σ^o_{ug} (because of the large number of trunks), and in attenuation (because of the very dense canopy); and c) the structure of Sundri and Gewa are similar to each other (Tomlinson, 1986). The returns from SG stands may be the lowest because the density of the canopy layer may be marginally greater than that of other canopies (see Table 2).

Radar returns between two different incidence angles: the mean (or median) are almost the same, but with the greater variance at the smaller incidence angle for which there is a shorter path length through the canopy, yielding less attenuation for the σ^o_{ug} term. The σ^o_{ug} is the main component producing the variance for mangal: a similar conclusion was reached by Sun and Simonett (1988) for pine forest.

Penetration of the canopy and delineation of flooded areas: there is penetration to some degree, but it is not very obvious. The difference of radar returns between the flooded and nonflooded areas of both the SIR-B images and modeling results are around 0.5 - 2.0 dB. This is probably not large enough to delineate the flooded boundaries with confidence, because of the high intrinsic variance in SIR-B images.

Relative importance of six model components: the dominant model components are σ^o_{ug}, σ^o_{vsu}, and σ^o_{vsl}, which are roughly equal for nonflooded surfaces. For flooded surfaces σ^o_{ug} is increased, and σ^o_{vsu} and σ^o_{vul} are unchanged. In areas where there are some natural clearings which lack the dense canopy layer, σ^o_{ug} is clearly dominant. We especially note the high returns from those forest edges which face the radar illumination direction. This feature has also been widely observed in short wavelength images of forest edges.

References

Attema, C. F. and F. T. Ulaby, "Vegetation modeled as a water cloud", Radio Science, vol. 13, pp. 357 - 364, 1978.

Chaffey, D. R., F. R. Miller, and J. H. Sandom, "A forest inventory of the Sundarbans, Bangladesh", British Overseas Development Administration Project # 140, Land Resource Development Center, Tolworth, Surbiton, Surrey, England, 1985.

Chapman, V. J. "Wet coastal ecosystems, Ecosystems of the world 1", Elsevier Scientific Publication Company, 1977.

Dobson, M. C. and F. T. Ulaby, "Preliminary evaluation of the SIR-B response to soil moisture, surface roughness, and crop canopy cover", IEEE Transactions on Geoscience and Remote Sensing, vol. GE-24, no. 4, pp. 517-526, 1986.

FAO, Food and Agriculture Organization of the United Nations, "Tropical resources assessment project (in the framework of the global environment monitoring system - GEMS)", Forest resources of tropical Asia, 1981.

FAO, Food and Agriculture Organization of the United Nations, "Management and utilization of mangroves in Asia and the Pacific", 1982.

Imhoff, M. L., M. Story, C. Vermillion, F. Khan, and F. Polcyn, "Forest canopy characterization and vegetation penetration assessment with space-borne radar", IEEE Transactions on Geoscience and Remote Sensing, vol. GE-24, no. 4, pp. 535-542, 1986.

Richards, J. A., G. Q. Sun and D. S. Simonett, "L-band radar backscatter modeling of forest stands", IEEE Transactions on Geoscience and Remote Sensing, vol. GE-25, no. 4, pp. 487-498, 1987.

Ruck, G. T., D. E. Barrick, W. D. Stuart, and C. K. Krichbaum, "Radar cross section handbook", Plenum Press, New York, 1970.

Sun, G. Q. and D. S. Simonett, "Simulation of L-band HH radar backscatter from coniferous forest stands: a comparison with SIR-B data", Int. J. of Remote Sensing., vol. 9, no. 5, pp. 907 - 925, 1988.

Tomlinson, P. B., "The botany of mangroves", Cambridge University Press, 1986.

Ulaby, F. T., R. K. Moore, and A. K. Fung, "Microwave Remote Sensing, vol. III: From theory to application", Artech House, 1986.

Ulaby, F. T and M. A. El-Rayes, "Microwave dielectric spectrum of vegetation - Part II: dual - dispersion model", IEEE Transactions on Geoscience and Remote Sensing, vol. GE-25, no. 5, pp. 550-557, 1987.

Table 1. Stand samples extracted from SIR-B images

Stands	Ground conditions	Samples (line * sample)	Total pixels
SGn	nonflooded	(10 * 10), (10 * 10), (11 * 6)	266
SGf	flooded	(4 * 12), (10 * 10), (12 * 12)	292
GSn	nonflooded	(7 * 20), (18 * 9)	302
GSf	flooded	(17 * 19), (5 * 12)	383
Gn	nonflooded	(10 * 10), (10 * 10), (10 * 5)	250
Gf	flooded	(7 * 25), (16 * 7)	287

SGn is a stand with Sundri dominant and with Gewa present for a nonflooded surface; GSn has Gewa dominant and with Sundri present on nonflooded surfaces; and so on (See also Table 2 for stand characteristics).

Table 2. DBH (cm) counts of each stand per hectare

Stands	Species	< 4.9	5.-9.9	10.-14.9	15.-19.9	20.-24.9	25.-29.9	30.-34.9	35.-39.9
SG	Sundri	42560	816	319	125	52	17	2	1
	Gewa	1340	493	250	72	16	5	0	0
GS	Sundri	42560	502	184	78	21	6	1	1
	Gewa	1340	1084	458	98	15	2	0	1
G	Sundri	42560	142	34	16	8	2	0	3
	Gewa	1340	1769	575	136	26	2	5	4

Source: Chaffey, et al., (1985).

Table 4. Mean canopy biomass (kg) per tree for each DBH segment

DBH segment cm	Species Gewa	Sundri
< 4.9	2.9	2.6
5.0 - 9.9	6.1	13.0
10.0 - 14.9	12.9	39.8
15.0 - 19.9	27.6	74.9
20.0 - 24.9	58.8	86.3
25.0 - 29.9	125.4	86.3*
30.0 - 34.9	267.5	86.3*
35.0 - 39.9	393.5	86.3*

*: Two sources of data were used to construct this table. In the field *canopy* biomass measurements were made by a NASA and Bangladesh Science Team in October, 1984. For that data, there were no Sundri trees whose DBHs were over 25.0 cm. The second source of data was Chaffey, et al., (1985), for the same area. They mapped the area into segments of different DBHs, including DBH's exceeding 24.9 cm. However, they did not derive any measurements of canopy biomass. On merging the two data sets, the three highest DBH segments, lacked canopy biomass data. The same values as for 20.0 - 24.9 cm were therefore employed.

Table 3. Regression models

Height (m) on diameter (cm) for Gewa

	Regression models	R^2 value
Gewa	Hg = 5.549 + 0.258 * Dg	0.505
Sundri	Hs = 4.769 + 0.414 * Ds	0.672

Canopy thickness (m) on diameters (cm)

	Regression models	R^2 value
Gewa	Cg = 1.840 + 0.188 * Dg	0.679
Sundri	Cs = 1.768 + 0.331 * Ds	0.589

Fig. 1. Model components of Sundri and Gewa forest stands

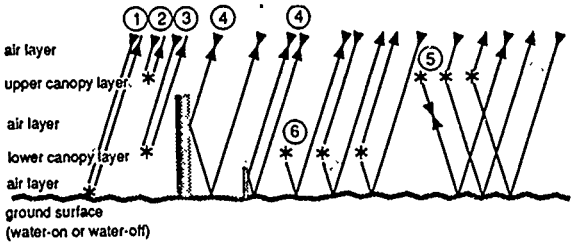
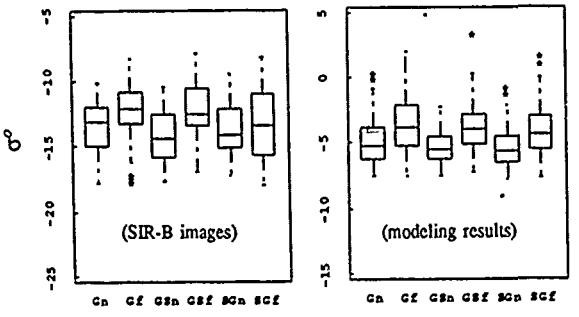
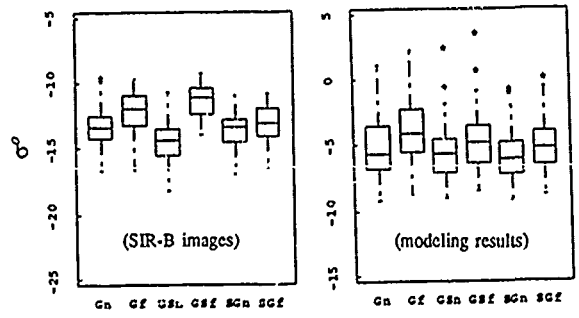


Fig. 2. Boxplots of forest stands for DT 120



Flooded and/or nonflooded surfaces

Fig. 3. Boxplots of forest stands for DT 104



Flooded and/or nonflooded surfaces

X-BAND BACKSCATTER MODELING OF ORCHARD CANOPY

N.S. Chauhan

The George Washington University
Washington D.C. 20052 USA

R.H. Lang[1]

NASA Goddard Space Flight Center
Greenbelt, Maryland 20771 USA

A microwave backscattering model in the X-band frequency regime has been developed for a tree canopy. The branches and tree trunks have been replaced by randomly oriented, finite, lossy, homogeneous dielectric cylinders. The leaves are represented by flat dielectric discs whose area and orientation distributions are prescribed.

The vector radiative transfer equations are used to calculate the backscattering coefficients as a function of incident angle. The phase and extinction matrices are computed by averaging the scattering amplitude of trunks, branches and leaves over their size and orientation distributions. The scattering amplitudes have been determined by using a physical optics approximation. For high frequency waves, the scattering amplitude of the disc peaks in the forward and reflected directions. In the case of a cylinder the scattering peaks in a cone around the cylinder. These peaked scattering amplitudes are integrated over their size and orientations to get the phase and the extinction matrices. Because of the peaked nature of the scattering amplitudes, the evaluation of the averages is simplified. The resulting phase matrix has a peak in the forward direction. Finally, including the peaked phase matrix in the vector transport equations, the backscattering coefficients for the canopy are obtained by following eigenvalue technique to solve the transport equations.

The theoretical backscattering calculations have been performed for both horizontal and vertical polarizations for realistic ground truth and canopy geometry parameters. Unlike at L-band, the leaves at X-band play an important role as a scatterer. For trees having a reasonable value of leaf area index (LAI), the leaves give a significant contribution to the backscattering. These results concur with the seasonal variations in backscatter observed by many researchers for deciduous trees. Also, as the angle of incidence becomes large, the skin depth decreases and this is reflected in the behavior of the backscattering coefficient. The computed and experimental values of backscattering coefficients at X-band show a reasonable agreement.

[1] On sabbatical leave from George Washington University.

ELECTROMAGNETIC WAVE EXTINCTION WITHIN A FORESTED CANOPY

M A KARAM & A K FUNG

Wave Scattering Research Center
Dept. of Electrical Engineering , Box 19016
Unversity of Texas at Arlington
Arlington , Texas , 76019 . USA

ABSTRACT

A forested canopy is modeled by a collection of randomly oriented finite length cylinders shaded by randomly oriented and distributed disc-shaped or needle-shaped leaves. For a plane wave exciting the forested canopy , the extinction coefficient for such a forested canopy is formulated in terms of the extinction cross sections in the local frame of each forest component and the orientation angles. The Eulerian angles of orientation are used to describe the orientation of each forest component.

For the finite length cylinders used to model the branches, their extinction cross sections in the local frame are obtained by employing the forward scattering theorem. For the disc-shaped and needle-shaped leaves their extinction cross sections in the local frame are obtained by the summation of the absorption and the scattering cross sections.

The behaviors of the extinction coefficients with the incidence angle are investigated numerically for both deciduous and coniferous forest. The dependencies of the extinction coefficients on the orientation of the leaves are illustrated numerically.

Key words: Branches, leaves, Extinction loss.

1 - INTRODUCTION

In a previous study (Karam and Fung,1989a) mathematical models had been developed to investigate the role of each forest components on the extinction of the electromagnetic wave in the microwave band. It was found that the finite-length cylinder model for the branches along with the forward scattering theorem gives a good estimate for the branch extinction loss. As to the leaves needle-shaped and the disc-shaped models along with the quasi-static approximation for the inner field can provide accurate extinction losses by summing the scattering and the absorption losses.

This study is aimed at developing a model for the extinction properties of a forested canopy. The canopy will be modeled as a collection of randomly oriented, finite-length cylinders for the branches and randomly oriented needles or discs for the leaves. Since the leaves and the branches are sparsely distributed Van de Hulst's formulation (Van de Hulst ,1957) will be used to obtain the canopy extinction coefficient in terms of the extinction cross section of the leaves and the branches. The branch extinction cross section will be formulated in terms of a branch local scattering amplitude in the forward direction. The leaf extinction cross section will be formulated as the sum of its scattering and absorption cross sections in the local frame.

Finally, numerical calculations will be presented to illustrate the extinction coefficient behaviors with respect to the leaf type , orientation, and the incident angle.

2 - PROBLEM FORMULATION

Consider a forested canopy which is composed of sparsely distributed leaves and branches. According to Van de Hulst (1957) the extinction coefficient is the sum of the extinction cross section of the individual leaves and the branches

per unit volume. For a plane wave exciting the canopy with q polarization ($q=v_i,h_i$) , the extinction coefficient is

$$\kappa_q(i) = n_b k_{bq}(i) + n_l k_{lq}(i) \qquad (1)$$

Where i is the incident direction ; n_b and n_l are the number of the branches and the leaves per unit volume. $k_{bq}(i)$ and $k_{lq}(i)$ are the branch and the leaf cross sections in the reference frame.

3 - THE BRANCH EXTINCTION CROSS SECTION

The branch extinction cross section can be obtained by applying the forward scattering theorem (Karam and Fung ,1982) yielding

$$k_{bq}(i) = \frac{4\pi}{k} \mathrm{Im}\left[\sum_{p=v,h} F'_{pq}(i,i) . \left(p'_i . q_i\right)^2 \right] \qquad (2)$$

In (2) , Im is the imaginary part operator , k is the background medium wave number, $F'_{pq}(i,i)$ is the branch local scattering amplitude in the forward direction and p'_i ($p'_i = v'_i$, h'_i) is the incident polarization vector in the local frame. Its relation with the orientation angle and the polarization vectors in the reference frame has been given by Karam and Fung (1989b). From (2) , the extinction cross sections for the vertical and horizontal polarization are

$$k_{bv}(i) = \frac{4\pi}{k} \mathrm{Im}\left[F'_{vv}(i,i) . (v'_i . v_i)^2 + F'_{hh}(i,i) . (h'_i . v_i)^2 \right]$$

$$k_{bh}(i) = \frac{4\pi}{k} \mathrm{Im}\left[F'_{vv}(i,i) . (v'_i . h_i)^2 + F'_{hh}(i,i) . (h'_i . h_i)^2 \right] \qquad (3)$$

The values of the scattering amplitudes in the forward direction are available from Karam et al(1988).

4 - THE LEAF EXTINCTION CROSS SECTION

The leaf extinction cross section can be obtained by summing the scattering $k_{sq}(i)$ and the ohmic $k_{aq}(i)$ cross sections as

$$k_{lq}(i) = k_{sq}(i) + k_{aq}(i) \qquad (4)$$

with

$$k_{sq}(i) = k'_{sv}(i) . \left[v'_i . q_i\right]^2 + k'_{sh}(i) . \left[h'_i . q_i\right]^2$$

$$k_{aq}(i) = k'_{av}(i) . \left[v'_i . q_i\right]^2 + k'_{ah}(i) . \left[h'_i . q_i\right]^2 \qquad (5)$$

$k'_{sq}(i)$ and $k'_{aq}(i)$ are the scattering and the absorption cross sections in the leaf local frame for q polarization ($q = v_i,h_i$) (Karam and Fung , 1989a)

5 - NUMERICAL RESULTS AND DISCUSSION

This section is aimed at investigating the effect of each forest component on the canopy extinction coefficient. The Eulerian angles (α ,β ,γ) are used to describe the branch and the leaf orientation (Fig.1). The cylinders modeling the branches are identical and have volume fraction v_{fb} ($v_{fb}=0.1\times10^{-2}$) , radius " a "(a=0.25m) , length "2h "(h=0.4m) and dielectric constant ε_b (ε_b =12.4-j 4.9). The leaves are modeled by circular discs or needles with volume fraction v_{fl}($v_{fl}=0.1\times10^{-2}$) , radius " a " and thickness "2t".
For discs we let a= 0.02m , t= 0.125mm ; and for needles we let a=1.25 mm , t=0.1m. The leaf dielectric constant has been given by Ulaby and El- Rayes (1988) in terms of the leaf moisture content (m_g=0.46). In all the following figures , the branches and the leaves are assumed uniformly oriented so that $0< \alpha < 2\pi$ and $\gamma=0$. Fig. 2 illustrates the branch extinction coefficient behavior versus the incident angle. The branches are vertically oriented ($0^0 < \beta < 30^0$). It is clear that both polarizations have the same value for the extinction coefficient. By increasing the incidence angle , the extinction coefficient will increase but its value for horizontal polarization will be lower than its value for vertical polarization. The dependence of the extinction coefficient on leaf type is illustrated in Figs. 3 and 4. In Fig. 3 the disc-shaped leaves ($0^0< \beta < 90^0$) and cylindrical branches are considered. In

Fig. 4 the needle-shaped leaves with the same orientation , volume and volume fraction as the disc-shaped leaves are considered. Both the relative level and the trend of the extinction coefficient depend on the leaf type. In Fig. 5 disc-shaped leaves vertically oriented ($60^o < \beta < 90^o$) are added to the branches to illustrate the variation of the extinction coefficient with the leaf orientation.

6 - CONCLUSION

The dependence of the level and the trend of the extinction coefficient on each forest component is illustrated. It is seen that the extinction coefficient is sensitive to the leaf type and leaf orientation.

ACKNOWLEDGMENT

This work was supported in part by the National Aeronautics and Space Administration under Grant NAG5-486 and in part by Texas Advance Research Program under Grant No. 1861.

7- REFERENCES

1- Karam, M. A. and A. K. Fung, "Vector forward scattering Theorem," Radio Science . vol. 18 ,pp 557-565, 1982.

2- Karam, M.A. and A.K.Fung, "The extinction properties of forest components", Int. J. Remote Sensing , Submitted for Publication, 1989a.

3- Karam , M.A. and A.K.Fung, " Leaf shape effects in elecrtomagnetic wave scattering from vegetation", IEEE Trans Geoscience and Remote Sensing Submitted for Publication, 1989 b.

4- Karam, M.A. :A.K.Fung, and Y.M.M.Antar," Elecromagnetic wave scattering from some vegetation samples," IEEE Trans. Geosci. and Remote Sensing , Vol. 26, No.6, pp 799-808 ,1988.

5- Ulaby F.T. , and M.A. El-Rayes, " Microwave dielectric spectrum of vegetation -part II: Dual dispersion model," IEEE Trans. Geocsi. Remote Sensing , vol. 25 ,No. 5 ,pp550-557, 1987.

6- Van de Hulst, H.C.,Light Scattering by Small Particles, Dover Publications , Inc. New York pp32-34, 1957.

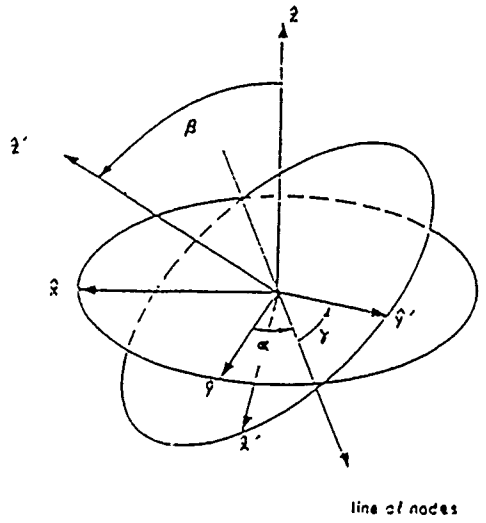


Fig. 1 The Eulerian angles of orientation

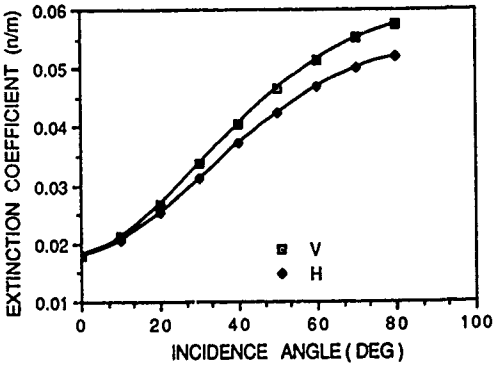


Fig.2 The extinction coefficient for branches

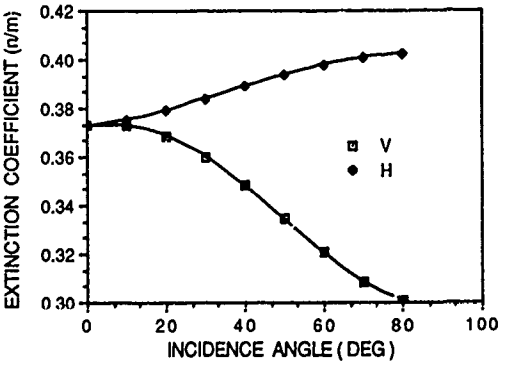


Fig.3 The extinction coefficient for branches and disc-shaped leaves

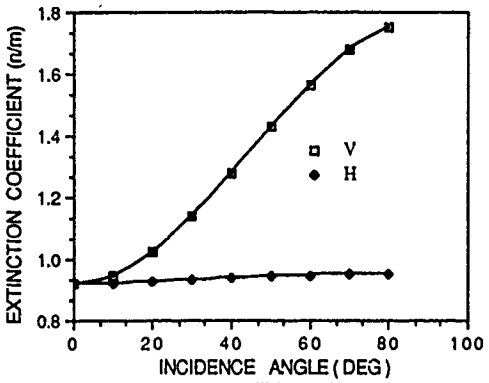


Fig.4 The extinction coefficient for branches and needle-shaped leaves

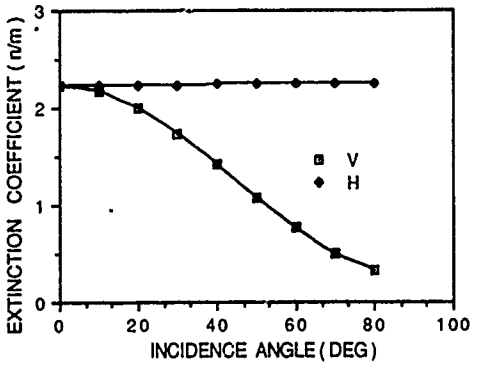


Fig.5 The extinction coefficient for branches and disc-shaped leaves

SPECTRAL REFLECTANCE PROPERTIES OF LICHENS: REMOTE SENSING DISCRIMINATION OF LICHEN AND ROCK

Edward A. Cloutis
Department of Geology
University of Alberta
Edmonton, Alberta T6G 2E3
Tel: (403) 492 3265

The Canadian Shield and other high-latitude regions are ideally suited for mapping and exploration by remote sensing means. To date, the laboratory spectral information necessary for effective analysis of remote sensing data from cold desert regions has been lacking. An examination of the spectral reflectance properties of a range of materials common to the Canadian Shield was undertaken in order to provide an initial data base for more effective high-latitude and alpine spectral remote sensing.

The focus of the project has been on geological materials, but lichen can cover significant portions of outcrops and are the most common obscuring material in these regions. The laboratory spectral reflectance properties of a number of lichen were measured for comparison with adjacent, lichen-free, weathered rock spectra. Eventually it is hoped to be able to separate the rock and lichen spectral signatures and individually map the distributions of different rock units and lichen species by remote sensing means.

Lichen spectra, regardless of type, have a number of features in common. The abrupt increase in reflectance in the visible-near infrared seen in vascular plants ("red edge") is very suppressed in lichen spectra, and appears as a gradual rise in reflectance towards the near infrared. The shortest wavelength absorption band is a distinct band or shoulder at ~0.68μm due to chlorophyll absorption. A more definable absorption feature is a broad, ubiquitous band near 1.45μm, probably due to the water present in the lichen. Its wavelength position and shape differs markedly from a similar band in weathered and unweathered rock spectra. A band of variable strength is seen near 1.7μm, and has previously been identified as combinations and overtones of various C-H stretching fundamentals. It is present in all lichen, the vast majority of biological organism, and fossilized hydrocarbon spectra.

Beyond 1.7μ are a number of absorption bands superimposed on a gradual decrease in reflectance. These bands are present near 1.9, 2.1, 2.3, and 2.5μm. The 1.9μm band is due to water and is also seen in hydrated minerals such as clays, and hence is not uniquely diagnostic of lichen, although high spectral resolution data can differentiate them. The bands at 2.1, 2.3 and 2.5μm are cellulose and other organic absorption bands. The wavelength positions of these bands do not vary among different lichens. A number of minerals also have absorption bands in the 2.1-2.5μm region, most notably carbonates and clays, but the shapes and positions of the mineral bands do not exactly coincide with those of the lichen and hence are potentially resolvable.

Lichen spectra are sufficiently unique that identification of lichen by remote sensing means is feasible, given adequate spectral resolution. The 1.7μm absorption band in particular may prove useful in determining areal lichen coverages because it is not significantly affected by common mineral absorption bands. The higher wavelength cellulose absorption bands may also prove useful for lichen discrimination, but interferences by mineral bands will require the application of sophisticated spectral deconvolution techniques. Discrimination of different lichen types may be possible using features in the visible spectral region, although the effects of mineral absorptions may be problematic. The preliminary results are promising enough to suggest that quantitative geological/biological remote sensing mapping of high latitude and alpine regions is an attainable goal.

Acknowledgements: This project was supported by Grants-in-Aid from the University of Alberta Boreal Institute for Northern Studies, and the Geological Society of America (#3741-87).

CONTRIBUTION OF 0.57μm/0.88μm RATIO TECHNIQUE IN GEOBOTANICAL STUDIES

A. P. Cracknell and A. K. Saraf
Department of Applied Physics and Electronic & Manufacturing Engineering,
University of Dundee
DUNDEE DD1 4HN, U. K.

ABSTRACT

Previous investigations have successfully used the near-IR/red ratio to map the concentration of hydroxyl-bearing minerals in imagery of sparsely-vegetated terrain and semi-arid regions where there is a wide variation in vegetation density. However, difficulties arise in applying this near-IR/red ratio particularly for geological mapping in tropical regions where the area is masked by vegetation. For areas which are covered by vegetation and where rock outcrops are rare, the standard image processing techniques do not improve images for geological interpretation. A disadvantage of ratioing, related to its ironing-out of topographic effects, is that it suppresses differences in albedo. However, this 0.57μm/0.88μm ratio technique preserves the topographic details and has proved useful in geological studies in a vegetated terrain. In the present study band 3/band 7 (0.57μm/0.88μm) data from the Airborne Thematic Mapper (ATM) of the Assynt district, in north-west Scotland, is used in geobotanical studies and has enabled us to improve the existing geological maps. In this study this ratioing technique has successfully been applied in the identification of geobotanical relationship.

Keywords: Near-IR/red ratio, band 3/band 7 ratio, Airborne Thematic Mapper, geobotanical relations.

1. INTRODUCTION

Band ratioing is a standard image processing technique that suppresses spatial radiance variations that are proportionally constant between bands and are generally attributable to terrain illumination, ground albedo and look-angle effects. Band ratioing commonly thereby enhances spatial radiance variations that are not proportionally constant between bands and are generally more informative in terms of surface composition (Crippen 1988). This paper explains the band 3/band 7 ratioing technique of ATM data and discusses its geological applications.

The most important property of a ratio image is that features in the spectral signature curve of a particular surface material are accentuated. If the bands used are chosen to cover peaks, absorption troughs and changes in slope on the curve, then they can be combined in pairs as ratios to express aspects of the spectral signature of the material. The simplest example relates to vegetation. Plant spectra are characterised by a combination of the chlorophyll absorption feature in the red part of the spectrum and the strong reflection of infrared in the 0.7-1.2 μm range by living cells. In images of Landsat MSS band 5 and band 7 the pixels containing a high proportion of plants will appear dark and light, respectively. In a standard false-colour composite image they will stand out from bare soil or rock as strongly red coloured areas. Within these fairly obvious areas may be all sorts of variations. The proportion of a single species relative to soil may change from place to place. Soil moisture or other environmental variations may affect the health and spectral properties of the species. In images expressing only data from various wavebands these differences are often concealed, simply because the human eye has a limited range of perception. However, the effect of the surface variations is to change the structure of the spectral reflectance curve. In particular, the depth of the chlorophyll absorption feature, the peak height of infrared reflectance and the gradient of the curve between red and infrared will change. Very small changes in these parameters have a disproportionately large effect on the ratio of band 7 to band 5, so they become easier to evaluate on the ratio image. Ratio images of different kinds are used in vegetation studies to estimate the proportion of leaf cover and to help in classifying and discriminating between different species (Drury 1987). But when band ratio techniques are being applied in

geobotanical studies, where the emphasis is on the identification of the relation between vegetation and bedrock geology, then the vegetation index provided by the infrared/red ratio image does not provide an opportunity for the identification of geobotanical information. However, in the present study the red/infrared ratio images have been found more useful for the identification of geobotanical information.

Plants depend on the underlying soils and rocks for their supply of water and nutrients. In their natural state they are sometimes excellent indicators of geology in humid terrains. This may be due to certain rocks providing favourable conditions for growth, but in some cases the substrate actually threatens the health of plants. This may be due to anomalously high concentrations of toxic elements, as might be the case in the vicinity of certain mineral deposits. More commonly, however, stressed vegetation is merely a response to either waterlogging or a shortage of water, in which case it is generally a seasonal phenomenon. Different plants have different tolerances to environmental stress, so another geologically related feature of vegetation is the development of plant communities specific to particular geological conditions. Both types of biological indicator can be examined most successfully using ratio techniques (Drury 1987).

To demonstrate the utility of this visible/infrared ratio technique an area of part of the Assynt district, north-west Scotland, has been selected.

2. THE STUDY AREA

In the Ordnance Survey National Grid the Inchnadamph area comes on sheet number 15 (Scale 1:50000) between 223000 to 229000 easting and 919000 to 924000 northing. Apart from small patches of bare outcrop and active screes, the area is completely vegetated, the dominant cover being a mixture of ericaceous heaths, sedges, fescue grasses, sparse thyme bushes and variable proportions of mosses.

The area is a part of a large antiformal stack within the duplex of the Moine Thrust Zone. The lithologies include Archaean to mid-Proterozoic acid-intermediate banded gneisses and mafic-ultramafic complexes of the Lewisian, late-Proterozoic feldspathic sandstones of the Torridonian, deformed Moinian semi-pelitic metasediments, Cambrian to Ordovician quartzites, pelites and carbonates, and a variety of silicic and peralkaline intrusive masses of Silurian to Devonian age, as well as numerous basic dykes, ranging in age from early-Proterozoic to Tertiary.

3. ATM DATA OF INCHNADAMPH AREA

Digital data which have been used in this study comprise N-S overlapping swaths from the Daedalus-1268 line scanner flown on behalf of the Natural Environment Research Council (NERC) and characteristics of the instrument, which is described as the Airborne Thematic Mapper (ATM), are described by Williams (1984).

4. RESULTS AND DISCUSSION

Figures 1 to 4 show band 3, band 7, band ratio 3/7 and band ratio 7/3 images, respectively. Comparison between images of the band ratios 7/3 and 3/7 clearly illustrate the improved quality of the band ratio 3/7 image (Figure 4). The reverse of band ratio 7/3 reverses the grey values of pixels. As discussed earlier, the band ratio 7/3 reduces the topographic effects but does not improve the image for geobotanical interpretation. Sometimes when the terrain under investigation is not very rugged then the shadow effects provide better opportunity in geological interpretation. In these circumstances an image provided by the band ratio 7/3 is less helpful whereas the band ratio 3/7 image provides both topographic and vegetation information.

Table 1: Statistical analysis of band ratios 3/7 and 7/3.

Variables	Mean	Standard Deviation	Minimum value	Maximum value
Band 3	72.2	17.0	5	252
Band 7	80.3	26.3	5	180
Band ratio 3/7	101.2	44.86	0	255
Band ratio 7/3	110.3	28.80	24	255

A statistical analysis of both ratio images reveals the distribution of pixel values between 0 and 255. The minimum and maximum pixel values of the band ratio 7/3 image are 24 and 255 respectively whereas in the band ratio 3/7 image these values are 0 and 255 respectively. A comparison between the standard deviations for the two ratios suggests that the pixel values of band ratio 3/7 are more widely distributed than those of band ratio 7/3, see table 1.

The correlation between these two band ratios reveals the nature of the distribution of pixel values in the range 0 to 255. A bivariate plot of both ratios provides an opportunity to understand the relation between these band ratios, see Figure 5. As Figure 5 illustrates, these ratios are not simply inversely proportional to one other. At both the top and bottom range of grey levels the relation between these ratios varies. This also emphasises that the grey values in both ratio images are not exactly the reverse of one other. The

variations in the lower range of grey levels of the pixels for these ratios also explains the occurrence of variations in the standard deviations and the images themselves.

Visual interpretation of both ratio images and their comparison with existing geological maps illustrates that the band ratio 3/7 image reveals more information then the band ratio 7/3 image. For example, Fucoids beds (see Figure 6) south of Inchnadamph in the 3/7 ratio image can be recognised whereas with surrounding rocks these beds do not provide contrasting signatures (Figure 3). The boundary between Lewisian Gneiss and Basal Quartzite (see Figure 6) can be identified easily in the band ratio 3/7 image; however, this boundary is not clear in 7/3 band ratio image. Other important features such as folded beds of Basal quartzite and Lamprophyres can be mapped in detail in the band ratio 3/7 image. Recent sediments and slightly older sediments which are covered by vegetation provide contrasting spectral signatures in the band ratio 3/7 image. The area south-east of Inchnadamph is partially covered by peat, but in recent years due to more exploitation of peat beds the underlying Durness Limestone is being exposed and therefore it is difficult to compare this particular area in images with existing geological maps. However, on the basis of information available from recent remotely-sensed images, existing geological maps can be improved. It is important to notice that these ratio images also provide information about recent removals of peat and relatively old removals of peat. The light-toned areas represent areas where recent peat removal activity has been performed.

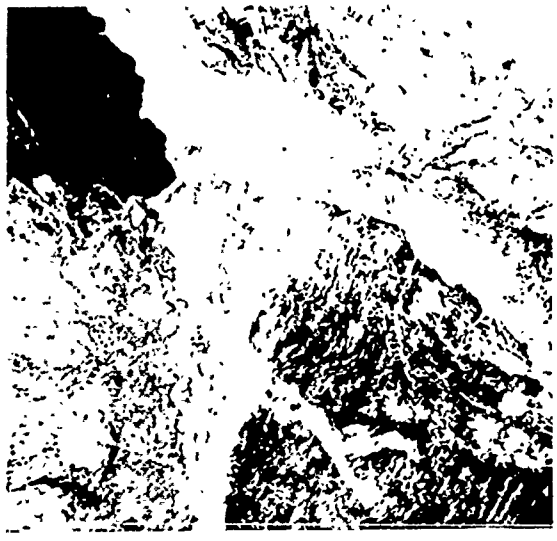


Figure 1: Linear stretched image of band 3 of Inchnadamph area.

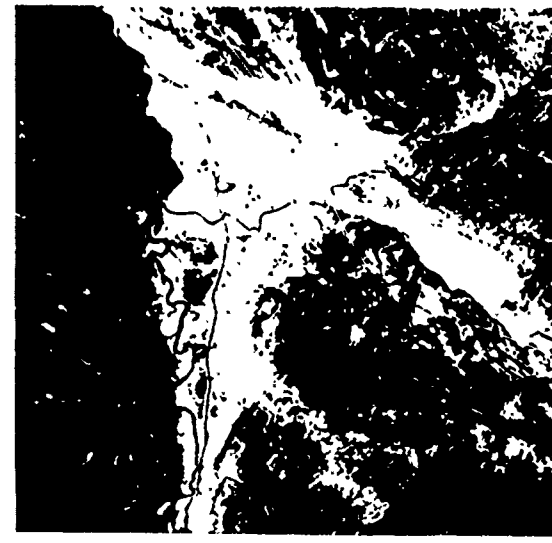


Figure 2: Linear stretched image of band 7 of Inchnadamph area.

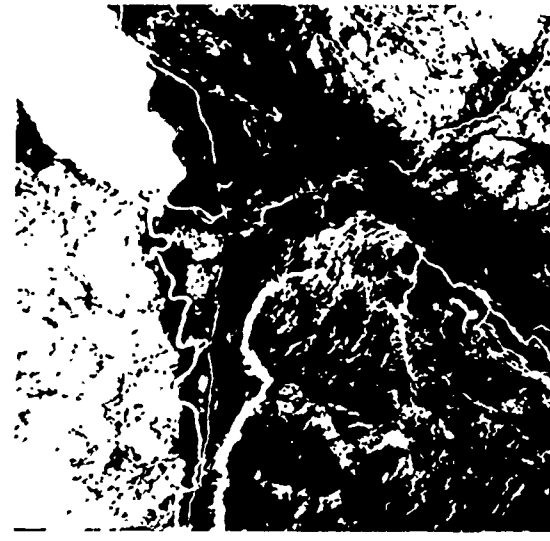


Figure 3: Linear stretched image of band ratio 3/7 of Inchnadamph area.

5. CONCLUSION

This study manifests the application of the 0.57μm/0.88μm ratio technique and provides more ground details than other usual infrared/red band ratio images in a vegetated terrain. In those conditions when the availability of bands is limited this technique can be applied to improve existing geological information. For areas which are relatively less rugged, and do not suffer too much from shadow effects, the application of this technique will sometimes preserve topographical details sometimes which are quite useful, as well as improving ground details in a vegetated terrain.

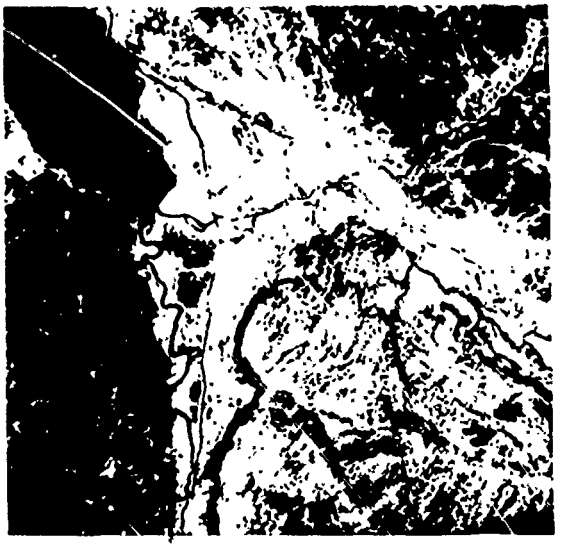


Figure 4: Linear stretched image of band ratio 7/3 of Inchnadamph area.

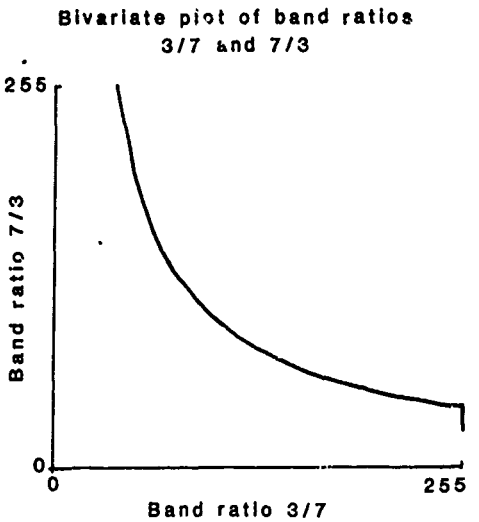
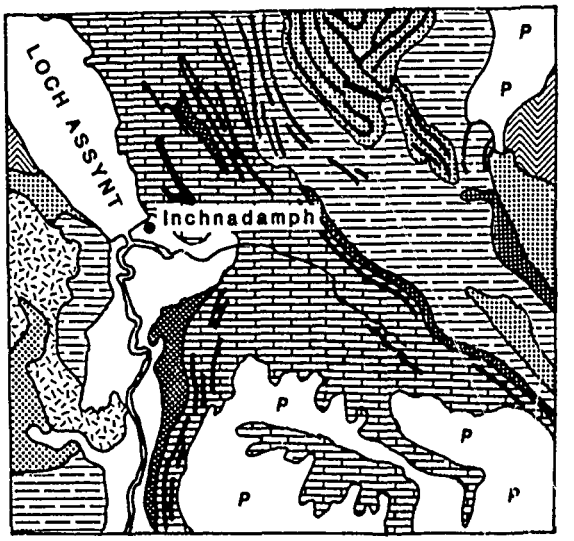


Figure 5: Bivariate plot of band ratios 3/7 and 7/3.



P Peat

Alluvium

Durness Limestone

Fucoid Beds

Pipe-rock

Basal Quartzite

Lewisian Gneiss

Felsite

Basic dykes

Figure 6: An extract from the geological map of the Assynt district, parts of sheets 107, 108, 101 and 102 (British Geological Survey, 1965). The map shows the area south and east of Loch Assynt, around Inchnadamph.

6. ACKNOWLEDGEMENT

We wish to thank the National Environment Research Council (NERC) for supplying the ATM data.

7. REFERENCES

1. British Geological Survey (BGS), (1965), Geological map of Assynt district, Geological Survey of Great Britain (Scotland), parts of sheets 107, 108, 101 and 102, Director General of the Ordnance Survey, Chessington, Surrey, 1965.

2. Crippen, R. E., (1988), The dangers of under estimating the importance of data adjustments in band ratioing, International Journal of Remote Sensing, 9, 767-776.

3. Drury, S. A., (1987), Image Interpretation in Geology, (Allen and Unwin Ltd., London).

4. Williams, D. F., (1984), Overview of the N.E.R.C. Airborne Thematic Mapper Campaign of September 1982, International Journal of Remote Sensing, 5, 631-634.

A BACKGROUND GEOBOTANY MODEL APPLIED TO GEOLOGIC EXPLORATION IN TROPICAL FOREST ENVIRONMENTS: FIRST RESULTS FROM LANDSAT-TM AND DTM FOR THE POJUCA Cu-Zn DEPOSITS, CARAJÁS PROVINCE, BRAZIL

W.R.PARADELLA [1]; W.D.BRUCE [2]; J.K.HORNSBY [3]; C.KUSHIGBOR [3]

[1] INPE (INSTITUTO DE PESQUISAS ESPACIAIS) - São José dos Campos, São Paulo, Brasil - 12201
[2] CCRS (CANADA CENTRE for REMOTE SENSING) - Ottawa, Ontario - Canada - K1AOY7
[3] INTERA TECHNOLOGIES LTD. - Ottawa, Ontario - Canada - K155H4

ABSTRACT

A Background Geobotany approach has been applied for geologic exploration in the Archean metavolcanic Pojuca Belt, Carajás Province. The test site is characterized by rugged topography; thick latosoil profiles and almost absence of bedrock outcrops. The vegetation is typical of tropical rainforest. The analysis foccused on community level and has been tested through digital image processing of orbital RS (TM-LANDSAT) and DTM data. The preliminary results of the investigation have shown that 11 RS biophysical classes were detected; the majority with spatial correlations with previous geological units. In addition, elevation and slope have presented an important role in the characterization of the geobotanical background patterns.

Key words:
GEOBOTANY; TM-LANDSAT; DTM; AMAZON REGION.

1. INTRODUCTION

Satellite-based remote sensing has been used successfully in tropical semiarid environments for lithological discrimination (Paradella et al, 1987). Results in tropical rainforest regions are scarce however, relying on spatial attribute analysis or having site specific conditions where iron laterites cause changes in biomass which result in spectral contrasts (Paradella , 1986) . At more temperate latitudes, a Background Geobotanical (BG) approach was developed for extracting geologic information in vegetated terrains (Bruce & Singhroy, 1984). It is based on the hypothesis that the nature and distributions of plant communities occur as a result of environmental conditions and therefore, variations in distribution from the established norm of an area may be indicative of changes in the geology (Hornsby et al, 1988).

In Brazil, recent discoveries of Cu-Pb-Zn, Ni, Cr, W and Sn have been made in the Carajás Province of the Amazon craton, in addition to the known Fe, Al, and Mn deposits in the area. The province is characterized by heterogenous topography, deep chemical weathering producing thick latosoil and little outcrops. Vegetation cover consists of a complex and multi-level canopy composed of numerous species.

This paper reports on the first results of a research program which has been established to develop remote sensing applications in this kind of environment. The following questions are to be answered: 1) to what extent is it possible to extract geological information from satellite-based remote sensing data, given the commonly held view that tropical rainforest vegetation is so complex as to appear homogeneous and featureless? 2) what are the controls on spectral patterns and of what comparative importance are geological controls among these? and 3) what is the optimum approach to extract geologic information from the most commonly and economically available data set (remote sensing and topography) in the planning and implementation of exploration activities?

1.1 - STUDY AREA DESCRIPTION

A 230 square kilometre test area was selected, centered on the Pojuca Cu-Zn deposit located in the Carajás mountain range. The oldest of these Archaen rocks are a sequence of medium to high grade metamorphic rocks known as the Xingu Complex. This complex occurs in the northeast portion of the study area and has characteristic low relief. Discordant with this basement is a volcano-sedimentary sequence characterized by west-northwest trending parallel ridges. This is the Igarapé Pojuca Group consisting mainly of greenschist to amphibolite phases rocks with several Cu and Cu-Zn deposits. Above the Igarapé Pojuca Group are widespread low-grade clastic sediments of early Proterozoic age called the Rio Fresco Group, and located in the southwest of the study area. It is characterized topographically by a plateau with small intersecting hills. Tertiary laterite and Quaternary alluviums complete the geological sequence (Figure 1).

The vegetation cover is typical for tropical rainforest. Three ecological regions can be recognized: 1) Montana tropical forest associated with the highest altitudes (Rio Fresco Group); 2) Sub-Montana tropical forest with woodland/broadleaved forest (Pojuca Group and part of Xingu Complex and 3) Mixed woodland forest (Xingu Complex) (Veloso et al, 1974). A total of 552 specimens with 119 species, 89 genera and 39 families were identified from one hectare transect in this area (Silva et al, 1987).

2. METHODOLOGY

Figures 2 and 3 illustrate the methodological approach. Landsat Thematic Mapper (TM) data from May 31, 1984 and a digital elevation model (DEM) were

used in the investigation. The 512 x 512 pixel area was geometrically corrected to a UTM projection base. Elevation, slope, and aspect images were derived from the elevation model. This data set was analysed through a combination of enhancements (frequency equalization, band-ratios, principal component transformation) and unsupervised (clustering) technique..A feature selection routine called the optimum index factor (OIF) was used to aid in the selection of the best colour composite image (Chavez et al, 1982). Eight biophysical maps were interpreted and integrated into a final integration map with 11 biophysical classes. This map was compared with supporting geological information to identify geobotanical relationships.

3. RESULTS

The final integration map is shown in Figure 4. Figure 5 is the best enhancement product obtained with TM data. This was a PC transformation of the input bands 3/2, 4/1, 4/3, 3/5, 7/4, 5/4, 5-4/5+4, 5-7/5+7, 4-3/4+3. From Figure 4 the following class associations can be made: class A = laterite; B = Rio Fresco Group; C = Rio Fresco Group; D = Igarapé Pojuca Group; F = Igarapé Pojuca Group; F = Granitic areas; G = Xingu Complex, H = Quaternary sediments. There is insufficient information about the substratum to identify geologic associations with classes I and J. In the colour composite different tonal patterns are detected, of particular interest, one correlated with the Pojuca trend (area A).

Figure 6 is the result of clustering the elevation and slope data. More than half of the biophysical classes have spatial correlations with the thematic class produced from the elevation and slope data. The associations are as follows: class A with 2; B with 3; D partially with 6 and 5; G with 7; H with 8, 9 and 10; I partially with 9.

4. CONCLUSIONS

The geologic characteristics of the terrain play an important role in controlling the spatial distribution of the biophysical remote sensing classes. In addition, geology also controls in part, the main geomorphometric terrain descriptors of elevation and slope. The role of aspect, which appears much more complex, was not evaluated in this phase of the research. The detection of these geobotanical patterns will help to improve the delineation of main geological units elsewhere in the Carajás Province beyond the area of detailed mapping. This is particularly important in the case of the Pojuca Group which hosts the Cu-Zn mineralizations. Future development of this methodology is planned in order to produce an operational tool for geological exploration in tropical environments using thematic mapper and topographic data.

BIBLIOGRAPHY

1. Bruce, W.D.; Singhroy, V.H. Landsat data for mineral exploration: the canadian experience. RESORS #1050207, 12p., 1984.
2. Chavez Jr., P.S.; Berlim, G.L.; Sowers, L.B. Statistical method for selecting Landsat MSS ratios. Jour. of Appl. Phot. Eng., 8. p.23-30, 1982.
3. Docegeo. Revisão litoestratigráfica da Província mineral de Carajás. Cong. Bras. Geol. 35, anexo/anais. p. 11-54, 1988.
4. Hornsby, J.K.; Bruce, W.D.; Harris, J.; Rencz, A.N. Implementation of background and target geobotanical techniques in mineral exploration. Them. Conf. RS for Expl. Geol., VI, 11 p. 1988.
5. Paradella, W.R. Avaliação preliminar de dados TM Landsat para exploração mineral na região amazônica: intrusivas de Maraconaí e Maicuru (Pará). Cong. Bras. Geol., 34. p.2951-2964. 1986.
6. Paradella, W.R.; Vitorello, I.; Liu, C.C.; Mattos, J.T.; Meneses, P.R.; Dutra, L.V. Spectral and spatial attributes evaluation of SPOT data in geological mapping of Precambrian terrain in semi-arid environmentas of Brazil. Colloque SPOT 1. Image utilization, Assessment, Results, Paris, p. 851-860. 1987.
7. Silva, M.F.F.; Rosa, N.A. Análise da vegetação sobre a jazida de cobre na Serra dos Carajás, Pa. Relatório do Convênio CNPq-MPEG/CVRD. 32p. 1987.
8. Veloso, H.; Japiassu, A.M.S.; Goes Filho, L.; Leite, P.F. As regiões fitoecológicas, sua natureza e seus recursos econômicos. Estudo fitogeográfico da área abrangida pelas Folhas SB-22 Araguaia e SC-22 Tocantins. Projeto RADAM, vol. 4, 119 ps. 1974.

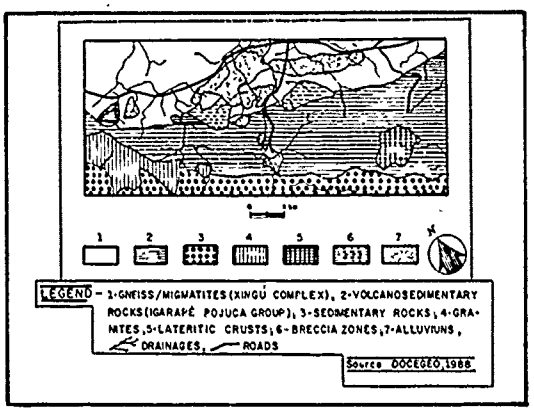


Figure 1. Geology of the Pojuca area.

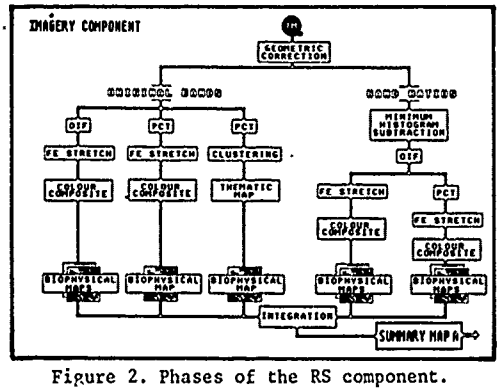


Figure 2. Phases of the RS component.

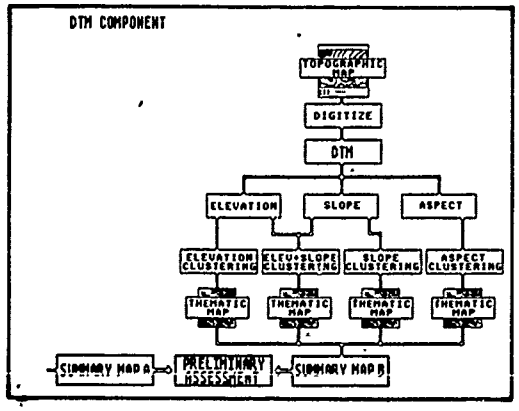


Figure 3. Phases of the DTM component.

Figure 4. Final integration map.

Figure 5. Best enhancement product.

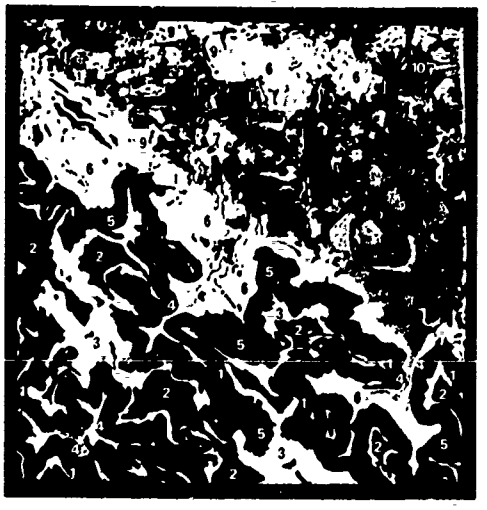


Figure 6. Unsupervised classification of elevation and slope.

GEOBOTANICAL AND BIOGEOCHEMICAL ANOMALIES OVER A FLUORITE MINERALIZATION IN KERIO VALLEY, KENYA

B.D.Odhiambo [1] J. Gaciri [2] and A. Falconer [3]

[1] Department of Geography, University of Waterloo, Waterloo, Ontario N2L 3G1, Canada

[2] Department of Geology, University of Nairobi, P.O. Box 30197, Nairobi, Kenya

[3] Regional Centre for Services in Surveying Mapping and Remote Sensing, P.O. Box 18118, Nairobi, Kenya

Abstract

A study has been undertaken to determine whether or not unique vegetation characteristics are associated with fluorite mineralization. The study included analysis of remotely sensed images, field data and existing maps. These thematic data sets were overlain and Geobotanical Units identified. The results indicate that major vegetation communities coincide with distinct geomorphological, pedological and lithological units. The plant species *Harrisonia abyssinica, Vangueria acutiloba*, and *Balanite aegyptiaca* localize the fluorite mining sites because of their anomalously high concentration levels of the path-finder elements in their ash. Furthermore, the parent rock, which is a hydrothermally-metamorphosed limestone, coincides with Geobotanical Unit I'; a unit characterized by a distinct vegetation community, and a distinct photographic and spectral anomally. It is concluded that fluorite deposits can be prospected for rapidly by an analysis of indicator species.

Keywords: Geobotany, Kenya, Landsat imagery, Mineral exploration

Introduction

It has been demonstrated that certain mineral concentrations can be identified using Landsat data. The mineral concentrations are localized by a change in reflectance of the vegetation affected by the mineralization, by a change in the density of the vegetation over the mineralization or even by the presence of some unique indicator plant species growing around the area of mineralization.

In Kenya, fluorite is one of the major commercial mineral deposits. In this study, an analysis of the fluorite path-finder elements has been undertaken to determine whether or not unique vegetation characteristics are associated with the fluorite mineralization and, if so, whether or not these areas can be identified using Landsat imagery.

Study Area

The study area is part of the Central Kenya Dome located on the western flank of the East African Rift Valley. It covers an area of approxiamtely 7.5 km x 3.7 km and is bounded geographically by latitudes 0°18'N and 0° 22'N and by longitudes 35° 40'E and 35° 38'E. From the point of view of mineral rights, the area falls under the jurisdiction of the Kenya Fluorspar Company Limited.

Geologically, the area is composed of the Precambrian Shield rocks, Upper Miocene volcanics, Lower Miocene sediments, Pliocene sediments, and Quarternary to Recent deposits. Fluorite, or fluorspar, the dominant economic mineral deposit is a result of hydrothermal emanations replacing calcite in limestone. However, the general lack of high-temperature minerals along the zone of alteration suggests emanation of hydrothermal fluids at low temperatures (Nyambok and Gaciri, 1975).

Materials and Methods

The geobotanical and biogeochemical study included analysis of remote sensing images, field data and existing maps. The remote sensing data consisted of Landsat Multispectral Scanner (MSS) and Landsat Thematic Mapper (TM) transparencies recorded on June 14, 1984. The satellite data were digitally and photographically enhanced and enlarged to a working scale of 1:50,000. The photographic enlargements included 70 mm x 70 mm chips for multi-band analysis using the ANALOG Color Additive Viewer. Digital analysis was performed on the PERICOLOR 1000 interactive digital image processing system . The original imagery had been processed using a supervised classification (LeDurand, 1983). Use was made of a Diskette Classification Program PC © SPOT and a colour composite.

It has been determined that the TM Band 3 image best enhances the geological features (Odhiambo, 1988). Radiometric values of the Band 3 image were systematically recorded along two transects (Figure 1) eastwards from the Kimwarer mine site, through the mill tailings across the study area. Using a Portable Field Radiometer, whose spectral bands were set to duplicate those on board the Landsat satellite, spectral measurements were made of the most dominant plant species in the geobotanical units. The radiant energy incident upon the plants, $Ei(\lambda)$, and that reflected from

it, Er(λ), were measured. These spectral measurements were recorded in each of the four spectral bands and normalized using a white target (assumed to be a standard reflectance surface). The percent spectral reflectance (R%) was then calculated using an equation modified from Lillesand and Kiefer (1987). Panchromatic aerial photographs recorded in December 1969 were used as an ancillary source of information. These photographs were stereoscopically analysed for vegetation/habitat, morphological and geological detail, to delineate the eleven Geobotanical Units (GBUs).

Plant and soil samples were collected in each GBU for quantitative element analysis. The plant samples were flushed with distilled water to remove any soil particles and were then oven-dried at 120° C for about six hours. Next, they were placed in a 25 ml squat PYREX beakers and were ashed in a muffle furnace (Carbolite, Model ESF 2) at approximately 450° C, for 8 to 12 hours. The ashed plant material and the fused finely-pulverized soil samples were diluted by a factor of about five using ANAL cellulose. Using a pressure, pellet making dies, and a visor, thin pellets were pressed out at 3000 kg/cm for subsequent irradiation by the Energy Dispersive X-ray Fluoresence Analytical (EDXRFA) system. The overall accuracy of the EDXRFA system was estimated to be between 5% and 10%, while its sensitivity ranged from 0.2% for aluminum (Al) to 38 ppm for uranium (U), in solid samples Lavi (1984).

Results and Discussion

The distribution of plants in the geobotanical units suggests that *Croton dichogamus* is the most widely distributed (dominant) plant species, followed by *Combretum molle*, *Ipomoea spathulata* and then by *Acacia tortilis*. *A tortilis* forms the overstorey, while *Croton dichogamus* forms most of the understorey. Lesser plants species are indicators of specific geobotanical units; for instance, *Ficus natalensis* and *Cleorodendron* spp. are indicator species for GBU VII, while *Commifora* spp., *Commbretum molle* (with a swollen stem) and *Harrisonia abyssinica* are the geobotanical indicator species for GBU I'.

In the Geobotanical Units (GBUs), 54 plant species were identified, 53.7% of which were selected randomly and assayed for their element concentrations. It is notable, however, that the most dominant plant species assayed comprise only 20.4% of the total number of species sampled. The rest of the plants, assayed only once, make up 33.3% of the plant samples and comprise the lesser-occuring plants in the GBUs. Emphasis for trace element analysis was, therefore, on the lesser-occuring species - which are indeed the indicator species for the various GBUs and localize the fluorite deposits. The dominant species on the other hand constitute the major vegetal component of the study area, and define approximately the underlying geological formations.

The quantitative trace element analysis for the elements calcium (Ca), zinc (Zn), manganese (Mn), rubidium (Rb), strontium (Sr), zirconium (Zr) and uranium (U) was done on the ash of all the plants assayed using EDXRF analysis. The results are summarized in Table 1. These biogeochemical data were smoothed using a moving average technique, as

Table 1. Mean Element Concentrations and Standard Deviations in Plant (68) and Soil (21) Samples

Element	Ca(%)	Cu	Zn	Mn	Rb	Sr	Zr	U
Plants (68)								
Mean	16	262	394	1596	185	2942	233	17
Sdev.	8	358	619	2956	472	4058	337	20
Max. Val.	37.7	2369	5148	2.21%	3720	2.96%	2152	101
Soils (21)								
Mean	1.5	426	213	2122	1[illegible]	449	526	15
Sdev.	2.1	589	64	1053	336	220	196	19
Max. Value	9.7	2924	387	4130	1630	975	846	85

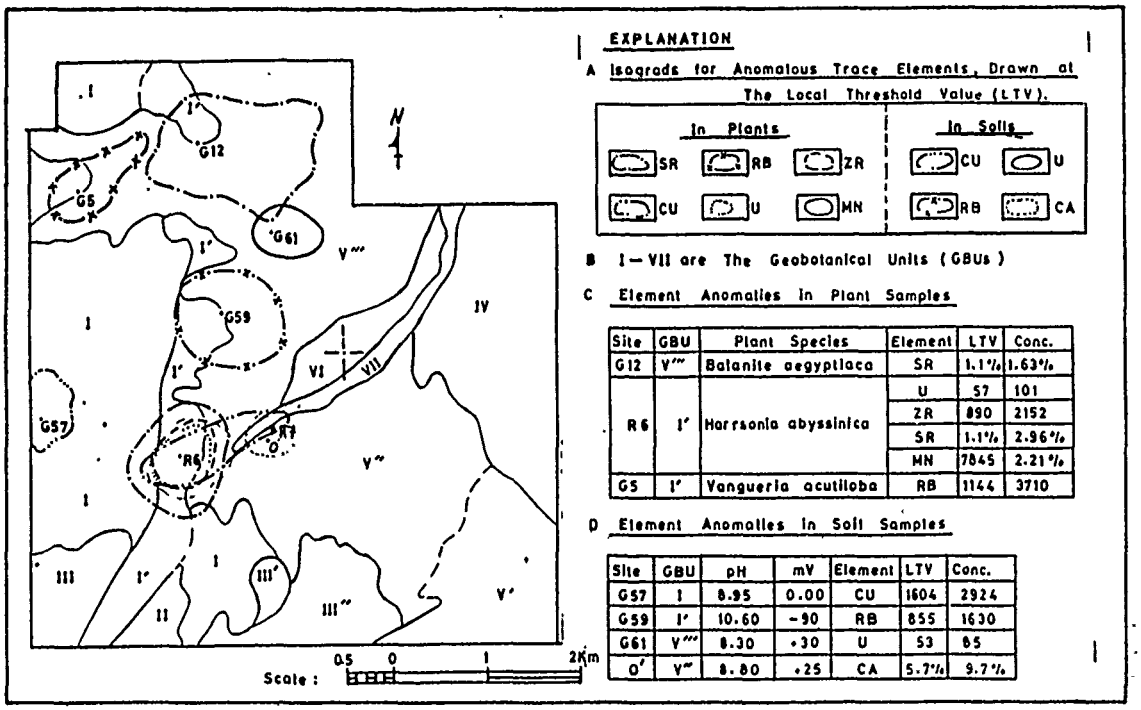
recommended by Tooms and Jay (1964), Malyuga (1964) and Sayala (1979). Briefly, it is notable that the indicator plant species have concentrations above the local threshold value. For example, *Ipomoea spathulata* contains 1740 ppm Cu, *Vangueira acutiloba* has 3720 ppm Rb, *Balanite aegyptiaca* contains 16300 ppm (1.63%) Sr. Of most interest was the fact that *Harrisonia abyssinica* contains 29600 ppm (2.96%) Sr, 2924 ppm Cu, 22100 ppm (2.21%) Mn, 5148 ppm Zn and 101 ppm U. *Harrisonia abyssinica*, therefore, stands out as the best fluorite indicator species (in terms of the number of anomalous minerals detected).

In the soil samples, anomalous concentrations of calcium occur over the mill tailings, while high uranium values occur below the Choff deposits. In both these cases, the Eh values are +25mV and +30mV, respectively (indicating a mildly reducing geochemical environment). At Site G59, rubidium concentration is above the LTV, but the Eh value is -90mV (a strongly reducing environment). These data, which are summarized in Figure 2, indicate anomalous trace element concentrations are associated with GBU I', on the Precambrian paragneisses with the fluorite mineralization. The concentration plots are also displayed in Figure 1, above the mineralized geological unit. It is observed that anomalous calcium concentrations localize the friable calcareous sandstone deposit east of the area. Over the fluorite deposits, the concentration is far below the local threshold value and even below the mean concentration value. This clearly indicates a negative biogeochemical anomaly of calcium over the fluorite deposits.

From the radiometric values along the two transects (Figure 1), it is evident that the mine sites have a higher reflectance than the mill tailing. This is interesting since the tailings appear brighter (with a higher reflectance) than the exposed mine sites. Field visits revealed that the tailings were wet and/or covered with a thin film of water (1 - 2 cm). It seems, therefore, that the digital image processing technique applied to the satellite imagery actually discriminated the tailings from the mine sites.

Conclusions

The case study establishes that the geobotanical and biogeochemical techniques applied in this



Site	GBU	Plant Species	Element	LTV	Conc.
G12	V'''	Balanite aegyptiaca	SR	1.1%	1.63%
R6	I'	Harrsonia abyssinica	U	57	101
			ZR	890	2152
			SR	1.1%	2.96%
			MN	7845	2.21%
G5	I'	Vangueria acutiloba	RB	1144	3710

Site	GBU	pH	mV	Element	LTV	Conc.
G57	I	8.95	0.00	CU	1604	2924
G59	I'	10.60	-90	RB	855	1630
G61	V'''	8.30	+30	U	53	85
O'	V'''	8.80	+25	CA	5.7%	9.7%

Figure 2. Summary of significant element concentrations in plant and soil samples, as related to GBU 1'.

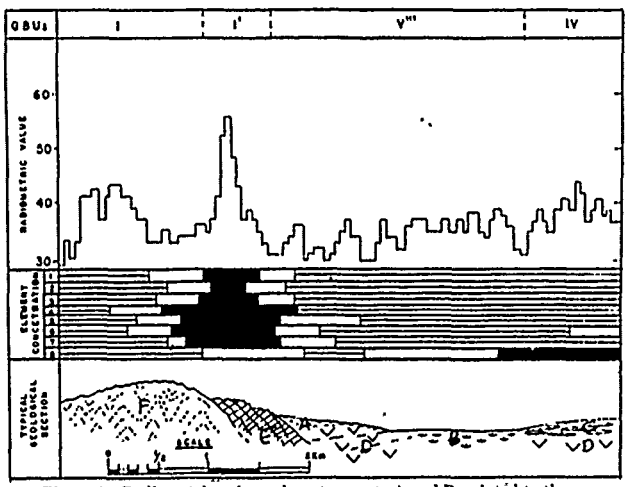
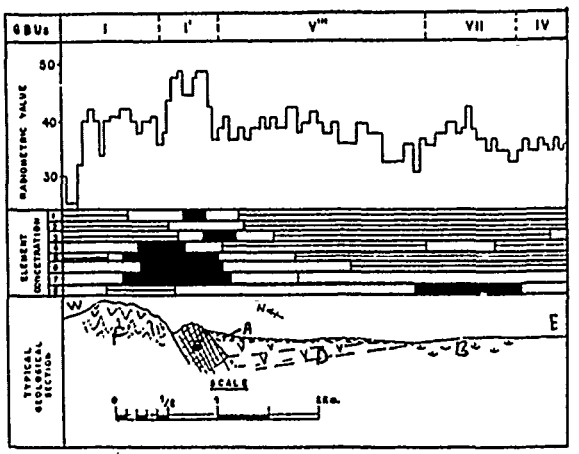


Figure 1. Radiometric values along transects A and B, related to the geology, biogeochemistry and to the geobotanical units. Values obtained from the Digital Imagery Processing System for Channel 6 only.

Lithology

A Quaternary to Recent alluvial deposit
B Flood plain of the Kimwarer River
C Soy Sandstone deposit
D Tertiary Volcanics
E Fluorite Mineralization
F Precambrian Para-gneisses



Element Concentration Range

Below X + Sdev.

From X + Sdev. to X + 2xSdev.

Above X + 2xSdev.

N.B. 1 = Zr, 2 = Rb, 3 = U, 4 = Sr, 5 = Zn, 6 = Cu, 7 = Mn, 8 = Ca

research localized the fluorite deposits along the foot slopes of the Kerio Escarpment, in GBU I'. This GBU is itself associated with hydrothermally-altered limestone deposits in the Precambrian Formation. The plant species *Harrisonia abyssinica*, *Balanite aegyptiaca*, and *Vangueria acutiloba* appear to be the potential fluorite indicator species in the Kerio Valley area. The spectral data obtained from field measurements and from digital image processing, demonstrate the potential of using remote sensing techniques in geobotanical and biogeochemical methods for rapid exploration of mineral deposits, especially in developing countries.

Acknowledgements

The authors are indebted to the Government of Kenya for financial assistance through a University of Nairobi Scholarship Grant , and to the Regional Centre for Services in Surveying Mapping and Remote Sensing, Nairobi. Mr. David K. Andere, Director of the Department of Resource Surveys and Remote Sensing, Nairobi, kindly made equipment available.

References

1. Lavi, M.M., "Practical Approaches in Quantitative X-Ray Fluoresence Analysis", Unpubl. M.Sc. Thesis, University of Nairobi, 78.pp, 1984.

2. Le Durand, C., "Landuse Mapping in Kericho Town Area Using Digital Analysis of Satellite Data", KREMU Tech. Rep. No. 114, 20 pp, 1983.

3. Lillesand, T.M. and R.W. Kiefer, "Remote Sensing and Image Interpretation", 2nd Ed., John Wiley & Sons, New York, 721 pp., 1987.

4. Nyambok, I.O. and S.J. Gaciri, "Geology of the Fluorite Deposites in Kerio Valley, Kenya", Econ. Geol., Vol. 70 No. 2, pp. 299-307,.1975.

5. Malyuga, D.P., "Biogeochemical methods of prospecting" (English translation from the Russian), Consl. Bureau, New York, 205 pp., 1964

6. Odhiambo, B.D.O.,"A Geobotanical Study Around the Fluorite Mineralization in Kerio Valley, Kenya", Unpubl. M.Sc. Thesis, University of Nairobi, 101 pp., 1988.

7. Sayala, D., "Regional Geochemical Exploration for Mineral Deposits in the Applachians of Central N.W. Virginia". Unpubl. Ph.D. Thesis, Washington University, In Dissert. Abs. Int., Vol. 40, No. 4, pp. 1587-1588, 1979.

8. Tooms, J.S. and J.R. Jay, "The role of the biogeochemical cycle in the development of Cu/Co anomalies in the freely drained soils of the Northern Rhodesia Copperbelt", Econ. Geol., Vol. 59, pp. 826-834, 1964.

MONITORING VEGETATION REGROWTH ON PLACER MINE TAILINGS, BONANZA CREEK, YUKON TERRITORY

J.K. Hornsby
Intera Technologies Ltd., Ottawa, Ontario. K1Z 8R9

B. Bruce
Canada Centre for Remote Sensing, Ottawa, Ontario. K1A 0Y7

G. Mackenzie-Greive
Environmental Protection Service, Whitehorse, Yukon Territory. Y1A 3V1

ABSTRACT

A feasibility study was undertaken to evaluate the use of satellite-based remote sensing data and techniques for monitoring vegetation regrowth on placer mine tailings. The intention was to develop a methodology which could enable repetitive monitoring by reducing the dependence on costly ground based methods which are currently used.

The research objectives were to develop a remote sensing-based method for monitoring natural vegetation regeneration on placer mine tailings, and to identify the underlying relationships among environmental, terrain, and human factors controlling vegetation regrowth. The study area was located along Bonanza Creek near Dawson City in the Yukon Territory. This has been, and still is, an area of extensive placer mining activity since the turn of the century. As a result the tailings are of varying ages and represent the result of a variety of mining methods.

A methodology was designed and tested using Landsat Thematic Mapper data. These data provided sufficient resolution for the identification of ten classes of terrain. Field investigations revealed that time was not the dominant factor controlling the regrowth of the vegetation rather the method used in mining the placer gold and the structure and composition of the resulting tailings were the more important factors. Working from an established base of knowledge in an area, the monitoring of the vegetation classes would be limited to the cost of the imagery, the processing of the data, and limited field inspection for quality control. It should be recognized however, that to detect subtle changes in growth from one year to the next, it would be advisable to user higher resolution airborne data and more extensive field sampling.

Keywords: Vegetation Regeneration, Placer Mining, LANDSAT

1. INTRODUCTION

Placer mining activity in Yukon Territory has occurred extensively since the turn of the century. The Klondike gold creeks, located immediately southeast of Dawson City, have been mined since the Klondike Goldrush and are still producing. In 1988 placer mining activity in the Yukon recovered 474 kg. (152,555 crude oz). This mining activity resulted in the stripping of vegetation from the valley bottoms and sides, and the subsequent deposit of tailings. In many areas claims have been reworked several times over the years.

Regrowth of vegetation has occurred at varying rates on the areas mined (Brady, 1984). The progress of this regrowth leading to an eventual return to a natural forest cover, is of concern to the Environmental Protection Service of Environment Canada. A need was therefore identified, for the monitoring of natural revegetation of placer mine tailings. Remote sensing data and methods were considered a possible means for monitoring the vegetation. Two main objectives were identified to investigate this possibility; 1) to develop a romote sensing-based method for monitoring vegetation regeneration on placer mine tailings, and 2) to identify underlying relationships among environmental, terrain, and human factors controlling vegetation growth.

1.1 Study Area Description

The study area is focussed on Bonanza and Eldorado Creeks (Figure 1). The area is approximately 50 square kilometers in size. The topography is rugged with elevations exceeding 1128 metres asl, and incised non-glaciated valleys, to a low of 305 metres at the Yukon River. Vegetation cover varies from non-existent to continuous cover of Betula papyrifera (birch), Populus tremuloides (aspen/poplar), Picea glauca (white spruce), and Poa sp. (grasses).

Unconsolidated materials consist of fluvial sands and gravels occurring naturally as valley fill and terraces on the valley slopes. Bedrock consists of quartzofelspathic, carbonaceous and mafic schistose rocks. Two major structural lineaments run the length of Bonanza and Eldorado creeks.

The placers in the Klondike gold fields are Tertiary in age (Milner, 1984). The Bonanza Creek watershed has provided the greatest amount of gold of all the creeks. Much of this gold has come out of the White

Channel pay streak; a zone of coarse-grained, poorly sorted gravels.

2. METHODOLOGY

The methodology developed uses standard image processing techniques with supporting field information, to produce a final landcover map representing significant plant communities and non-vegetated surface materials (Figure 2) (Lillesand and Kiefer, 1979). The methodology has been scaled to a PC-based hardware and software system for cost effectiveness and ease of operation.

Landsat MSS and TM data was acquired for the study. The MSS was used to produce preliminary enhancements which could be taken into the field to determine the significance of variations in the spectral responses of the different ground covers. When TM data became available (image data June 14, 1986), it was used for the post fieldwork processing and classification since the increased spatial resolution was necessary to produce transferable results. Contrast stretch, principal components, and unsupervised classification operations were first applied to the visible and reflected infrared bands. These enhancements aided in the identification of training sites with the support of the field data.

Field data were collected to document the extent and nature of different vegetation types, soils, and drainage conditions found in the area. These investigations were required to determine the classes of ground cover to be extracted from the remote sensing data. Collection of this information enabled geobotanical associations to be determined among vegetation, terrain and mining activities over time.

3. INTERPRETATIONS

Ten significant cover types were identified from the results of the field investigations. These were based on two criteria; geobotanical associations, and material composition where vegetation cover was absent (Table 1).

The spectral responses of each class were relatively distinct (Table 2). The active mining class had high mean reflectance values in all bands due to the lack of vegetation and the physical and textural properties of the gravels. The recent tailings class surrounds the active class. While it is also non-vegetated it retains a distinct response having generally lower mean reflectance values. The dredge tailings in the Bonanza Creek valley have a similar response to the recent tailings class particularly in the shorter wavelengths. This leads to some confusion between the two classes. It does however have higher reflectance minima in bands 4 and 5. The Klondike River valley tailings were clearly identified with only minor confusion with more recent tailings on the south side of the valley. The hilltop tailings class represent workings high on the valley slopes. These have a discontinuous vegetation cover. The deciduous and coniferous classes are extensive over the area. The classes are actually mixed in composition with the difference being in the dominance of either conifer or deciduous trees. The valley fill class, having a unique vegetation composition of conifers, grasses and shrubs, also had distinct spectral responses from the deciduous and coniferous mix classes.

4. CONCLUSIONS

The methodology developed in this project has been applied successfully with results which support the use of satellite-based remote sensing data as a means for monitoring vegetation regrowth on placer mine tailings. Discrimination of six vegetation classes was possible corresponding reliably to geobotanical associations with underlying materials and topographic influences. Three classes of non-vegetated tailings types were also separable, having distinctive tailings characteristics. It was observed that age did not always play a significant role in the amount of revegetation which has occurred. The extremely coarse nature of dredge tailings and relatively high associated local relief does not allow for fines to collect in sufficient quantity to support vegetation. Without active intervention to address this problem these areas can be expected to remain much the same over time.

Operational implementation of this methodology would require limited expenditure in time and cost once preliminary field verification confirms the location of training sites for each class. Costs could be limited almost entirely to acquisition of the data and a minimum of image processing. Higher resolution airborne sensors may be used to define more specific ground covers and/or more precise boundaries, however costs would be increased greatly.

5. REFERENCES

Brady, M.A., 1984: Natural revegetation of mining disturbances in the Klondike Area, Yukon Territory; MSc.Thesis, Faculty of Forestry, University of British Columbia, 109p.

Lillesand, T.M., and R.W. Kiefer, 1979: Remote Sensing and image interpretation; John Wiley & Sons, 612p.

Milner, M.W., 1984: Geomorphology of the Klondike Placer Goldfields; Final Report, Contract OSV 5-007, Dept. of Indian Affairs and Northern Development, 157p.

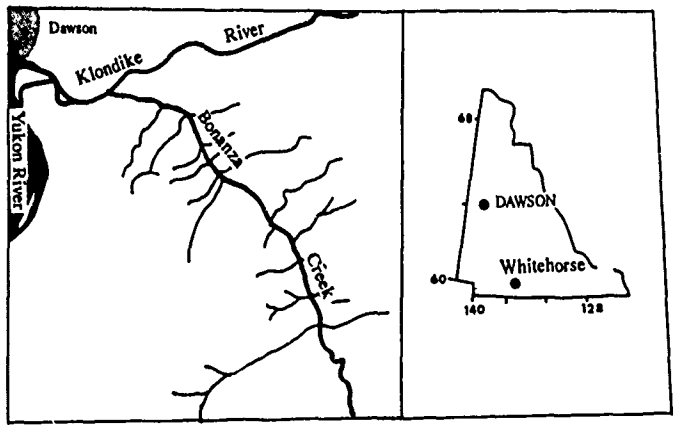


Figure 1. Location map of study area.

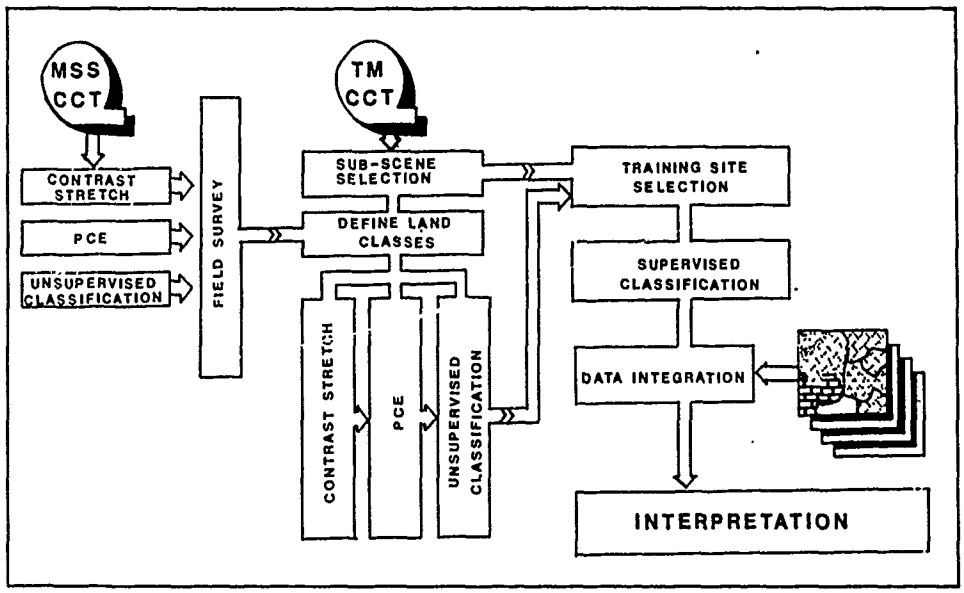


Figure 2. Image analysis methodology.

Table 1: Classes with vegetation cover, surficial material and topographic characteristics.

CLASS	VEGETATION	MATERIALS	TOPOGRAPHY
Active Mining	none	White Channel gravels	East facing valley slopes and gravels on valley floors.
Recent Tailings	none	coarse sandy gravel	Long ridges and flat areas.
Bonanza dredge tailings	sparse grass	coarse sandy gravel	Flat to hummocky.
Klondike dredge tailings	none	coarse gravels	Hummocky, in Klondike River valley.
Hilltop tailings	fireweed, grasses, spruce	gravels	High on hill slopes.
Tailings ponds	Populus - balsamifera, Betula papyrifera, Poa & Salix sp.	coarse gravels with fines	Depressions between hummocks.
Open water	none	n/a	n/a
Deciduous forest	Betula papyrifera, Populus tremuloides Alnus sp.	sand & gravel	Southeast facing valley slopes.
Coniferous forest	Picea glauea	sand & gravel	Northwest facing slopes.
Valley fill	Picea glauea Pao sp., Alnus sp.	sand, gravel & silt	Valley bottoms & lower slopes.

Table 2: Mean spectral response values and general distribution of each mapped class.

CLASS	MEAN REFLECTANCE						OCCURRENCE
	TM1	TM2	TM3	TM4	TM5	TM7	
Active mining	195	120	120	102	178	98	Bonanza Creek & off Klondike R.
Recent tailings	97	48	54	54	92	48	Around areas of active mining.
Bonanza dredge tailings	131	70	82	74	124	135	Bonanza Cr. valley bottom
Klondike dredge tailings	94	43	47	42	60	33	Klondike R. valley
Hilltop tailings	66	28	25	57	72	26	Around active mining; some SE facing slopes.
Tailings ponds	71	29	27	52	45	18	Tailings ponds; shadows.
Open water	73	30	30	18	9	4	Yukon and Klondike rives; tailings ponds.
Deciduous forest	89	43	45	71	97	42	SE valley slopes.
Coniferous forest	62	24	18	50	37	126	NE facing slopes.
Valley fill	68	27	25	52	42	60	

CORRELATION STUDY OF REMOTE SENSING AND GEOLOGICAL DATA FOR MINERAL EXPLORATION, FARLEY LAKE AREA, MANITOBA

by

Vijay Singh[1], Wooil M. Moon[1], Mark Fedikow[2] and Andy N. Rencz[3]

[1]Geophysics, The University of Manitoba, Winnipeg, Manitoba, Canada R3T 2N2;
[2]Geochemistry, Department of Energy and Mines, Winnipeg, Manitoba, Canada R3C 4E3;
[3]MRD, Geological Survey of Canada, Ottawa, Ontario, K1A 0E8.

ABSTRACT

Mineral exploration techniques involving geophysical, remote sensing, geochemical and biogeochemical data have been used successfully in the past with varying degree. Any mineral exploration approach which can integrate as many of these different techniques as possible is likely to increase the success rate and eventually have a substantial economic benefits. In this study an attempt has been made to analyze different types of remote sensing and geological data available for the Farley Lake area, Manitoba.

The Farley Lake area forms a part of the Agassiz Metallotect, and has promising economic potential for a gold deposit. Consequently, extensive exploration work has been carried out by provincial and federal governments under the mineral development agreement (MDA), and the area has been surveyed in detail by using remote sensing methods and by surface methods. The remote sensing data includes the airborne multidetector electro-optical imaging scanner (MEIS-II) data. These data provide the spectral reflectance information in different wavelength windows, and were obtained to detect the signs of vegetation stress. The other remotely sensed data set is the airborne magnetic field gradiometer data. The surface data sets were obtained as 'ground truth' and include the geochemical and biogeochemical data. For geochemical data clay-sized fraction and heavy mineral fraction of the basal till samples have been anaylzed for eight elements, whereas the six trace element concentrations of the black spruce needles constitute the biogeochemical data.

In order to test the usefulness of each of these data sets, correlations among different variables in a particular data set have been calculated, as well as the relationships among the different data sets have been determined. The results of this study indicate that the airborne MEIS-II data in general show fairly low correlations with the geochemical and biogeochemical data. However, there is a moderate correlation between the band1 and band2 of the MEIS-II and copper concentration at one of the biogeochemical sample grids. In the till geochemical data there is high correlation among the ferrous elements and the zinc concentration in the basal till. Moreover, the elemental concentration in the basal till data lacks any correlations with the aeromagnetic data. This would be expected in areas covered with thick glacial overburden. The strong magnetic anomaly in the western half of the study area is oriented roughly north-south direction, and does not correspond with the published subsurface bedrock geology map. If a magnetic anomaly is indicative of bedrock geology, as in many cases, the subsurface geology map produced from a few distant outcrops would require further confirmation by detailed geological mapping and probably modification to the available geology maps.

The digital processing and correlation, in this study, of different geological, geochemical, geophysical and airborne remote sensing data provides a quick and reliable basis, on which we can quantitatively weigh each different data set for future systematic integration of remote sensing and geological data.

Key Words: Digital processing, Correlation, Geophysical, Geochemical, Biogeochemical, Remote Sensing, Exploration.

INTRODUCTION

The future success of any exploration program will largely depend on the examination of large potential areas and the rapid evaluation of all possible available data, geophysical, geochemical, and remote sensing, from the present and past exploration programs. The presence of a thick overburden and vegetation cover is usually considered to be a hinderance in geologic studies. However, it is believed that two thirds of the land area which is covered with moderate to heavy vegetation and probably thick overburden would provide the future supplies of minerals. Therefore, it is imperative that multi-media data such as geological, geochemical, biogeochemical and remote sensing be collected for the areas of interest, and their usefulness be evaluated rapidly.

Whenever multi-media data sets are available, rather than trying to estimate individual parameters separately as represented by each data sets, there is an attempt to integrate these different data sets for a meaningful relationship. The correlation among different variables in a particular data set and the usefulness of the individual data sets can be readily assessed if the hierarchical relationship among the various parameters measured can be readily determined. This would allow the removal of redundancy and more optimum use of the information extracted from these data sets. In this study, multivariate statistical techniques, such as cluster analysis and principal component analysis have been applied to determine the hierarchical relationships among the different variables in a data set and the relationships among different data sets.

STUDY AREA

The study area is centered on the Farley Lake, in northern Manitoba. It forms a part of the Agassiz Metallotect, and lies in the northern belt of the Lynn Lake greenstone belt. The bedrock is composed of mafic to intermediate volcanic flows, flow breccias, mafic tuffs, pyroclastic breccias, and chemical sedimentary rocks (Gilbert et al., 1980). The sedimentary rocks include mostly the iron formation, which are made up of banded chert and magnetite, along with some detrital argillite and siltstone. The iron formation is interbedded with a characteristic high Mg-Ni-Cr basalt. This iron formation is believed to produce a large scale, regional magnetic anomaly roughly extending east-west, which was used to delineate the Agassiz Metallotect (Fedikow, 1984).

The area is covered with surficial deposits of glacial and post-glacial origin, mostly by glacial till, glaciolacustrine silts and clays, and littoral sand and gravel. More than 60 percent of the area is covered with fen and bog veneer (Nielson and Graham, 1985). At Farley Lake, distinctive glacial dispersion trains have not been observed, and their absence may be attributed to the relatively short distance of dispersion of the local component of till, relatively large sample spacing as compared to the size of anomalies, or to the thickness of overburden present.

TYPE OF AVAILABLE DATA

The Farley Lake area due to its promising economic potential for a gold deposit has been extrensively surveyed both by federal and provincial governments, by using surface and remote sensing methods.

The remote sensing data includes the airborne multi-detector electro-optical imaging scanner (MEIS-II) data. These data provide the spectral reflectance information in different wave length windows,and were obtained to detect the signs of vegetation stress due to the abnormal concentrations of certain elements. The airborne remote sensing data have an advantage over the satellite borne data, as it provides better spatial resolution and the different spectral windows can be easily selected.

The surface data sets include geochemical and biogeochemical data. For geochemical data (Fig.1), basal till samples were collected and the non-magnetic heavy minerals fractions (specific gravity greater than 2.95) were analyzed for eight elements, namely Cu, Pb, Zn, Co, Ni, Cr, Mn, and Fe; whereas for the clay-sized fractions nine elements were analyzed that is Cu, Pb, Zn, Co, Ni, Cr, Mn, Fe and As (Nielson and Graham, 1985). The biogeochemical data consists of the analysis of black spruce (Picea mariana) needle samples collected at grid locations (Figure 2), and were analyzed for six elements, that is Cu, Cr, Zn, Ni, Mn and Fe.

ANALYSIS TECHNIQUES

Techniques of multivariate statistical analysis have been applied in this study, and include the cluster analysis and principal component analysis, which are based on the satisfactory definition of similarity between the samples and variables. A number of similarity and dissimilarity coefficients can be calculated ranging from correlation coefficient to a simple nonquantitative subjective measures, and therefore the choice of an appropriate similarity coefficient is important for a meaningful interpretation. Similarity is calculated by making pair-wise comparisons, and when a series of comparisons are made between all possible pairs, a square matrix of similarity coefficients is obtained. In this study Pearson-product moment correlation was used, which is expressed as:

$$r_{jk}=\frac{\Sigma_{i=1}^{n}X_{ij}X_{ik}-(\Sigma_{i=1}^{n}X_{ij}\Sigma_{i=1}^{n}X_{ik})/n}{\sqrt{\{\Sigma_{i=1}^{n}X_{ij}^{2}-[(\Sigma_{i=1}^{n}X_{ij})^{2}/n]\}\{\Sigma_{i=1}^{n}X_{ik}^{2}-[(\Sigma_{i=1}^{n}X_{ik})^{2}/n]\}}}$$

where r_{jk} is the correlation coefficient between two variables X_{ij} and X_{ik} and n is the number of data points. The correlation may be made between each pair of variables (R-mode) or between different pairs of samples (Q-mode).

Cluster Analysis: Parks (1966) states that 'cluster analysis is a simple form of correlation analysis, a method of searching for relationships in a large symmetrical matrix'. On the basis of computed similarity coefficient, a nucleus of cluster is being formed by joining the samples (Q-mode) or the variables (R-mode) with the highest similarity coefficient, and gradually adding more samples or variables depending on the situation as the similrity coefficient is lowered. Eventually other clusters are initiated until all the samples or variables are linked. A dendrogram is usually the end product of clustering which is a two-dimensional diagram showing the hierarchical relationships. Some of the characteristics of dendrograms are as follows: the number of clusters defined increases as the criteria of similarity are increased; the discontinuities or 'natural breaks' in the groups are readily apparent; it can be used to eliminate the highly correlated or redundant variables, thereby allows the expression of a large number of variables in terms of fewer interpretable ones; furthermore, the anomalous samples because of their uniqueness in the group would appear to be the least correlated.

Principal Component Analysis: According to McCammon (1966) 'principal components provide an effective method for reducing the number of variables in large-scale correlation studies'. The purpose is to determine the relationship among a number of interrelated variables in terms of fewer,composite of the original variables which are pair-wise uncorrelated. Therefore principal components are the linear combinations of original variables obtained such that the successive principal components have a smaller variance, and usually first few of the principal components account for a large part of the total variance. The coefficients of the principal components are termed principal components loadings, and the measurements of the principal components on each of the samples are known as principal component scores. These principal component scores provide the orthogonal projection of the samples on to the axes defined by the principal components, which can be displayed as scatter diagrams.

The consistency of results obtained from cluster analysis can be verified by comparing them with the results of principal component analysis. In Q-mode cluster analysis, a group of samples form a cluster on the basis of similarity and the different groups represent the natural breaks in the data set, which are characterized by the predominance of certain combination of interrelated variables. The results of R-mode principal component analysis should display the same linear combinations of variables (principal components), and therefore, on the basis of these principal components one should be able to determine the different groups of the samples.

INTERPRETATION OF RESULTS

A general flow chart showing the data processing

steps for this study is shown in Figure 3, details of which are given by Singh et al. (1988). The biogeochemical and geochemical map data were digitized, and each sample location was assigned longitude and latitude value. For the airborne remote sensing image data, corrections were applied to individual images where necessary, and a mosaic image was produced by combining a few of the images (Figure 4). The individual pixels of the mosaic were also assigned a longitude and latitude coordinates by comparing it with the topographic map. This allowed different data sets to be registered to a common base and thus correlations could be calculated. However, the basal till geochemical samples were found to be absent in the vicinity of biogeochemical sample grids, therefore correlations were calculated between the airborne remote sensing image data and geochemical data, and the remote sensing data and biogeochemical data.

The z-score transformations were applied to the data matrix, column wise for the cluster analysis, so that the mean value of each column variable becomes zero, and the standard deviation equals to one. This type of tranformation is necessary so that different types of data can be compared, as this does not alter the distribution, however, the effects due to different measurement units are eliminated. A matrix of correlation coefficient was calculated, and unweighted-pair group method was used for clustering. In this situation, the first linkage is formed between the most similar members and the subsequent additions are made at the average similarity of all the members in that cluster. This method weighs each member in the cluster equally and the results showing hierarchical relationships are displayed as a dendgrogram (Figure 5). Similarly, correlation matrix was used to calculate the principal components, rather than the variance-covariance matrix, again this is to eliminate the influence due to units of measurement.

Biogeochemical Data

The ash content of the black spruce needle samples, in general displays a negative relationship with most of the elements analyzed, except at the biogeochemical grid location 3, where there is a positive correlation between the ash content and the zinc concentration. This general negative correlation of the ash content with the elemental concentrations may be due to the closed nature of the system. At grid location 1, copper displays some correlation with Ni, Mn, and Fe, whereas the Zn and Cr show lack of correlation with these elements. At grid location 2, Fe and Ni have a positive correlation, which in turn is moderately related to Cu, whereas Fe and Ni lack significant relationship with Mn. At these locations the MEIS-II data does not indicate significant correlation with any of the elements analyzed. The concentrations of Cu, Mn, and Fe in the black spruce needles display good correlations at the biogeochemical grid location 3. Furthermore, the band2 of MEIS-II data indicates good correlation with the Fe, and the band1 and band2 show a moderate relationship with the copper concentrations.

However, the correlations of the MEIS-II and biogeochemical data at the grid location 3 should be interpreted with caution, because the number of samples is small as compared with the other biogeochemical grids. Moreover, an examination of correlation matrices and dendrograms for the biogeochemical data reveals that some of the element pairs in different biogeochemical grids do not have the same degree of linear relationship. This makes the interpretation of biogeochemical anomalies difficult. Some of the factors which may have pronounced effect on the accumulation of elements in plants include: health of the plant, depth of the root system, age of the plants or organs sampled, pH of the soil, drainage, rainfall, availability of elements, variable shading, and the antagonism of other elements. A lack of correlation of the MEIS-II data with the biogeochemical data is probably due to the complex nature of spectral reflectance from plants. This is because the reflectance from plants depends on the number and nature of leaves arrangement, spectral characteristics of trunk and stalks, solar zenith and azimuth angle, and the reflectance properties of the background, such as soil, rocks, understory shrubs,and vegetation litter.

Geochemical Data

In the clay-sized fraction of geochemical data, element Zn is highly correlated with the ferrous group of elements, namely Fe, Cr, Ni, Co, and Mn. This can be interpreted as the mobility of zinc is limited by the organic activity and this tends to precipitate with limonite (Hawkes and Webb, 1962). Cobalt displays a strong relationship with Mn, Fe, and Cr. However, Cu and As show fairly low correlations with other elements analyzed for the clay-sized fraction of the till. This is probably due to their high mobility in acidic solutions, which are likely to exist in areas with decaying organic activity. Whenever this acidic solution mixes with fresh water, it would result in a change of the pH and thus cause the precipitation of certain elements. The MEIS-II spectral reflection data do not show any significant correlation with the concentration of elements in the clay-sized fraction of the till data.

The PCA indicates that the first principal component, PC1 accounts for about 37.3 percent of the variations in the clay-sized fraction of the till data, and this is mostly due to the ferrous elements and the concentrations of zinc. The PC2 explains about 13.6 percent of the variation, and that is due to negative loadings of Cu and As, and the positive loadings of Co and Cr, and to some extent of Mn. The PC3 accounts for about 12 percent of the variation, and this is because of the positive loadings of Co and Cr and the negative loadings of Pb and to some extent Zn.

The heavy mineral fractions of basal till have been believed to be derived from the underlying rock formations in the Farley Lake area (Nielson and Graham, 1985). In this data, Fe has strong correlations with Mn and Ni, and also with Cr. The element Zn shows a moderate relationship with the ferrous elements analyzed. The base metal elements Cu and Pb display poor correlations with the ferrous elements (Figure 5, 6, and 7b). Furthermore, the MEIS-II spectral reflectance data do not show any significant correlation with the geochemical nature of the heavy mineral fractions of the basal till data.

When the results of cluster analysis are compared with PCA, the PC1 explains about 44.5 percent of variation in the data, and most of this is due to ferrous elements and Zn, whereas PC2 accounts for about 20.5 percent of the variation and that is because of Ni and Cr, and Pb. In the samples when there is an increase in the Ni and Cr concentrations, there appears to be a decrease in Pb concentration. The PC3 explains about 12.5 percent variation of the data, which is again mostly due to positive loadings of Ni and Cr, and negative loading of the Cu.

Aeromagnetic Data

The airborne magnetic data are available as vertical field gradiometer rather than as total field magnetic data, therefore it can not be compared directly with the other available data sets by coregistering it to the common base of longitude and latitude coor-

dinate system. Furthermore, magnetic data were collected with a vastly different spatial resolution and direct correlation with the other data could not be easily accomplished. However, it was interpolated to a regular grid and a gray level map was obtained (Figure 7a). The high gradient values in western half of the study area indicates the presence of highly magnetized formations beneath the surficial glacial deposits. These anomalous features represent either a complex geological structure or a simple geological formation with highly magnetized material distributed irregularly in the matrix. A number of quantitative map comparison methods are available (Davis, 1973), but a visual comparison of the airborne magnetic field gradiometer data with the iron concentration in heavy mineral fraction of the basal till (Figure 7b), and the available geology map does not show much similarity. In the Farley Lake area, overburden is thick and no significant correlation between geochemical and airborne field gradiometer data was expected (Figure 6 and 7). Nevertheless the strong gradient anomaly in the west is expected to represent subsurface geological magnetic features. As it is generally believed that the features on the magnetic field gradient map represent subsurface boundary of basement lithologic units with varying degree of magnetization, the gray level map of this data suggests that the available geology maps may have to be modified with detailed field verification.

CONCLUSIONS

The study of Farley Lake area suggests that airborne remote sensing data have low correlation with the biogeochemical and geochemical data. However, there appears to be some positive relationship between iron concentration and band2 of the MEIS-II data, and band1 and band2 indicate moderate correlation with MEIS-II data at biogeochemical grid location 3. There is high correlation among ferrous elements and zinc in the basal till geochemical data. The elemental concentrations in the clay-sized fractions are effected by the ionic mobility whereas the elemental concentrations of the heavy mineral fractions represent resistant component of the till geochemical data, which are not effected by the mobility of ions in aqueous phase. The nature of the sampling plan adopted in the Farley Lake area makes it difficult to determine that the anomalous concentrations in till samples correspond to the anomalous concentrations in biogeochemical data. The interpolated airborne magnetic field gradiometer data do not have any significant correlation with geochemical data. However, airborne magnetic data suggests that the available subsurface geology maps may require modifications to account for the anomalous features.

A more general conclusion of this study is that whenever multi-media data are available, or when it can be obtained by digitizing data from previous exploration programs, it can be analyzed rapidly by digital processing. The correlation of different data set using the type of approach outlined in this study allows a rapid and effective evaluation of different parameters measured in an individual data set and the different data sets, and thus provides the first step for systematic integration of exploration data.

REFERENCES

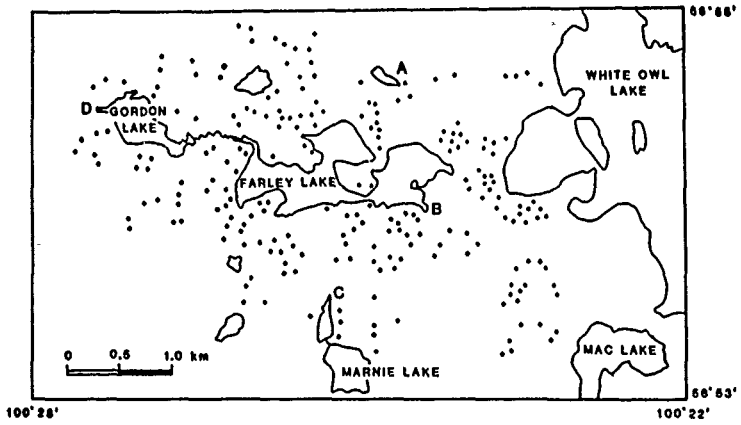
1. Davis, J.C.,'Statistics and data analysis in geology', John Wiley, 550 p., 1973.
2. Fedikow, M.A.F., 'Preliminary results of biogeochemical studies in the Lynn Lake area', Geological Services, Manitoba Energy and Mines, Open File Report, OF84-1, 1984.
3. Gilbert, H.P., E.C. Syme, and H.V. Zwanzig, 'Geology of metavolcanics and volcanoclastic metasedimentary rocks in the Lynn Lake area', Manitoba Mineral Resource Division, Geological Paper, GP80-1, 1980.
4. Hawkes, H.E., and J.S. Webb, 'Geochemistry in mineral exploration',Harper and Row, 415 p., 1962.
5. McCammon, R.B., 'Principal component analysis and its applications in large scale correlation studies', Journal Geology, Vol. 74, No. 5, p. 721-733, 1966.
6. Nielson, E., and D.C. Graham, 'Preliminary results of till petrographic and till geochemical studies at Farley Lake', Geological Services, Manitoba Energy and Mines, Open File Report, OF85-3, 1985.
7. Parks, J.M., 'Cluster analysis applied to multivariate geological problems', Journal Geology, Vol. 74, No. 5, p. 703-714, 1966.
8. Singh, V., W.M. Moon, and M. Fedikow, 'An integrated remote sensing study of mineralized zones associated with the Agassiz Metallotect, Lynn Lake, Manitoba (I) [MEIS-II and Biogeochemical study]', Geological Services, Manitoba Energy and Mines, Open File Report, OF88-3, 1988.

ACKNOWLEDGEMENTS

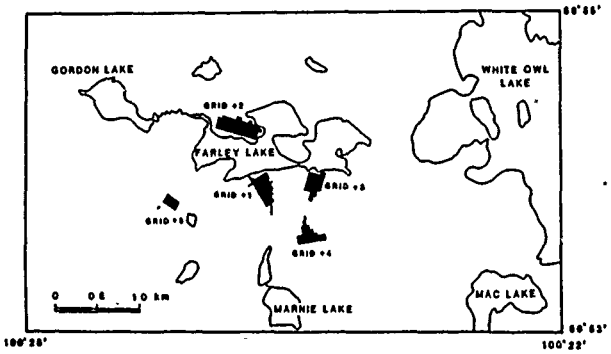
This study is a part of the Agassiz Metallotect project, funded through the Mineral Development Agreement (MDA) signed by the Province of Manitoba and the Federal Government of Canada. Most of the image processing and digital processing work was carried out at the Geophysical Image Processing Laboratory of the University of Manitoba. Some time on the mainframe computer was provided by the Faculty of Science, University of Manitoba.

FIGURE CAPTIONS

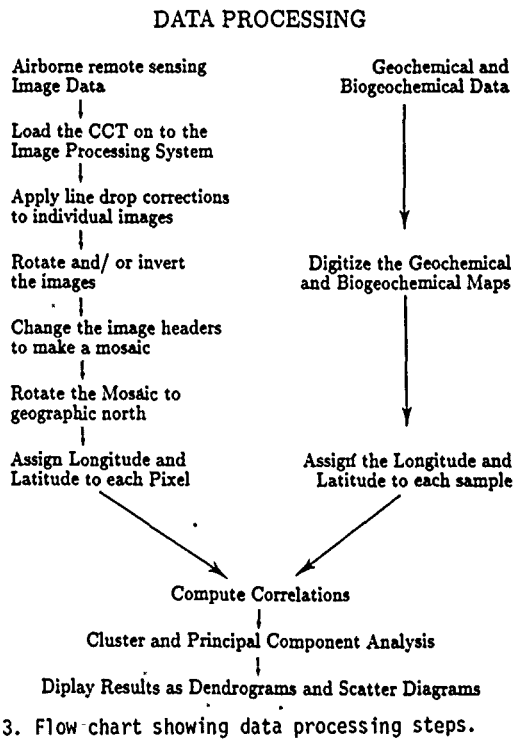
1. Location of geochemical samples at the Farley Lake.
2. Location of biogeochemical sample grids.
3. Flow chart showing data processing steps.
4. Mosaic of band1 of MEIS-II data, Farley Lake produced from four separate images.
5. Dendrogram of heavy mineral fraction of till geochemical data showing hierarchical relationships.
6. Gray level maps of element concentrations in heavy mineral fraction of till geochemical data (normalized to 16 gray levels): (a) Manganese (302-883 ppm), (b) Nickel (3-13 ppm), (c) Copper (3-26 ppm).
7. Gray level map (normalized to 16 gray levels): (a) Aeromagnetic vertical field gradiometer data (min -22.82 to max 88.27), (b) Iron concentration in heavy mineral fraction (0.986-3.155 percent).



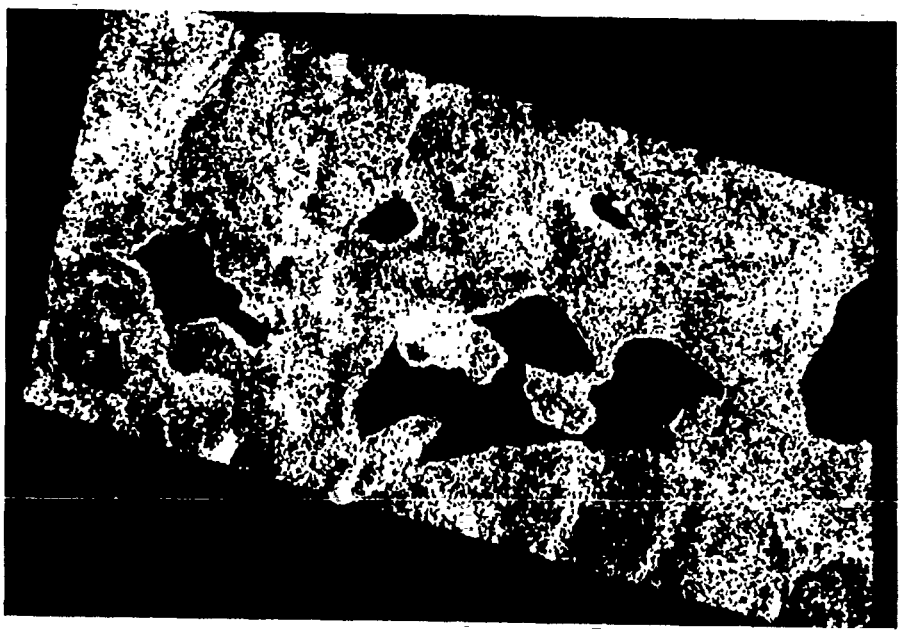
1. Location of geochemical samples at the Farley Lake.



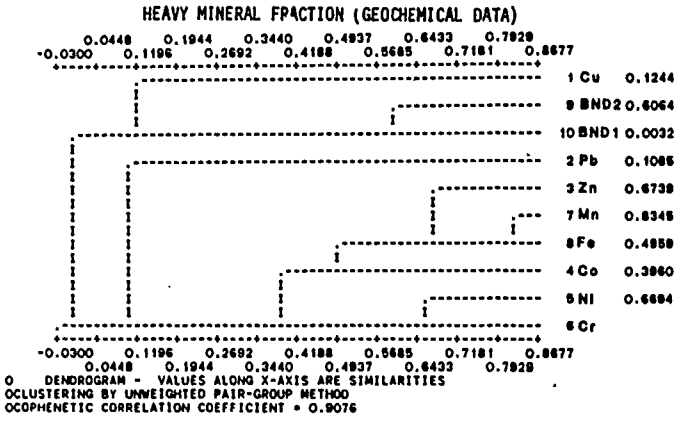
2. Location of biogeochemical sample grids.



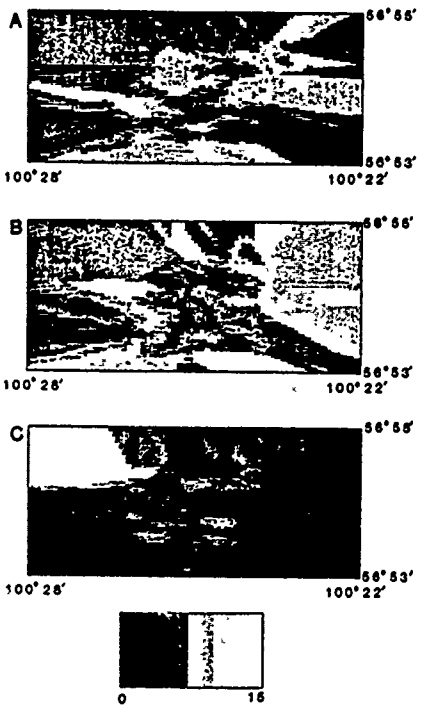
3. Flow chart showing data processing steps.



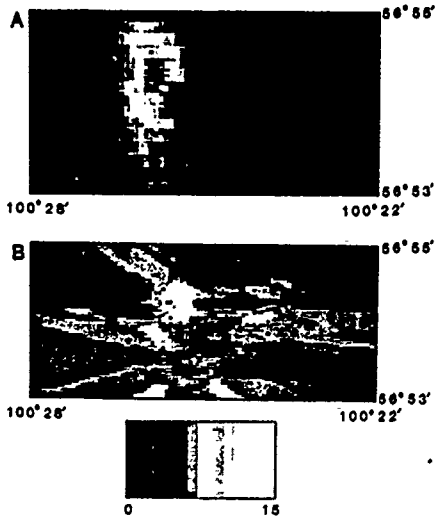
4. Mosaic of band1 of MEIS-II data, Farley Lake produced from four separate images.



5. Dendrogram of heavy mineral fraction of till geochemical data showing hierarchical relationships.



6. Gray level maps of element concentrations in heavy mineral fraction of till geochemical data (normalized to 16 gray levels): (a) Manganese (302-883 ppm), (b) Nickel (3-13 ppm), (c) Copper (3-26 ppm).



7. Gray level map (normalized to 16 gray levels): (a) Aeromagnetic vertical field gradiometer data (min -22.82 to max 88.27), (b) Iron concentration in heavy mineral fraction (0.986-3.155 percent).

SAR FOR GEOLOGY IN ALBERTA
(Surficial Geology/ Aggregate Exploration)

J.K. Hornsby & R.V. Dams
INTERA Technologies Ltd., Calgary, Alberta

I.A. Sutherland & A.M. Hougham
Alberta For. Lands and Wildlife, Edmonton, Alberta

Abstract

The purpose of this study was to demonstrate and investigate the use of synthetic aperture radar (SAR) imagery for geologic applications in Alberta. Regional structural-physiographic mapping, oil and gas exploration and surficial geology/aggregate exploration applications were investigated. This paper reports on methods and results of the surficial geology/aggregate exploration applications. The SAR data were digitally enhanced and analyzed, and hardcopy outputs were interpreted and compared with available geologic information. Results indicate that SAR provided useful geologic information for mapping of surficial geology and the identification of areas of high aggregate potential. Optimal digital enhancement techniques were identified.

KEYWORDS: SAR, Aggregate, Surficial Geology

1. INTRODUCTION

Remote sensing studies have been carried out for a wide variety of geologic applications: morphological and lithological mapping; recognition of tectonic structures; mineral and energy exploration; waste disposal and plant siting; and data input to the development of models for regional and continental tectonics (Goetz and Rowan 1981; Hardjoprawiro and Irsyam 1984). Yetabe et al. (1985) noted increased geological use of remote sensing in Canada, yet in their estimation, the technology was still under utilized.

While satellite sensors (visible and near infrared) are useful for geologic applications, airborne and spaceborne radar can provide additional information through: (1) acquisition under virtually any weather or light conditions; (2) sensitivity to surface roughness and dielectric properties; and (3) side-looking orientation which enhances terrain features, due to shadow and differential backscatter (Jackson, 1979). Thus, the interest in radar for geologic applications has grown rapidly, with both airborne and spaceborne systems utilized. The objective of this study was to evaluate and demonstrate the utility of an active imaging microwave sensor--synthetic aperture radar (SAR)--for geologic applications in Alberta. This paper reports on the application for mapping surficial geology and aggregate potential mapping.

1.1 Study Area

The study area was located in the vicinity of Fort McMurray (Tp.88-91, R.8-10) in northeastern Alberta (Figure 1). This area has gently undulating terrain incised by the Athabasca and Clearwater Rivers and their tributaries. The northern boreal forest includes pine forests on well-drained areas interspersed with poorly drained spruce and shrub bogs.

Bayrock and Reimchen (1973) mapped the surficial geology of the area at a scale of 1:250,000. Gully and creek valley deposits form a thin colluvium over bedrock on valley slopes, and thin alluvial materials occur along streams and the Athabasca and Clearwater Rivers. Stream alluvium (sand, silt, and clay deposits) is most extensive along the Athabasca River north of the Clearwater confluence, and along the Clearwater River. An area of fine-grained aeolian sand mapped north of the Clearwater River was developed from glacial outwash. Glaciolacustrine clay, silt, and sand occur extensively in the study area. Thick-bedded clay and silt with minor sand occurs over much of the level area south of the Clearwater and Athabasca Rivers and north of the Clearwater River. West of the Athabasca River a glaciolacustrine unit contains pebbles and till-like layers. Glaciofluvial meltwater channel outwash, in the central portion and south of the Clearwater River, are thin, flat sand and gravel deposits along the bottoms and terraces of meltwater channels. Lastly, the thin kinosis till-ground moraine (loam with numerous pebbles and boulders), has a rolling topographic surface expression.

2. METHODS AND MATERIALS

This study used common digital image analysis techniques and hardcopy interpretive methods which can be easily utilized by industry and government geologists.

2.1 Remote Sensing Data

SAR data were acquired on December 24 1985 from an altitude of 9,148m. (30,000 ft.) above ground level. The data were collected in high resolution mode flying in a northerly direction and looking west. A Landsat Thematic Mapper (TM) data set was acquired from a Landsat 5 pass on 18 September 1987 consisting of bands 3 (0.63 to 0.69 um), 4 (0.76 to 0.90 um), and 5 (1.55 to 1.75 um). This red, near-infrared, shortwave-infrared combination was identified as optimal for geobotanical investigations (Hornsby & Bruce 1986).

2.2 Image Analysis

Image analysis consisted of four stages: geometric correction, image integration, image enhancement, and image output (Figure 2). The ARIES II Image Analysis System at the Alberta Remote Sensing Centre was used for all digital image analysis tasks.

Geometric correction of the SAR imagery to remove image distortion involved the interactive selection of ground control points on the displayed image and on the corresponding UTM basemap. Residual errors in excess of two pixels (25m) were eliminated prior to resampling. Image to image registration was performed to correct the TM data, using the corrected SAR as the base. The result would allow for integration of the two datasets. The TM data was resampled to 12m. in order to preserve the resolution of the SAR during the integration phase.

The radar and Landsat datasets were merged by mathematically combining the radar data with each of the Landsat bands. This allowed all three bands of Landsat data as well as SAR to be retained in the final merged image. Two approaches were tried: (1) addition of the radar data into Landsat TM bands 3,4 and 5 with a radiometric rescaling of the data; and (2) multiplication of each of the TM bands by the radar data with radiometric rescaling of results. With the relatively low dynamic range of radar response in the Fort McMurray area due to the low topographic relief, the multiplication method (SAR x TM Band) provided the best output image. The radar data were clearly present but did not mask the TM data; with the addition method the radar was masked in the final output image.

Raw SAR data are very dark and have little contrast when displayed. In order to maximize interpretability, a linear contrast stretch increased the apparent contrast in the image by spreading the data over the full dynamic range (0 to 255 grey levels). Directional convolution filters were applied to the imagery to enhance changes in pixel values along straight of curved geologic features. In combination with a linear contrast stretch enhancement, directional filters of the principal compass directions (N-S, NE-SW, E-W, SW-NE) are good enhancements for extracting geologic information related to linear and curvilinear features (Moore and Waltz 1983; Hornsby et al. 1984). A filter size of 5 by 5 pixels provided optimum results; a smaller filter size did not as effectively reduce land cover noise, and a larger size eliminated some of the higher spatial frequency features.

The enhanced images were output to two different hardcopy image recording systems: (1) FIRE laser image recorder; (2) IMAPRO image recorder.

2.3 Image Interpretation

The enhanced images were enlarged and printed at a 1:50,000 scale. Interpretation was carried out on acetate overlays over the hardcopy images. Interpretation involved standard photo interpretive techniques (tone, texture, pattern, shape, size and association characteristics of the image).

Interpretation of the data set was carried out in two stages. The first stage involved the mapping of surficial geologic units. The interpretation was based primarily on the SAR-TM enhancement enlarged to a scale of 1:50,000. Directional filter enhancements and the contrast stretch of the radar data were also used during the interpretation process. The second stage consisted of producing a second overlay with shaded areas indicating areas of high, medium, and low potential. In addition, specific targets were identified for follow-up ground work in locating deposits which are suitable for cost-effective excavation.

3.0 RESULTS

While the Landsat TM provided detailed vegetation information which is strongly related to site drainage characteristics, the SAR enhanced subtle topographic features indicative of previous glacial activity and resultant surficial deposits.

3.1 Surficial Geology Mapping

Five major surficial deposit types were mapped from the imagery. These included recent alluvial and erosional deposits, and Pleistocene-aged glaciofluvial, glaciolacustrine, and glacial deposits. The glaciolacustrine and glaciofluvial units were further sub-divided on the basis of differing image characteristics. Specific descriptions of the major units were derived from supporting published material (Bayrock and Reimchen, 1973). Till-ground moraine mapped in the eastern half of the study area north of the Clearwater River appeared on the imagery as a distinctive reddish-purple colour with homogenous texture. The areas were more discontinuous than previous mapping at a scale of 1:250,000 indicated. The reflectance characteristics suggest coarser materials than the surrounding glaciolacustrine and glaciofluvial deposits.

The glaciofluvial deposits were divided into two types: (1) a silt, sand and gravel unit (2a); and (2) a clay, silt, sand unit (2b). These were distinguished on the imagery primarily on the basis of tone, texture, and pattern. Unit 2a had a reddish-brown colour containing curvilinear patterns interpreted as glaciofluvial channels (supported by the work of Bayrock and Reimchen (1973)). The radar data enhanced the subtle channel indicators while the multispectral data aided in defining the boundaries between the two units. The channel patterns detected by both the radar and the TM sensors showed the two units east of the Athabasca River and south of the Clearwater River to be glaciofluvial in origin.

Glaciolacustrine deposits occurred extensively in the study area. Three types of deposits were identified by comparing the mapped units with those interpreted from the imagery. A clay-silt type (3a) (seen as light brown) and a sand-silt type (3c) (seen as dark brown) could be separated on the basis of differing tones of brown. The third unit (3b) was not described by Bayrock and Reimchen, but was mapped from the imagery due to its contrasting colour of green compared to the light and dark-brown of the other two units. Several other undifferentiated glaciolacustrine (3) areas were mapped. These areas were thus considered because of their association with other glaciolacustrine units, even though they appeared slightly different in colour.

Recent alluvial deposits are extensive along the Athabasca and Clearwater River floodplains. Meander scars and channel bars are readily identifiable on both the radar and digitally merged SAR-TM images along the river. The materials consist of sand, silt, and clay (Bayrock and Reimchen, 1973). Meander scars occur as

distinct tonal changes in the imagery due to changing cover types controlled by the feature morphology. The arcuate shape is also characteristic of meander scars and oxbows. Recent erosional deposits along gullies, creeks and river valleys can generally be described as thin colluvium over bedrock on valley slopes. These do not represent significant accumulations of materials but are distinct due to the different geomorphic processes involved in their formation.

3.2 Aggregate Potential

Areas of high, medium, and low aggregate potential were mapped based on the surficial geologic units present. Areas of glaciofluvial deposition were considered to be of highest potential due to the coarser nature of the materials as compared to glaciolacustrine deposits. The high energy depositional environment also provided the opportunity for concentrations of deposits to occur in channels. These channels, located in Unit 2a, were identified as the targets for concentrated field exploration. The glaciofluvial Unit 2b was considered of moderate potential. Channels within this unit were also singled out as potential exploration targets. The areas of lowest aggregate potential were in the recent alluvial deposits along the river floodplains. These areas represent a source of sand but environmental concerns would likely limit the extraction of materials from these areas.

4.0 Conclusions

Interpretation of surficial geology and aggregate potential from the imagery indicates the radar and multispectral data are appropriate sources of information of this type. The combined SAR-TM enhancement was more useful than SAR alone because the TM provided additional information in areas of homogeneous radar response. The two data types are complimentary: SAR provides more information related to topographic features, while the TM is more sensitive to vegetation types and distribution, and provides information related to subtle changes in materials as reflected in the vegetation cover. Combined with conventional field investigations , this imagery would provide a means for rapidly and cost-effectively mapping. Subsequent field efforts could be optimized through preliminary image interpretation. A final refined map could then be produced from the preliminary interpretation based on field investigations.

The identification of targets for aggregate exploration does not necessarily require mapping of the surficial geology, although a knowledge of the geomorphology is required to make educated judgements on the location of potential aggregate sources. The imagery provides information on subtle changes within a given surficial unit that is not available through conventional photography. This does not suggest the exclusion of air photos in the exploration process, but rather their use in a complementary role.

Conclusions were made regarding the advantages and disadvantages of using SAR for geologic applications. The most significant advantage is that SAR provides geologic information not always available from more conventional visible and near-infrared remote sensors.

Advantages of SAR:

1) provides geologic information available from other remote sensors or geophysical techniques;
2) provides a continuous record and a regional view;
3) mosaicing provides large area mapping capability;
4) data may be collected independent of weather and light conditions;
5) relative low cost compared to other data types;
6) most useful in remote areas with little geologic information.

Disadvantages of SAR:

1) not collected on a regular basis;
2) geometric distortions requiring correction;
3) large data volume.

REFERENCES

Bayrock, L.A., and T.H.F. Reimchen, 1973:
Surficial Geology, Waterways NTS 74D; Alberta Research Council (with descriptive notes).

Goetz, A.F.H. and L.C. Rowan, 1981:
Geologic Remote Sensing, Science, v.211, pp781-791.

Hardjoprawiro, S. and M. Irsyam, 1984:
Application of geological remote sensing in Indonesia; Proc Remote Sensing for Geological Applications Seminar; Beijing, China, pp.47-58.

Hornsby, J.K. and B. Bruce, 1986:
Regional geobotany with TM: A Sudbury case study; Proc. 10th Cdn. Symp. on Remote Sensing, Edmonton, pp.601-610.

Hornsby, J.K., A.N. Rencz, and G.F. Bonham-Carter, 1984:
Comparison of techniques for enhancing geologic lineaments from Landsat data with application to the Nahanni River area, Yukon Territory; Proc. 9th Cdn. Symp. on Remote Sensing, St. John's, Newfoundland, pp.387-395.

Jackson, P.L., 1979:
Multichannel SAR in geologic interpretation: An appraisal; et Propulsion Laboratory, Pasedena, California, JPL Publication 80-61, pp.223-250.

Moore, G.K. and F.A. Waltz, 1983:
Objective procedures for lineament enhancement and extraction; Photogrammetric Eng. and Remote Sensing, v.49, pp.641-647.

Yetabe, S.M. and A.G. Fabbri, 1985:
The application of remote sensing to Canadian exploration: Promising and yet unexploited. Proc. 1st Atlantic Canada Symp. on Remote Sensing and Geogr. Info. Systems, Lawrencetown, Nova Scotia, pp.5-8.

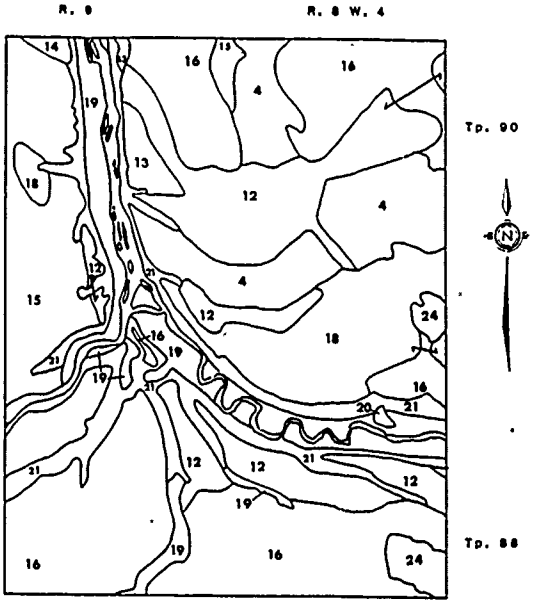


Figure 1. Surficial geology of the Fort McMurray study area (Bayrock and Reimchen 1973)

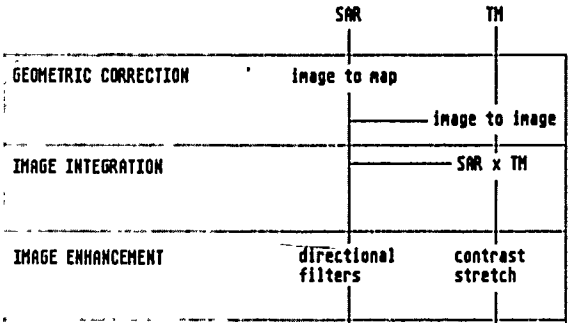


Figure 2. Image Processing Methodology

USING PASSIVE, THERMAL REMOTE SENSING TECHNIQUES FOR DETECTING SUBSURFACE GRAVEL ACCUMULATIONS IN VEGETATED, UNCONSOLIDATED SEDIMENTARY TERRAINS

GREGORY S. BURNS

NATIONAL AERONAUTICS AND SPACE ADMINISTRATION
JOHN C. STENNIS SPACE CENTER
SCIENCE AND TECHNOLOGY LABORATORY
STENNIS SPACE CENTER, MISSISSIPPI 39529-6000
(601) 688-1911
FAX 688-1925

DOUGLAS E. SCHOLEN

USDA FOREST SERVICE
SOUTHERN REGIONAL OFFICE
SOUTHERN DIRECTORATE
1720 PEACHTREE ROAD
ATLANTA, GEORGIA 30367
(404) 347-2349

ABSTRACT

Multiband radiometric data from an airborne imaging thermal scanner are being studied for use as an exploration tool to find buried gravel deposits. The techniques will be based upon measuring relative differences in the thermal properties between gravel-laden targets and the surrounding gravel-less background. These properties will be determined from modeling the spectral radiant emittance recorded over both types of surfaces in conjunction with ground measurements of the most significant heat flows above and below the surface. Thermodynamic properties of sampled materials from control sites will be determined, and diurnal and annual subsurface heat waves will be recorded. Thermal models that account for heat exchange at the surface as well as varying levels of soil moisture, humidity, and vegetation are needed for adaptation and modification to simulate the physical and radiative environments of this region. The models will attempt to relate the thermal and molecular properties of the materials to the apparent temperatures derived from their spectral emittance. Details of preliminary findings from this first six months of the three year project are presented. These results are from initial efforts to duplicate data processing and analysis techniques between research partners and to establish a common baseline of understanding from which to proceed.

(Key Words: Soil Physics, Emission Spectrometry, Remote Sensing, Gravels, Thermal Models)

INTRODUCTION

This paper discusses research underway between the National Aeronautics and Space Administration, Science and Technology Laboratory and the U.S. Forest Service - Southern Region. This project was made possible by an award recieved through a response made to the NASA Office of Space Science and Applications research announcement for Remote Sensing Applications and Commercialization. The emphasis is to locate gravel deposits underlying certain National Forest lands on the Gulf of Mexico Coastal Plains so that a rationale can be developed for their proper management as a construction material in concert with environmental and ecological management considerations. This non-renewable resource is of major importance in the overall management of the national forests and as such represents a significant economic asset. Knowledge of its location and abundance is essential.

APPROACH

This study will determine if surveys using airborne, imaging thermal spectroradiometers are warranted for locating subsurface occurrences of these materials. A six-band, thermal infrared multispectral scanner (TIMS) that is capable of resolving less than 0.2 degree differences between objects will be used for the study. The six bands span the 8 to 12 micron wave lengths in the mid infrared portion of the electromagnetic spectrum. This area of the spectrum is particularly useful for identifying geologic materials. High absorption features, at 10 microns caused by extreme electron vibrations from molecular bond stretching in the Si02 molecule prevalent in the chert gravels, and similarly at 11 microns from bond stresses in the AlOH radicals of clays, are easily recognizable.

The thermal properties of gravels and soils are of equal importance to their spectral characteristics in providing clues to the locations of substantial deposits of gravels. The higher thermal inertia of gravel-sand-clay mixtures relative to adjacent fine grained soils reduces the temperature variation of the graveled bodies from day to night and likewise at depth on an annual basis from summer to winter. This characteristic causes the graveled bodies to show a lower temperature than adjacent soils on imagery acquired during early morning hours at the end of Summer when ground radiation is at a maximum.

Full solution of the problem will require determination of the benefit of preprocessing the data for spectroradiometric calibration, atmospheric path radiance correction, signature extendibility, heat flux simulation, geometric rectification, geo-location, and verification of prospects. Supporting field work includes comparative monitoring of subsurface thermal parameters in gravel deposits and non-gravel

soils for maximizing differentials between target and background radiant emission in order to time airborne data acquisition. Also, thermal emission spectra will be measured at ground level in conjunction with mineralogic and textural studies for site characterization.

PRELIMINARY RESULTS

These physical attributes were exploited when processing archived TIMS data that was collected over the Kisatchie National Forest in October of 1983. Software at both installations has been used to perform a density slicing technique that permits isolation of pixels containing candidate spectral signatures within the emission temperature range sought. Initially these signatures were identified through correlation with known gravel bodies. The presence of gravel at other locations is being verified by field reconnaissance and test hole drilling with a Mobile B-60 drill rig located on the Kisatchie National Forest near Alexandria, Louisiana.

Bands 1,3 and 6 were used to delineate the typical quartz signature, showing the absorption feature in band 3. While it was possible to correlate the TIMS data with areas of known gravel concentration, the same signature would also delineate other areas without gravel. Subsequent studies of the 6 band signatures of these same areas revealed that wide variations in bands 4 and 5 existed without effecting bands 1, 3 and 6. Further work with 6 band signatures succeeded in deleting many of the non-graveliferous areas from the images.

In all cases, the areas of known gravel occurrences showed the lowest temperatures due to their high thermal inertia retaining heat flow to the surface when compared to adjacent non-graveliferous areas. This distinction was found to be of prime importance since a sandy surface without gravel also produces a similar quartz signature. However, even with this distinction, a few non-graveliferous areas in the images matched the signature and surface temperatures of the areas with gravel. Field reconnaissance of these areas found them to be a thin layer of sand overlying a moist sandy clay. This combination provides a like signature and thermal inertia to the graveled areas.

A review of the thermal properties of the two materials revealed that although the thermal inertia is nearly the same, the diffusivity of the sandy clay is only one-third of the diffusivity of the sandy gravel. The thin layer of sand overlaying the sandy clay prevents the diurnal solar heat from penetrating more than a few centimeters, resulting in a cool ground surface in the early morning hours when the imagery was flown. This situation differs markedly from that of the gravel deposits since the sandy clay does not contain comparable heat. Studies are now underway to determine a means to differentiate between these two materials. Options currently under consideration include varying the time of flight to alter the effects of the diurnal temperature variation superimposed, and upon the seasonal variation, simultaneous processing of multi-flight images, when available, to capture times of differing thermal response.

Additional TIMS imagery acquired September 28, 1984, over portions of the Desoto N.F. in southeastern Mississippi and the Oakmulgee Division of the Talladega N.F. in central Alabama have permitted a comparison with results obtained from the Kisatchie imagery. The predominate mineral signature of the Desoto and Oakmulgee imagery differs strikingly from that of the Kisatchie. The absorption feature shifted from band 3 to band 4, and the peak value moved from band 5 or 6 to band 2, indicating the high iron content in the surficial quartz sands of those areas. Only a few isolated areas, where surficial sands have been removed by land management activities or river erosion, show the typical quartz signature similar to the Kisatchie imagery.

Comparison of Oakmulgee TIMS data with field data again showed lower surface temperatures in areas known to contain gravel. All of these areas showed their absorption feature in band 4. The only quartz signature present in the Oakmulgee area emanates from sand bars adjacent to the river. Additional field verification will be required to determine whether or not the gravel signature also includes areas of other materials with high thermal inertia. The low surface temperature feature at known gravel deposits was noticeably absent from the Desoto imagery. Temperature variations appeared to be influenced most by vegetation and surface moisture. A review of the rainfall records for the two areas offered a probable explanation. Rainfall was recorded on the September 22, 23 and 24 in the Desoto area just four days prior to obtaining the imagery. In contrast, the rainfall records for Oakmulgee showed a high forest fire index which indicates exceptionally dry ground conditions. This presence of surficial moisture in the Desoto area effectively overwhelms the thermal inertial differences provided by gravel deposits.

While these initial successes in correlating processed images with known gravel deposits show promise, some refinement of the process will be required before this method offers real advantage over traditional means of locating new gravel deposits. Many of the indicated gravel deposits in the images have proven to be only a few feet in thickness and, therefore, of no commercial value. Those areas already explored in the field are easily accessible. Many other prospects remain inaccessible without considerable work involving clearing of underbrush and/or construction of access roads. Some additional assurance of deposit size is required before land managers are willing to undertake this kind of effort. The alternative is to wait until other land management activities coincidentally provide access to an indicated area.

In addition to this problem of differentiating between significant and insignificant deposits, another problem exists because several known substantial deposits remain hidden in the imagery due to exceptionally heavy vegetation which holds heat above the ground through the

night and prevents the ground surface from being cooled. This type of problem may be only solved through repeated surveillance by aircraft or satellite, thus taking advantage of the effect of cyclic land management activities.

Locating indicated gravel deposits in the field is another problem. In order to find a specific small area within the vast, generally homogeneous expanses of forested country, a means other than visual recognition is required. Actual surficial appearances of a substantial gravel deposit differ little from adjacent areas. Determination of the exact field location of any pixel is possible only with accurate georeferencing of the imagery during flight along with use of a portable GPS receiver on the ground during surveys. GPS satellites currently in orbit are too few to allow georeferencing at optimum times for data acquisition except in coincidental cases. GPS equipment is still in the developmental stage, expensive and lacking user convenience. The inertial navigation system currently in use on the NASA Lear jet does not account for all of the aircraft motion that is distorting the image. All of these deficiencies must be successfully dealt with by replacement and or substitution before accurate ground location will be possible.

Approximate locations from unreferenced imagery can be obtained by comparison with standard USGS map quadrangles and aerial photography, but this method often results in confusion and doubt in the field, when surficial evidence is sparse. Accurate registration between ground and imagery is required before positive verification is possible.

BIBLIOGRAPHY

Carson, J.E. and H. Moses, "The Annual and Diurnal Heat-Exchange Cycles in Upper Layers of Soil", Journal of Applied Meteorology, Volume 2, June, 1963, pp. 397-406.

Elachi, C., "Solid-Surface Sensing: Thermal Infrared", from Introduction to the Physics and Techniques of Remote Sensing, Chapter 4, pp. 114-141. Charles Elachi, au., 1987.

Hunt, C.B., "CMate on and in the Ground", from Geology of Sorts, Chapter 5, pp. 85-98, Charles Hunt, au., 1972.

Kahle, A.B., "Surface Thermal Properties", from Remote Sensing in Geology, Chapter 8, pp. 257-274. B.S. Siegal and A.R. Gillespie, eds. 1980.

Kahle, A.B., "A Simple Thermal Model of the Earth's Surface for Geologic Mapping by Remote Sensing", Journal of Geophysical Research, Volume 82, Number 11, pp. 1673-1680, 1977.

Nerry, F., J. Labed and M.P. Stoll, "Emissivity Signatures in the Thermal IR Band for Remote Sensing: Calibration Procedure and Method of Measurement", Applied Optics, Volume 27, Number 4, pp. 758-764, 1988.

BACKSCATTER ANALYSIS OF AIRBORNE RADAR: IMPLICATIONS FOR BACKGROUND GEOBOTANY

J. Harris and J.K. Hornsby

INTERA Technologies Ltd./CCRS

ABSTRACT

Multi-channel airborne radar data acquired by CCRS and INTERA Technologies Ltd. is assessed for its value in improving a TM-based geobotanical classification of the Mazinaw Lake area in central Ontario. Radar backscatter for each radar pass (total of three) is analyzed, with respect to surface parameters such as topography and to radar parameters such as incidence angle and look direction. Topographic variables such as slope and aspect were calculated from a DTM and compared to relative radar backscatter using multiple regression techniques. A method for identifying pixels whose backscatter showed a strong correlation with topographic variables (slope and aspect) was developed. These pixels which showed strong topographic effects were used to create a "topographic mask" to compare with the geobotanical classification derived from the TM data.

KEYWORDS: geobotany, radar, DTM, radar look direction bias

1.0 INTRODUCTION

The Geological Applications Group of the Canada Centre for Remote Sensing (CCRS) has initiated a research project in the Mazinaw Lake area of central Ontario to evaluate a remote sensing data set, including airborne radar, TM data and a DTM (digital terrain model), for the mapping of geobotanical anomalies potentially associated with regional gold and base metal mineralization (Hornsby et al., 1988). Hornsby et al., 1987, using TM data and supervised classification, have produced a geobotanical classification of the study area.

The objective of this interim report is to evaluate the general contribution of airborne radar for defining regional geobotanical anomalies. This involves studying the terrain factors which are responsible for radar backscatter. This is accomplished by comparing the radar digital number (DN) values to the DTM data, particularly slope and aspect. The end result is an assessment of the contribution of topography on radar backscatter. Secondly, a method for identifying pixels whose backscatter show a strong correlation with topographic variables (i.e. slope and aspect) is developed. These pixels are used to produce a "topographic mask" which is applied to the TM-based classification produced by Hornsby et al., 1987 to isolate pixels where the influence of topographic variables on solar reflectance is minimal, thus helping to identify areas that are potentially anomalous with respect to mineralization. This particular paper represents a summary of the results achieved to date. Research is on-going and further results are forthcoming.

2.0 DATA PROCESSING

A sub-area of the Mazinaw Lake study area was defined (see Fig.1) which is covered by three radar images, two CCRS C-SAR images and one INTERA STAR-1 image. The two CCRS C-band images were acquired in Oct., 1987 at a resolution of approximately 20 m. in range and 10 m. in azimuth and formatted with 12 m. pixels. Polarization was HH and incidence angles ranged from 45° in the near range to 85° in the far range. The INTERA X-band image was acquired in Jan., 1987 at a resolution of 12 m. in range and 4 m. in azimuth and formatted with 12 m. pixels. Polarization was HH and incidence angles ranged from 59° in the near range to 82° in the far range. Figure 2 is a schematic diagram showing the look direction of each radar image and the associated slope enhancement zone (ie ± 20° from the normal to the radar look direction) as well as the sun azimuth direction for the TM data. Note that the look direction for the STAR-1 image and the sun azimuth for the TM imagery are almost the same.

Each radar image was co-registered to a 12 meter pixel size and then median filtered (5x5) to help reduce the effects of speckle. A DTM and TM imagery were also resampled to 12 m. pixels and registered to the radar imagery. All image processing was undertaken using a micro-based image analysis system (Easi/Pace from PCI). Correlation coefficients for each radar image pair were calculated using the image analysis system. The lowest correlation was between the C-SAR (line 4, pass 5 -L4P5) and the STAR-1 image (-.004) as they are of opposing look directions (approximately 180 degrees divergent) while the C-SAR (L1P2) and the STAR-1 had the highest correlation coefficient (.55).

3.0 ANALYSIS

This section consists of two parts:

1. comparison of the radar data to slope and aspect data derived from the co-registered DTM;

2. comparison of the classified TM imagery to the radar imagery and slope and aspect data.

3.1 Comparison of the Radar Imagery to Slope and Aspect

The effect of slope aspect is shown graphically in Figure 3 where slope aspect, which has been divided into 36 classes (10° intervals), is plotted against the correlation coefficient between radar DN and slope for each aspect class. It can be seen that a higher correlation between radar DN and slope exists for slopes that are oriented approximately ± 20° to the normal of the radar look direction. Furthermore, DN values are on average 15 to 20 higher for slopes oriented ± 20° to the normal of the radar look direction than slopes outside this zone, verifying the radar look direction bias (Harris, 1984). Also note the effect of opposing look directions for the STAR-1 and C-SAR (L4P5) data on the results. The C-SAR (L4P5) is imaging the backslopes of the foreslopes imaged by the STAR-1 and vice versa.

Figure 4 shows a graph in which slope and local incidence angle is plotted against radar DN. The slope data were divided into 16 classes and the mean DN value for the STAR-1 and C-SAR (L4P5) was calculated. Only the STAR-1 and C-SAR (L4P5) radar data were used as their look directions approximated the TM sun azimuth. The slope classes and associated radar mean DN values were used in a linear regression of the data. The R^2 values are shown in Figure 4 for all slopes and only those slopes oriented ± 20° to the normal of the radar look direction. It can be seen that the correlation between slope and radar mean DN is generally low for all slopes but much higher (R^2 > .85 and significant at .01) for slopes facing the radar look direction. The regression indicates that, at least for slopes facing the look direction, greater than 85% (R^2) of the radar backscatter can be attributed to the effect of slope.

To further evaluate these effects the cumulative histograms for the STAR-1 and C-SAR (L4P5) data were divided into 16 density classes (DN) above approximately the 70% level. Graphs in which the average slope and aspect were plotted for each DN class were constructed. Multiple regression using DN as the dependent variable and slope and aspect as the independent variables for the data above the 75% level of the cumulative histogram indicates that slope and aspect account for approximately 80% of the radar backscatter on average (R^2 =.74 and .87 for the C-SAR (L4P5) and STAR-1 data respectively, significant at .01)

3.2 Comparison of the TM Classifications, Radar and DTM Data

Based on the results discussed above a map was produced showing areas where the effect of topography on the TM classifications would be expected to be the greatest. This map, shown as Figure 5, was produced by dividing the aspect data to include only slopes that are ± 20° to the normal of the STAR-1 look direction which, as mentioned previously, is in the same direction as the TM sun azimuth. This "topographic influence mask" was logically combined with each class of the TM-based classification to determine the intersection, thus establishing areas of the TM classifications where the effect of topography would be more prevalent.

Mean DN values for the three radar images and TM band 4 are plotted for each TM class in Figure 6. In general TM bands 4 and 5 were found to show more topographic influence than bands 1, 2 and 3, and the resulting graphs for bands 4 and 5 were similar. Therefore, only the results for TM band 4 are displayed in Figure 6. Figure 6 shows that the relative shapes of the curves for the STAR-1 data and the TM data are the same suggesting that influence of topography is similar for both types of imagery. Comparing the curves calculated using all slopes and slopes oriented ± 20° to the normal of the STAR-1 look direction shows that, although the overall shapes of the curves are similar, the absolute mean values increase by about 20 DN for the radar and approximately 5 DN for the TM data for ± 20° slopes. This again demonstrates the effect of slope aspect in relation to the radar look direction/sun azimuth direction. This effect is much stronger for the radar, as would be expected, as radar is an active unidirectional sensor.

Table 1 is a tabulation of the average STAR-1 DN values for all slopes and slopes oriented ± 20° to the normal of the look direction, as well as average slope and aspect for each TM derived class. In addition, the number of pixels in each TM class and the number of pixels which intersect the "topographic influence mask" expressed in absolute and percent values are tabulated. The topographic influence on each TM class can be determined from Table 1 and is the strongest for the classes that have the highest mean radar value and, concomitantly the greatest number of pixels intersecting the "topographic influence map". Note that the TM classes with the highest radar mean values also are associated with higher average slopes and slope aspects closer to the normal of the radar look direction. The classes that display the highest intersection with the "topographic influence map" are classes 1 (or/ms) and 5 (bw/po/ms/or). The areas of intersection between each TM derived class and the "topographic mask" were eliminated, leaving residual areas (pixels) not affected by topography. These areas are presently being evaluated with respect to anomalous areas defined by soil and till geochemical surveys.

4.0 CONCLUSIONS

1. The contribution of this particular radar data set to the definition of geobotanical anomalies (vegetation) appears to be minimal, as much of the variation in backscatter is attributed to topographic (slope and aspect) effects. This may be partially related to the time of data collection (late Oct./87 and Jan./87 for the C-SAR and STAR-1 data, respectively) as the seasons of data collection for vegetation discrimination were not optimal.

2. Radar in conjunction with the DTM data provided a useful tool with which to evaluate the topographic influence on the TM classifications and to produce maps where the effects of topography would be expected to be the most pronounced. The classification with topographic effects removed would be more useful in a mineral exploration programme as vegetational anomalies based on spectral response could be evaluated more easily with respect to underlying geochemistry and potential mineralization defined by ground based till sampling and known mineral occurrences. Work on this aspect of the analysis is on-going. Furthermore, the radar can also be used, in a relative sense, to map areas of relative slopes and aspects, thus helping to define topographic patterns that may be useful for evaluating regional geobotanical anomalies.

REFERENCES

Harris, J. (1984) Lineament Mapping of Central Nova Scotia

Using Landsat-MSS and SEASAT-SAR Imagery, Proceedings of the 9th Canadian Symposium on Remote Sensing, St. John's, Newfoundland, Aug., p 359 - 373.

Hornsby, J. K., B. Bruce and V. Malhorta (1987) Influence of Terrain in Geobotanical Modelling, Proceedings of the 11th Canadian Symposium on Remote Sensing, Waterloo, Ont., June, p 277-286

Hornsby, J. K., B. Bruce, J. Harris and A. Rencz (1988) Implementation of Background and Target Geobotanical Techniques in Mineral Exploration, Proceedings of the 6th Thematic Conference for Remote Sensing in Exploration Geology, Houston, Texas, May, p 511-521

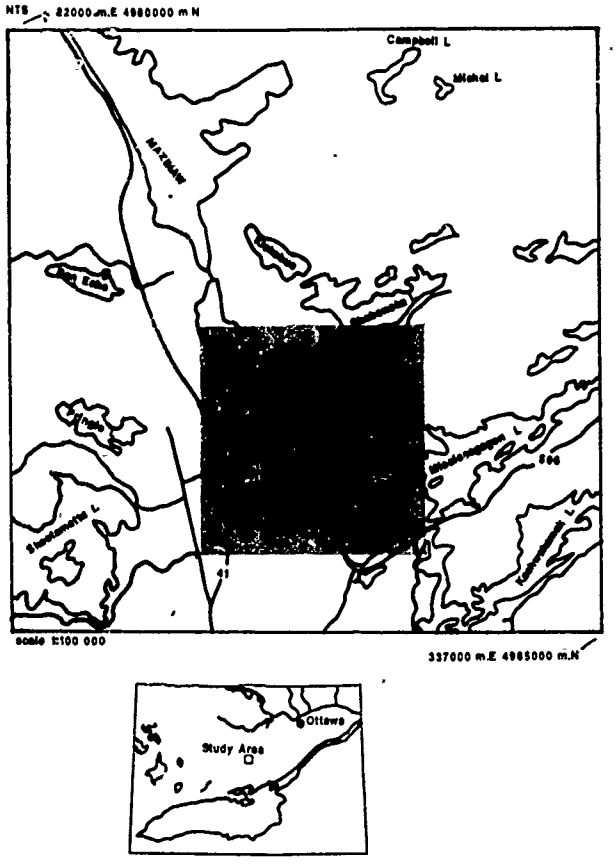


Figure 1 - Mazinaw Lake Study Area

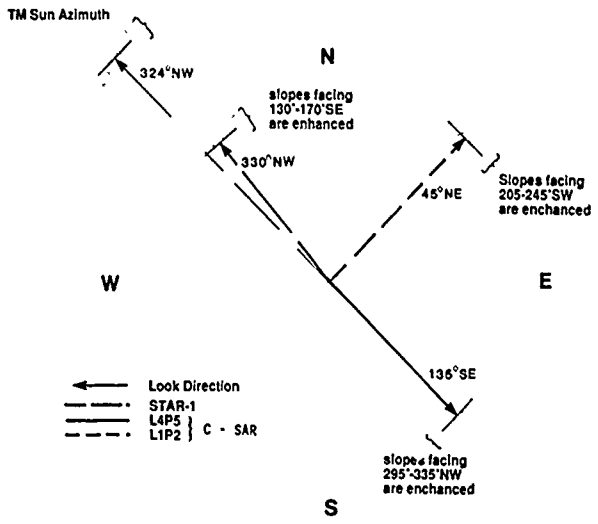


Figure 2 - Radar Look Directions, TM Sun Azimuth and Associated Slope Enhancement Zones

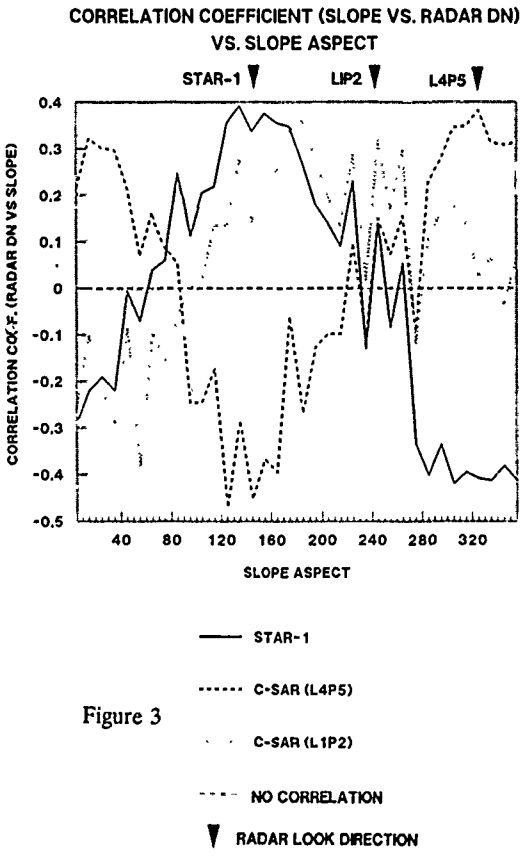


Figure 3

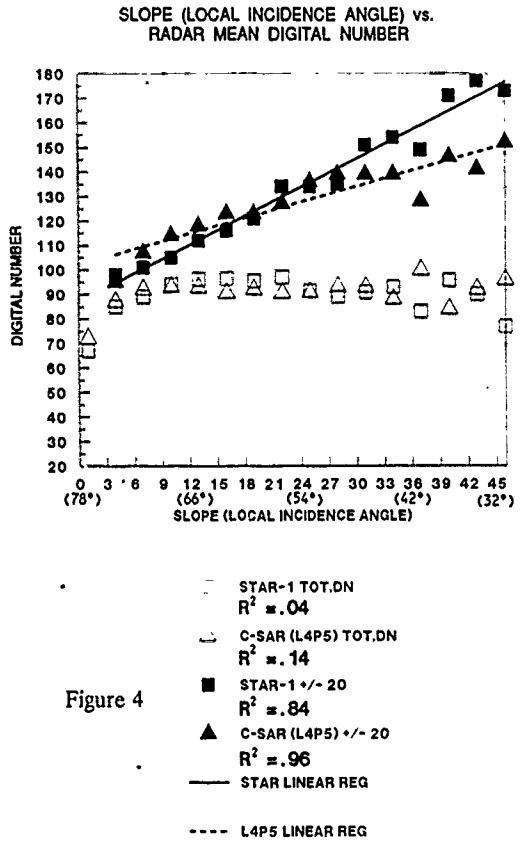


Figure 4

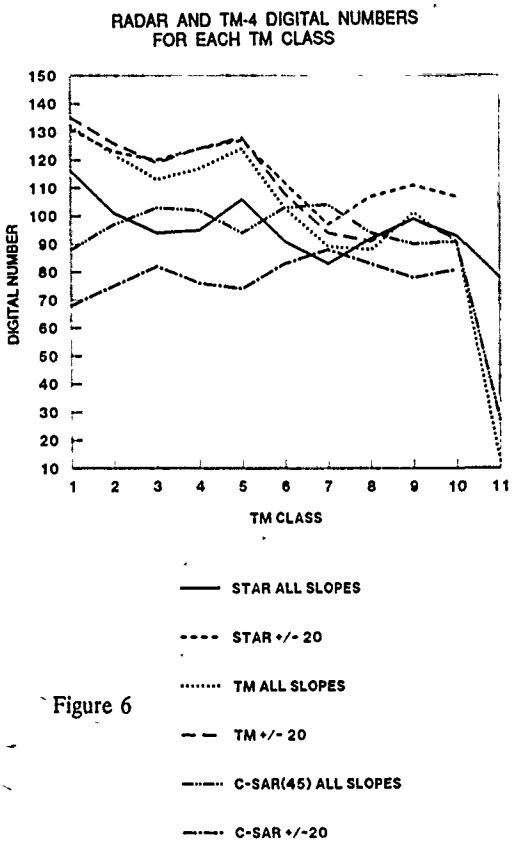


Figure 6

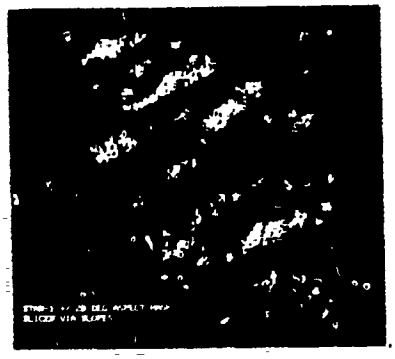


Figure 5

Topographic Mask Obtained From the STAR-1 data

Table 1 - Average Slope, Aspect, Radar Digital Numbers (DN) for TM Classes

CLASS+		STAR AVERAGE DN TOTAL ±20°		AVERAGE SLOPE° TOTAL ±20°		ASPECT° TOTAL ±20°		DIFF	TOTAL # PIXELS	INTERSECTION OF TM PIXELS WITH TOPOGRAPHIC MASK (# OF PIXELS)
OR/MS	1	116	131	13.7	17.8	160	150	10	6298	1589 (25%)
OR/MS Be	2	101	123	12.4	13.7	172	150	22	11676	1820 (15%)
BW	3	94	120	12	12.8	196	152	46	5317	578 (10.8)
BW/PO	4	95	124	11.9	12.2	190	152	40	6097	605 (9.9)
BW/PO MS/OR	5	106	127	13	14.8	164	150	14	7911	1425 (18.0)
BW/PO PW	6	91	112	10.6	10.6	192	150	42	18347	2212 (12.0)
PW/Be	7	83	97	8.7	8.1	206	146	56	7427	676 (9.1)
SWAMP	8	9 2	107	6.4	5.2	178	142	28	2425	310 (12.7)
OUTCROP	9	99	111	11.9	12.4	184	148	34	2805	382 (13.6)
DIST.	10	93	107	8.8	8.5	170	148	20	3672	304 (8.2)
WATER	11	78	-	-	-	-	-	-	-	--

+ TM Geo-botanical Classification (Hornsby et. al. 1987)

OR - Red Oak
MS - Sugar Maple
BW - White Birch
DIST - Disturbed
Be - Beech
PO - Trembling Aspen
PW - White Pine
Ce - White Cedar

DATA ACCESS AND ECONOMIC INTELLIGENCE FROM SATELLITE REMOTE SENSING FOR WORLDWIDE RESOURCE EXPLORATION, DEVELOPMENT AND ENVIRONMENTAL MANAGEMENT

Frederick B. Henderson III, Ph.D.
President
THE GEOSAT COMMITTEE, INC.
601 Elm Street, Room 438C
Norman, Oklahoma 73019
(405)325-3329

ABSTRACT

Worldwide satellite remote sensing has demonstrated operational exploration and engineering cost savings and reduced exploration risks for global resources through improved geological mapping integrated with other geophysical techniques. Experimental and operational remote sensing satellite systems under development by the United States, France, Japan, India, Canada, ESA, Russia, China and others, will considerably increase our ability to explore for, develop, and manage energy and mineral resources throughout the world. Satellite remote sensing data's use will significantly expand by using improved geographic information systems (GIS) to integrate with other geophysical, geochemical, and geologic data. Timely, non-discriminatory access to satellite data is vital to continued international development of satellite remote sensing technology. Also important for global resource information and economic intelligence will be the growth of a global data acquisition and distribution network capable of handling the expanding volumes of data that will be produced by the large number of remote sensing satellite systems in the 1990's. Growth in the international development of satellite remote sensing for global resources will depend on international observance of the "open skies" and non-discriminatory access to data principle. It will also be dependent on international agreements for cooperative remote sensing R&D with full reciprocity of experimental data and on data derived from operational systems.

KEY WORDS

data access, economic intelligence, environment, "open skies", resources

PRESENT STATUS OF INTERNATIONAL CIVIL REMOTE SENSING

Present Systems

Operational civilian land and ocean remote sensing programs from 1972 to 1989 have consisted mainly of the U.S. NOAA GOES/TIROS, NASA/NOAA Landsat, and the French SPOT systems. These satellite systems were supplemented by the U.S. Seasat and U.S. Navy Geosat ocean satellites and more recently by Japanese MOS-1 and Indian IRS-1 satellites. In the Soviet Union, remote sensing satellite programs have been developed, but data from these systems has not been available up until 1989.

Access and Distribution of Data

Data from these satellite systems have been accessible through various government research programs and agencies and commercial suppliers. Access to these data must first come through acquisition methods which include direct downlink to ground stations, inter-satellite links between observing and telecommunication satellites and ground stations (i.e., Landsat-TDRSS), and tape recorder acquisition and ground station data dumps. These means to acquire global data coverage all include some degree of control over who has access to the data stream and how it is distributed. Early distribution of NASA and NOAA data was by government agencies under the principles of "open skies" and "non-discriminatory access to data" at the "cost of reproduction." With the advent of privatization and commercialization of Landsat and SPOT in 1984 and 1986, access to these data was transferred to government supported, commercial companies (EOSAT and SPOT IMAGE) for non-research data dissemination for "commercial uses." While Landsat and SPOT data prices rose considerably, both companies operate under the principle of timely, non-discriminatory access to data.

Users and Applications

During the past 17 years, the use of satellite remote sensing data has grown throughout many government and private sector user communities. These users and the applications of these data have recently been summarized in depth in two of the NOAA subcontractor studies on the future of the Landsat Program in the U.S. (KRS, 1988 and TASC, 1988), released in 1989. Remote sensing is used operationally by industry and agencies in international exploration and development of energy and mineral resources as well as for international agriculture, timber, and fishing. Other emerging uses and markets for satellite remote sensing include shipping, engineering, real estate, land use, cartography, and environmental study and management. It is important to note government usage of the data generally exceeds 65% worldwide.

Ground Segment Technologies

In addition to the developing user community, the use of satellite remote sensing has spawned a wide

array of ancillary remote sensing technologies and markets. These include ground receiving, data processing, data mangement, interactive work stations, data integration, geographic information systems, and value-added services and other technologies for information extraction from "raw" data provided by remote sensing satellites.

Market Development

In the 1970's, NASA worked with the user community to develop Landsat applications and transfer the use of this technology to both government and private sector users. With the advent of privatization in 1979 and the transfer of the Landsat program from NASA to NOAA in 1982, cooperative applied research, applications demonstrations and other technology transfer efforts of the U.S. government diminished markedly. Since 1972, similar efforts have been established by various government and user organizations around the world such as CSIRO (Australia), CCRS (Canada), Institute Frances du Petrol and CNES (France), RESTEC (Japan), the Geosat Committee, and many university institutes.

Much applied research and development has been done "in house" by large user companies such as some members of the Geosat Committee. In 1987, the Geosat Committee and the NOAA/University of Oklahoma Cooperative Institute for Applied Remote Sensing (CIARS) conducted a workshop to review 15 years the development of geologic remote sensing methods (Williams and Henderson, 1987) by some members of the Geosat Committee. At present, there appears to be a general lack of aggressive market development by the data suppliers and more significantly by the value-added industry. More international industry--government cooperative basic and applied research and applications development will be needed during the 1990's. This is particularly true in consideration that the present satellite remote sensing data is of relatively low spatial and spectral sensitivity when compared to data to be produced during the 1990's.

SATELLITE REMOTE SENSING DEVELOPMENT, 1990-2000

International development of improved operational and research experimental systems during the 1990's will have broad economic implications for the future (Henderson, 1989). An abundance of new satellite data will be produced requiring new and improved means of data distribution. This will include new data of greater spatial and spectral sensitivity and discrimination. Specific details of the remote sensing capabilities of these new systems are listed in the 1987 NASA/NOAA report to the U.S. Congress, entitled "Space-Based Remote Sensing of the Earth" (NASA/NOAA, 1987).

Operational Versus Research Systems

The general characteristics of the civil operational and research satellite systems for the 1990's are summarized in Table 1. Operational and experimental systems include U.S. TIROS/NOAA satellites and Landsat 6, and perhaps 7, France's SPOTS 2-5, ESA's ERS 1-2, Radarsat, Japan's MOS and JERS, India's IRS 2-3, and the Soviet Meteor and Resource satellites. Research systems include NASA's Earth Observation System (E.O.S.), Japan's ADEOS and ITIR, and other polar orbiting platforms proposed by NOAA, Canada, ESA, and Japan. Operational satellites imply more or less continuous global coverage over the life of the satellite. Experimental systems carry new, first time experimental sensors in an operational mode. Research systems provide more limited coverage by suites of new experimental sensors.

The Coming Data Crunch

These new systems will produce tremendous amounts of "raw" satellite data. The satellites listed in Table 1 will provide over 430 different VNIR, SWIR, and THIR spectral bands of data in various band widths and spatial resolutions over various swath widths. They will include at least 10 different radar bands and new ocean sensor data from radar altimeters, wind scatterometers, etc. The proposed NASA polar orbiting platform itself will provide 1 terrabit/day (NASA/NOAA, 1988). The data crunch expected from the remote sensing satellite of the 1990's will require the development of advanced data/information management systems on the ground. Compounding the satellite data problem will be the increased availability of airborne imaging remote sensing systems. International policies will be needed for access to and distribution of data from research, experimental, and operational satellite systems for both research and commercial purposes.

Economic Potential and Global Resource Information

Present satellite data is of relatively low spatial and spectral sensitivity with little inherent economic value as data. More sensitive, narrow band airborne systems such as NASA/JPL's AIS (128#s) and AVIRIS (224#s) have demonstrated greater discrimination of specific minerals which may have significance as indicators of deposits of economic importance. Similar high spectral sensitivities in vegetation remote sensing may have significance in observing previously unknown leaking hydrocarbon deposits. Increased spectral sensitivity will also be of great public and economic value in studying and monitoring global change and environment management. In any case, the increased sensitivity of multi-narrow band satellite systems such as NASA's HIRIS (spaceborne equivalent of AVIRIS with 192 #s) instrument for the E.O.S. platform will provide data of significant economic value for global resource exploration and environmental management.

The increased economic potential of more sensitive satellite data will enhance the development of civil remote sensing systems for both government and commercial use. Exploration for global energy and mineral deposits is greatly impacted by timeliness of data in obtaining competitive land or offshore positions (i.e., leases, concessions, etc.). If a government or industry exploration entity has discriminatory or proprietary access to such data, it could have a significant economic competitive advantage. Similar economic advantages could accrue to entities with discriminatory access to improved data for international agriculture, timber, or fishing economic intelligence (i.e., commodity markets). Discriminatory access to more sensitive data by a satellite country could provide it with significant international economic advantage with, or in, less developed countries. Of key concern here, is whether the expanding satellite capabilities in some countries could lead to more nationalistic use of new sensor systems to the detriment of international satellite remote sensing. It is no surprise that the satellite producing countries all have need of international access to global resources and markets and see their satellites as providing an assured supply of global economic resource intelligence. Most satellite producing countries such as France and Japan, have designed their satellite systems to meet both industry and government needs for international economic intelligence.

International Observance of Open Skies and Non-Discriminatory Access to Data

Until recently, some countries have not observed the principle of timely, non-discriminatory access to data. For example, international access to Landsat data received in India was virtually unavailable. Moreover, the Soviet Union did not provide any access to their satellite data. Conditions are changing. With the 1988 launch of its IRS satellite, India has provided international access to their data. In 1989, Soyuz-karta and Glavcosmos have made moves to open some limited commercial access to Soviet photographic and electro-optical satellite data.

In 1987, Japan's government research and space organizations, STA and NASDA, launched the experimental MOS marine observation program. In 1979, the Ministry of International Trade and Industry (MITI) determined that an advanced geological satellite might provide data useful for Japan's critical industrial needs for global energy and mineral resources and set about to develop the JERS-1 satellite. The possibility that such an industry oriented satellite might not provide data on a timely non-discriminatory basis led to several years of study and debate in Japan between various agencies and interests leading to the joint NASDA/MITI development of the JERS-1 satellite. The JERS-1 satellite will be a major advancement for geological exploration as it provides most of the 1976 recommendations of the Geosat Committee for geologic sensors (see Table I), combining high spatial resolution, stereo, spectral and radar sensors. JERS-1 will provide greater SWIR spectral sensitivity with three bands in the 2.1-2.4 micron region compared with one band for Landsat 6. This will greatly increase its clay and soil mapping capabilities among others.

Public questions by the Geosat Committee and others led to a statement of NASDA's intent to provide timely, non-discriminatory access to the JERS-1 data (1988 letter communication to the Geosat Committee). While NASDA will provide such access to JERS data to the U.S. government, ESA, and probably some other governments as research data (i.e., under the NASDA/NASA MOU of 1988), it is not yet clear what the Japanese policy for global distribution of JERS-1 data will be for "commercial uses." Agreements to exchange satellite data on a non-discriminatory basis between governments as research data and potential commercial access thereto is a problem not only for the Japanese JERS, but also for ESA's ERS and Canada's Radarsat. International arrangements to exchange new research data and access to such data for "commercial uses" remain as major national and international problems to be worked out during the 1990's.

What must be decided is whether governments will continue to observe the principles of open skies and timely, non-discriminatory access to all data, or whether governments will allow open or covert discrimination in data access for national advantages. The principle of non-discriminatory access encourages the development of international complimentary, compatible, and perhaps joint systems, commercial or otherwise. This will keep remote sensing satellite costs down by minimizing unnecessary duplication and encouraging international competition "on the ground," rather than in space. Restrictive policies might spur some proprietary commercial development, but it would encourage satellite development only for those companies and governments who can afford to build expensive competitive systems in space, effectively leaving out the rest of the world remote sensing community.

SOME INTERNATIONAL POLICY ISSUES FOR THE 1990's

Preparing for Remote Sensing in the 1990s

During the 1990s, several major trends will develop in international civil remote sensing:

1. Operational, experimental, and research systems will provide vast amounts of new data of increased sensitivity and enhanced economic value for global resource intelligence;

2. Remote sensing data will be used more as data integrated with other multiple data bases in a growing information industry; and

3. Commercialization will continue to expand as the economic sensitivities of remote sensing systems increase. Commercialization will first dominate the ground segment, while continued government support for long-term research and for space segment satellites will be needed throughout much of the 1990's until a viable commercial remote sensing market is developed.

Countries developing civil satellite remote sensing systems during the 1990s will do so for three primary reasons (Keystone, 1989):

1. To provide timely global resource economic intelligence;

2. To compete internationally in remote sensing and related information systems technologies; and

3. To study global change and monitor the balance between man's continued need for global resource development and man's need to better manage his impact on the environment.

Policy Issues for the International Remote Sensing Community

In the United States, the U.S. government is undergoing a major review of its politically troubled Landsat program. This review is being conducted by the newly formed National Space Council which will consider all aspects of continuing a civil land and ocean remote sensing program in the United States including the roles of continued government support, commercialization, and international cooperation. Japan continues to evaluate its various government and industry programs. France is reviewing the commercial viability of the SPOT program. ESA is already caught up trying to determine the roles of government requirements, research needs, and the role of commercial sales for its ERS satellite, yet to be launched, while similar deliberations over Radarsat are being conducted in Canada. Similar considerations are being given to other satellite programs in the Soviet Union and India and in other countries planning satellite programs. In all these satellite remote sensing deliberations, some basic policy issues will have to be decided during the 1990's. They include:

1. The relation of long-term, high risk government research and development with short-term, lower risk commercialization;

2. The relation between inter-governmental cooperative remote sensing research and the access to operational systems developed therefrom;

3. National and international access to government and inter-government remote sensing research data for non-research, "commercial uses,"; and most importantly

4. International observance of open skies and non-discriminatory access to data.

BIBLIOGRAPHY

1. NASA/NOAA, "Space Based Remote Sensing of the Earth: A Report to the Congress, Prepared by National Oceanic and Atmospheric Administration/U.S. Department of Commerce and National Aeronautics and Space Administration (Washington, D.C.: Superintendent of Documents, U.S. Government Printing Office), 1988.

2. Keystone, Keystone Landsat Policy Dialogue, Final Report, The Keystone Center, P.O. Box 606, Keystone, CO, 80435, 1989.

3. KRS, "Study for An Advanced Civil Earth Remote Sensing System," 3 Vol.; Kodak Remote Sensing, Divsion of Eastman Technology, 1200 Caraway Court, Landover, MD, 20785, 1988.

4. TASC, "A Study of An Advanced Civil Earth Remote Sensing System," 2 Vol.; TASC, The Analytical Sciences Corporation, 55 Walkers Brook Drive, Reading, MA, 01867, 1988.

5. Henderson, F.B. III, "Remote Sensing Strategies for Global Resource Exploration and Environmental Management", in press, 1989.

6. Williams, T.H. Lee and Henderson, F.B. III, Geologic Remote Sensing Methods: An Assessment of Status, Needs, and Research Direction. The Geosat Committee, Inc., 601 Elm Street, Room 438c, Norman, OK, 73019, 1987.

TABLE I

CIVIL/LAND OCEAN REMOTE SENSING SATELLITES, 1989-1999

Geosat Committee Recommendations 1976	TIROS-NOAA U.S.	Landsat 6* U.S.	EOS 1,2 NASA & NOAA U.S.	GEOSAT U.S.	TOPEX Poseidon U.S. & France	SPOT 2-5 France	ERS-1,2 ESA	E-POP ESA	RADARSAT Canada	MOS 1b Japan	JERS Japan	ADEOS Japan	IRS 2 India	Meteor-Priroda 2 & 3 U.S.S.R
VNIR		5#				4#	7#				4#	12#	8#	
SWIR	5-6#	2#	310#			1#		40#	5#	8#	4#	1#		3#
Thermal IR		1#										3#		4#
RADAR			C,X,L#				L#	X,L#	C or L#		L#			X#
10-30m IFOV	1.1km	15m, 30m, 120m	30m, 0.5-1km			10-20m	25m	10m, 20m, 50m, .25-1km	15-30m 1km	50m, 0.9-2.7km	17m	16m 0.7km	36-73m	30m, 1-2km
STEREO		down-track				cross-track					down-track			?
Other		SeaWifs	altimeter scatter-ometer	altim-eter	altim-eter 3# micro-wave			altim-eter scatter-ometer		2 micro-wave		add'l non-Japan-ese #s TBD		

* Landsat 7 under study

AIRBORNE AND SATELLITE VERIFICATION OF ARMS CONTROL AGREEMENTS: PAST, PRESENT, AND FUTURE

T. W. Peter Brogden
Dept. of Electrical Engineering,
Ryerson Polytechnical Institute,
350 Victoria St.,
Toronto, Ontario,
M5B 2K3

A. Walter Dorn,
Dept. of Chemistry,
Lash Miller Laboratories,
80 St. George St.,
University of Toronto,
Toronto, Ontario,
M5S 1A1

Douglas Scott,
The Markland Group,
201-93 Bold Street,
Hamilton, Ontario,
L8P 1T8

ABSTRACT

Within the context of a general model of arms control treaties, we discuss the significance of remote sensing by airborne and satellite methods. This general model has the features of a closed loop control system, with remote sensing contributing through data collection to the process of verification.. In common with the behavior of all closed loop control systems, any increase in the capability of verification must be matched by improved mechanisms of promoting compliance when events of clear violation are detected, so that a stable result is achieved.

We classify arms control treaties as bi-lateral (exclusively between superpowers) or multi-lateral, with the latter group being subdivided into global or regional treaties. The forms of effective compliance mechanisms differ between the two classifications, with multi-lateral treaties generally requiring international agencies performing both the verification and compliance promotion functions.

Earlier proposals for an International Satellite Monitoring Agency (ISMA) have not found the necessary support, in that direct linkage with a particular arms control regime was not included. However, with the advent of RADARSAT, the prospects of an effectively verifiable arms control or de-militarization treaty for the Arctic as a region will be much greater. We conclude with some comments on the technical and institutional developments which such a treaty would require.

Keywords: Verification, Compliance, Arctic.

MODEL OF ARMS CONTROL TREATIES

Arms control treaties are frequently written in such arcane legal jargon as to be bordering on incomprehensibility to all but international lawyers and career diplomats. However, to engineers and scientists, a control treaty need be no more complicated than a straightforward closed loop control system with feedback, as shown in Fig. 1.

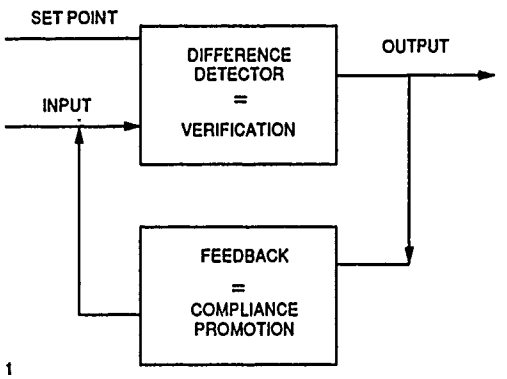


Fig. 1

A difference detector describes the function of "VERIFICATION", and the feedback part of the loop will promote "COMPLIANCE" with the set point. "NON-COMPLIANCE" describes the state when a clear difference exists between the observed input and the set point, and, to be effective, the feedback mechanism must lead to elimination of the difference.

VERIFICATION

Data collection is the primary activity of the verification process, and this is the part where remote sensing makes its largest contribution. Set points can obviously correspond to numbers of missiles, warheads, or aircraft, etc, or magnitude of equivalent tonnage of TNT for the Threshold Test Ban Treaty. In a listing of about 15 treaties negotiated since 1925, nearly all have provisions for national technical means of verification, which include aerial observation in the 1961 Antarctic Treaty, and seismic sensing in the 1974 Threshold Test Ban Treaty. The use of satellite surveillance in detecting the Krasnoyarsk radar antenna is well documented, but the subsequent negotiations to counter the perceived non-compliance with the ABM Treaty are illustrative of options available in such bi-lateral treaties. The Krasnoyarsk incident also emphasized the need for on-site monitoring and inspection to supplement remote sensing, and these techniques are built into most of the multi-lateral treaties. In general, the techniques available for verification, that is, data collection and evaluation, would appear to exceed the power of the compliance mechanisms available under most existing treaties.

COMPLIANCE PROMOTION

Just as a chain is no stronger than its weakest link, so a control loop with feedback must be designed for its total response. If verification is improved, so must compliance promotion as the response to a clear violation. Two of the present authors (Scott and Dorn, 1989) have recently examined compliance procedures for both bi-lateral and multi-lateral treaties. As an example, we consider the negotiations between the superpowers over what should happen vis à vis the Krasnoyarsk antenna, seen by the US as a clear violation of the bi-lateral ABM Treaty. The treaty had established a Standing Consultative Commission, at which the matter was discussed. Various diplomatic moves have been made, including a threat to withdraw from the ABM Treaty, to persuade the USSR to correct the non-compliance situation. However, these moves are strictly between the parties involved, and do not directly involve other members of the international community. We call this process "self-help for the powerful", and it can be sufficiently flexible to allow a stable outcome to be reached.

For multi-lateral treaties, the principle of self-help is not effective, and the current negotiations for the Chemical Weapons Convention (CWC) include setting up a UN agency called the Organization for the Prohibition of Chemical Weapons (OPCW). The experience of the International Atomic Energy Agency in administering the 1968 Non-Proliferation Treaty (NPF) is a good guide as to how the OPCW should operate. All the functions of verification and compliance promotion will be the responsibility of the OPCW, with specific agreements as to how on-site inspections would be carried out. The OPCW will report to the UN Secretary General, with full publication of its activities. The evaluation and response procedures still need clearer definition, and we must make the point that the use of a majority vote for decisions on whether violations have occurred and what should be done about them will make a much stronger compliance promotion process, than one based on consensus. By becoming a state party to the treaty, each ratifying state goes on record as agreeing to the majority vote rule, with no single state having a veto.

It was early in 1988 that the superpowers changed from insisting on the self-help style of compliance promotion, to supporting a Consultative Committee. This change in attitude in the regime of the CWC probably reflects the rapidly spreading ability to manufacture chemical weapons by states with hitherto limited capabilities of any chemical production, in the war zones of Africa and the Middle East. Even the big boys were running scared, and have realized the vital importance of a broadly-based and strong CWC. It must be remembered that the 1925 Geneva Convention on Chemical and Bacteriological Weapons has no provision for verification or compliance promotion, and does not ban production or stockpiling, but only the use of such weapons.

The well-publicized satellite pictures of the putative Libyan chemical weapons factory only serve to emphasize the limitations of the 1925 Geneva Convention. It is not illegal to build such a factory, nor to have it in full production. Any unilateral action by another state to eliminate the factory would be in direct contravention of international law. The alleged death toll due to the use of chemical weapons by Iraq in the war against Iran did not produce vigorous diplomatic response. However much one may deplore this situation, it is in accordance with the existing treaty. The only legal actions that could be taken were against violators of export controls, and we note that the rigorous enforcement of such controls will always play a strong part in compliance promotion.

Full scope verification within the CWC will have to use a wide array of techniques for data collection, as remote sensing by satellite cannot get close enough. Other technical means for sensing the pre-cursor and waste products as well as the agents themselves could, in some cases, be called remote sensing, but it appears that to be most effective, the participation of citizen reporters must be included.

INTERNATIONAL SATELLITE MONITORING AGENCY (ISMA)

The development, and hopefully early conclusion, of the CWC at the Conference on Disarmament in Geneva will give new credibility to full scope verification systems in arms control treaties. The original ISMA proposal by France in 1978 was for a very broadly based monitoring agency, but was felt by the superpowers to infringe on their territory of satellite use for spying. Although many nations supported the proposal, there was always sufficient opposition from the superpowers to prevent it coming into being.

In another context, it was only one of up to 12 initiatives to create some sort of International Verification Agency. These initiatives came about because the so-called middle powers were dissatisfied with much of the information being disseminated about the arms build-up by the superpowers. These events are chronicled in the monograph by one of the present authors (Dorn, 1987). Judged against the progress in the CWC negotiations in getting agreement on a full-scope verification system or agency, these proposals did not have sufficiently clear boundaries to their regimes, nor the urgent threat of proliferation. Though touted generally as confidence-building measures, the detractors were always able to undermine confidence that their perceived security would be improved by the creation of any such agency.

AN ARCTIC INTERNATIONAL MONITORING AGENCY

We propose that an Arctic International Monitoring Agency is appropriate now, and would combine and carry forward a number of actions already evident in the international arms control scene.

Firstly, the superpowers have independently moved to support the creation of a second full-scope verification agency reporting to the UN, with the OPCW following on from the IAEA. On-site inspections and more openness will be the order of the day under the CWC. The USSR has expressed a willingness to discuss proposals on arctic arms controls (Gorbachev, 1987), and these should be followed up vigorously by the other northern nations.

Secondly, the operation of RADARSAT will provide vastly extended technical capabilities in surveillance over the Arctic as a region.

Thirdly, real economic pressures are constraining the maintenance of the armaments of all nations, especially the superpowers.

Fourthly, recent commercial and military accidents have heightened public awareness of the fragility of the arctic environment, and the need for more knowledge of what is going on.

This proposal is outlined in greater detail by one of the authors (Dorn, 1989). The recommended order of development would be

different again from the IAEA and OPCW, reflecting the fact that this agency would have regional monitoring responsibilities broader in scope than the fissionable and chemical materials of the other two. The strength of the agency must be carefully considered, in light of the arguments on the decision-making processes discussed earlier. We have to recognize that institutional developments have to match technological capabilities. In this case, we believe that, following the initial commitment by the circumpolar nations to the principle of an Arctic International Monitoring Agency, its growth can both be relatively modest and cost-effective, when balanced with reduced arms expenditures.

BIBLIOGRAPHY

Dorn, A.W., "Peace-keeping Satellites: The Case for International Surveillance and Verification". Peace Research Review 10 (5 & 6) Dundas, Ont. Peace Research Institute, 1987.

Dorn A.W., "An International Monitoring Agency for the Arctic", Information North, Vol.15, No.1, Jan. 1989.

Gorbachev, M, *The Speech in Murmansk* (1 Oct. 1987) Moscow: Novosti Press Agency, 1987.

Scott, D. and Dorn A.W., "Making Arms Control Treaties Stronger", International Perspectives, Vol. XVIII, No. 1, Jan./Feb. 1989.

STUDY OF AN ADVANCED CIVIL EARTH REMOTE SENSING SYSTEM
RESULTS OF AN ASSESSMENT OF TECHNOLOGY AND MARKETS THROUGH 2000

David Okerson
Science Applications International Corporation
P.O. Box 1303
McLean, Virginia, U.S.A., 22102
Tel: (703) 734-5542
Fax: (703) 734-2899

NOAA recently commissioned three parallel studies of the economic and technical options for an "Advanced Civil Earth Remote Sensing System". This paper describes the results of one of these studies; the results are consistent with the findings of all of the studies. Current and planned remote sensing systems were considered, together with the known and anticipated market for remotely sensed information and the projected improvement in technical capabilities and system costs. The "market" was considered to include all users of the information, including U.S. government purchases (in particular, the Department of Defense). The "system" was understood to include all parts of the launch, space, and ground segments, as well as the marketing/business arrangements.

In essence, the studies conclude that the purely commercial prospects for a business which would service all the costs of a remote sensing satellite program are untenable. This conclusion is essentially independent of any satellite sensor characteristics or any credible changes in market or technology and associated costs within the period in question. In business terms, the "net present value" of an investment in the system is a large negative number.

Therefore, if a U.S. civil space remote sensing program is to continue, government funding in some form will be necessary. A range of indirect funding techniques are possible to make the funding politically acceptable, but the economics are quite simple. Among the more interesting alternatives to reduce the direct cost, while maintaining a useful program are:

- The U.S. program might usefully pursue sensors which complement, rather than directly mimic the characteristics of alternative international sources of remote sensing data;
- The U.S. program might dilute costs by arranging a shared program with foreign participation.
- The characteristics of the system should be chosen to address as wide a range of information users as possible while limiting the cost. This tends to favor instruments capable of serving a wide range of information needs rather than a "single-purpose" or a very high priced instrument.
- An investment in ground processing systems, including data processing, algorithm development, archiving, and information distribution systems (as opposed to the space segment) appears to provide the greatest return on investment.

The study specifically identified microwave remote sensing as capable of addressing a very wide range of anticipated market needs. U.S. participation in a program such as RADARSAT would be directly consistent with our recommendations.

Abstract for:

"PLANNING AN ADVANCED COMMERCIAL REMOTE SENSING SYSTEM: EVALUATION OF SPACEBORNE, ENVIRONMENTAL REMOTE SENSING SENSOR TYPES WHICH ADDRESS THE CIVIL AND MILITARY MARKETS."

Dr. D. Brown, C. A. Hood
Dr. D. Okerson, Dr. M. Schwaller, R. Suresh
Science Applications International Corporation
400 Virginia Ave, SW, Suite 810
Washington, DC 20024
Phone: 202-479-0750
Fax: 202-479-0856

Dr. S.W. McCandless
User Systems Eng., Inc.
4608 Willet Drive
Annandale, VA 22033
Phone: 703-978-8898

Choosing an optimal suite of sensors as a payload for a commercial remote sensing satellite requires an analytic method to help sort through the various sensor combinations. The method should be able to integrate sensor performance parameters with data and information requirements derived from market surveys and projections. This integrating methodology will yield payload selections that are both technology-driven and market-driven. This paper describes such a methodology and presents some results based upon sample market data.

Information on an exhaustive set of over 150 existing and planned spaceborne, environmental remote sensing instruments and sensors was collected. Instruments were grouped into twelve separate categories based on their spectral and spatial resolutions (such as Infrared Sounders, Moderate Resolution Visible and Infrared Radiometers, etc.) As a parallel effort, a list of remote sensing products that are responsive to the information requirements identified in civil and military market projections was generated. For each product, the extreme values for spatial resolution, revisit time, and product turn-around were delineated based upon their expected utility to the civil and military applications. A figure of merit for the feasibility of utilizing each sensor type for the generation of each product was calculated based upon the following metrics: stand-alone capability, coverage parameters, all weather-capability, processing requirements, and accuracy. The results were arranged into an matrix whose dimensions were (p x s), where p = the number of products, and s = the number of sensor types. Summing the columns gave an indication of which sensor types were most responsive to the entire set of products from a purely technical perspective. Sensor types with the highest scores were those that are very versatile (i.e., they could be used to generate many of the desired products.) Such instrument types include Synthetic Aperture Radars (SAR), Visible and Infrared

Radiometers of varying resolutions (Moderate, High, and Ultrahigh), and Low Frequency Microwave Radiometers (LFMR.)

A second analysis was performed which introduced the effect of prioritizing the list of products based upon their relevance to either the civil or military markets. This process produced a column vector (1 x p) which contained the market-derived product weights. Products having a spatial resolution of 30 m or better were ranked high due to a large number of high-resolution information requirements in both the civil and military sectors. When these market weights were combined with the technical feasibility results (i.e., the technical feasibility matrix multiplied by the product weighting vector), a combined market-driven and technology-driven sensor ranking was generated. Our results showed that the sensor rankings exhibited little elasticity. We attribute the similarity in sensor rankings to the fact that the most versatile sensor types also have high enough spatial resolution to meet anticipated market demands.

The results of this investigation served two purposes. First, they identified candidate sensor types for an optimal commercial remote sensing payload; second, they discouraged the inclusion of some sensor types which rank high for a specific product but rank low in meeting overall civil or military market needs (i.e., Altimeter, Scatterometer.)

GRS-S PARTICIPATION IN INTERNATIONAL SPACE YEAR (ISY)

C. Specter

Florida International University
Miami, Florida 33199 USA

ABSTRACT

The major theme of ISY is Earth observation from space for human benefits. Through the use of remote sensing technology, we seek a better understanding of the factors contributing to global environmental changes, as well as those contributing at national and regional levels to degradation of the environment. As arguably the most relevant IEEE Society concerned with observations of environmental conditions of the Earth, it is appropriate that our professional interests and capabilities be brought to bear on this significant undertaking. Members of the Geoscience and Remote Sensing Society (GRS-S) are able to become involved in ISY through a wide range of events and activities.

1. INTRODUCTION

Members of GRS-S were issued a challenge at IGARSS '88. It came in the form of a keynote address delivered by Sir Hermann Bondi. He reminded us that the real value of remote sensing technology is to be found in its contributions to the prosperity and well being of mankind. The dramatic advances of technology during this century have been accomplished through the diligence of the scientific and technical communities in overcoming technical problems associated with experimental systems. Now the same energy, good will, and drive must be expended toward overcoming legal, institutional, and political obstacles if remote sensing technology is to realize its full potential (Bondi, 1988).

Technology and diplomacy must go hand in hand if the output of remote sensing systems is to be integrated into national and international decision making processes concerning the development and conservation of the planet's natural resources. Likewise, political and technical bodies must cooperate to assess and monitor global environmental changes if contributing forces are to be understood and controlled. For their part, scientists and engineers need to demonstrate that information derived from remotely sensed data makes a valuable contribution in these areas, and that its value outweighs its costs. International Space Year (ISY) provides GRS-S members with the opportunity to respond to these critical global demands.

2. BACKGROUND ON ISY

Plans for ISY began in 1985 when the U.S. Congress passed a resolution recommending its establishment, and suggesting ISY take place during 1992, commemorating 500 years since Columbus' discovery of the New World. The plan was endorsed by President Reagan and NASA in 1986. Preliminary proposals were prepared in February 1987 by the Space Science Board of the U.S. National Academy of Sciences and the Intercosmos Council of the Soviet Academy of Sciences. These early plans included space exploration missions as well as Earth observation missions. Even in these early plans, the U.S. proposals focused on science, applications and education, including increased public awareness (McLucas and Meyerson, 1988). Over time, Earth observation utilizing satellite remote sensing has emerged as the major theme of ISY. The ultimate purpose of the endeavor is to inspire and channel international response to the growing environmental challenges we face.

Countries participating in ISY include the U.S., Canada, numerous European states, Japan, the USSR, the People's Republic of China, Brazil, and Pakistan. The list is expected to grow.
Professional scientific and engineering societies have endorsed the concept. The Committee on Space Research (COSPAR), the International Council of Scientific Unions (ICSU), the International Astronautical Federation (IAF), and our own Institute of Electrical and Electronics Engineers (IEEE) exemplify the wide range of support. ICSU views the objectives of ISY as complementary to those of its global change program, the International Geosphere-Biosphere Program. The ISY timeframe has been extended to 1992 - 1994.

Given the increased number and sophistication of Earth observation missions scheduled over the next decade, it is expected that a flood of data will be available for analysis and interpretation, coming from such systems as ERS-1 (ESA), JERS-1 and ADEOS (Japan), TOPEX/Poseidon (U.S.-France), UARS (U.S.), IRS-1 (India), the U.S.S.R. RESOURCE system, Brazil's MECB system, advanced weather satellite systems, SPOT, and possibly Landsat. Beginning in the mid 1990s, a series of polar

platforms will provide additional sources of information: e.g., RADARSAT, scheduled to be launched in 1994, and the Earth Observing System, an international cooperative effort supported by NASA, NOAA, and the NSF, in partnership with ESA, and hopefully Australia, Canada, and Japan. The first of four platforms is scheduled for launch 1995.

At the 1988 ISY Mission to Planet Earth Conference, six working groups prepared recommendations regarding ways in which these systems could be best utilized to address growing concerns regarding the causes and impacts of global change (Meyerson 1988). One series of recommendations related to the need for coordination among national space agencies. Information must be collected concerning the capabilities of current and future systems, gaps in coverage need to be addressed, data must be standardized, and worldwide accessibility to the output must be achieved so that the global user community can integrate the information with other sources. As a mechanism to implement these objectives, the Space Agency Senior Officials working group established the Space Agency Forum on International Space Year (SAFISY) as their coordinating body.

Two working groups focused their attention on the Global Information System Test (GIST), a system for coordination of collection, calibration, and distribution of Earth observation data. GIST would provide demonstration projects of an "end to end" use of data, a necessary first step in establishing global standards for global change studies. Suggested pilots projects included one focusing on the greenhouse effect and one on tropical deforestation. Likewise, the Environmental Monitoring Objectives Working Group recommended pilot projects in these same areas, along with a global data set production project.

Concomitant with these global interests, ISY is to incorporate initiatives of particular relevance to developing countries. Given the pressing social and economic conditions facing many of these countries today, it is necessary that ISY projects provide information that is directly relevant to development planning within the context of sustainable environments. The Social and Economic Development Objectives working group stressed the importance of transforming data into useful information so that end users- project officers and managers will be persuaded to use it. Educational programs for political decision makers will be required if they are to be made aware of the potential for applying remote sensing technolcgy to development programs, and specialized technical training for indigenous personnel will be necessary if they are to use the technology once the idea is accepted.

3. AREAS OF POTENTIAL INVOLVEMENT FOR GRS-S

Members of GRS-S are able to become involved in ISY in numerous ways. Individuals may be employed by government agencies or private sector organizations that have significant roles to play in relation to one or more of the Earth observing missions scheduled over the next decade. Three additional areas of potential interest to GRS members are discussed below: i) participation in GRS-S sponsored educational activities; ii) involvement in the IEEE ISY International Telecommunications Network; and iii) participation in one or more projects which demonstrate the value of applying remote sensing technology to global, national, and regional problems.

3.1 Educational Activities

The IEEE, represented by then President-Elect, Russell Drew, participated in the Pacific ISY Conference in 1987. Seven technical panels outlined possible activities that could take place under ISY auspices (McCord, 1987). Two of these panel reports are of particularly interest to the remote sensing community. The Earth observation panel began to formulate plans that were expanded upon in 1988 at the Mission to Planet Earth Conference described above. The Professional Associations Working Group for ISY outlined eight basic categories of activities suggested for adoption by professional associations, stressing their role as disseminators of information. These categories are: topical/specialty conferences and sessions; publications/newsletters; lectures/displays; a global communications network; advertising, design competitions; educational programming; and briefings to governments. The point was made that this list was not intended to be all-inclusive, suggesting that other opportunities for involvement might arise over time.

GRS-S has already responded to many of these suggestions. IGARSS '89 sessions are covering topics such as global change and other ISY related subjects. ISY topics will have increasing prominence at future IGARSS. The theme of IGARSS '92 will be "Mission to Planet Earth: International Space Year." GRS publications already reflect an ISY focus. For example, the March 1989 IEEE Transactions on Geoscience and Remote Sensing was a special issue on the Earth Observing System (EOS) and a special issue dedicated to ISY is being considered during 1992. Likewise, the IEEE Geoscience and Remote Sensing Newsletter has run several feature articles on EOS, and one article geared to arouse membership interest and involvement in ISY activities. In relation to the idea of a lecture series, the GRS-S Adminsitrative Committee is considering the inclusion of a series on ISY for IEEE section meetings. Also, the GRS-S Administrative Committee has approached the SSP/AESS, offering to assist in planning and implementing the ISY International Telecommunications Network. This effort is described in more detail below.

3.2 The ISY International Telecommunications Network

A Global ISY Communications Network was proposed at the 1987 ISY Pacific Conference. By October 1987 it had been determined that IEEE would be involved in this activity, and the lead society would be the Space Systems Panel of the Aerospace Electronics and Systems Society (SSP/AESS). Participation and input from other IEEE societies is expected. The ISY Telecommunications Network is responding to a House of Representatives document that calls for a system "utilizing a network of existing national and international communications satellites for a series of

coordinated seminars on major worldwide scientific/applications concerns or challenges. Each program might be sponsored by a country, group of countries, or international organizations (e.g., WMO), but would be simultaneously available to the whole net."

The original intent of the ISY International Telecommunications Network was to provide videoconferencing as well as transmission to and from space sources, and digital, voice, or facsimile transmission. The current plan is to use the ISY International Telecommunications Network as a "test bed" to explore successful information transfer, to provide a system for exchange of information among scientists, and a mechanism to transfer data directly from satellites to the scientific community, through the use of the Integrated Services Data Network (ISDN). NASA's Advanced Communications Technology Satellite, ACTS, scheduled for launch in 1992, could be integrated into the network as well. Since GRS-S represents 2,000 engineers and scientists concerned with Earth observation from space, it is hoped that we will play a role in system planning and design.

3.3 ISY Demonstration Projects

According to the 1988 ISY Mission to Planet Earth Conference report, ISY efforts to strengthen coordination and standardization are to emphasize global issues, but also regional initiatives of particular relevance to developing nations. The best method for achieving these objectives are sharply-focused pilot projects that set standards for worldwide application. Earth observation initiatives are to extend beyond research to include immediate and direct applications for social and economic development (Meyerson, 1988). One such pilot project proposed at the Mission to Planet Earth Conference was a deforestation project, to be chosen from several experiments currently planned. The satellites mentioned in relation to the study are SPOT, LANDSAT, GOES, and METEOSAT, as well as those that will be flying in 1992.

Two members of GRS-S have proposed a tropical deforestation project which differs from the above in three major areas. First, the project would utilize C-Band satellite radar systems; second, the project would be long-term, spanning a ten-year period; and third, it would go beyond the demonstration stage to provide an operational global monitoring system. The project, "Observation of Tropical Forests by C-Band Radars: Proposed Ten-Year Project Embracing the International Space Year is discussed in more detail below (Specter and Raney, 1989).

Tropical deforestation is one of the most critical problems confronting humankind. The majority of these forests are found in developing and newly industrialized countries. Tropical forests comprise only 7-8% of the Earth's land surface, but almost half of the global wood stock, and at least two fifths of the Earth's species. Tropical forests play a major role in the precarious carbon dioxide balance of the Earth's atmosphere. Destruction of this global resource is taking place at an alarming rate. A review of forecasts concerning the rate of deforestation are indicative of the lack of knowledge concerning the problem. Some predict that these forests will disappear, except for small areas preserved as parks, by the year 2050; others maintain they will be gone by 2020.

The effects of tropical deforestation may include detrimental impacts on vital global processes, as well as on economic and social development programs focused at the national level. In addition, tropical deforestation can be linked to increased human suffering, in terms of loss of livelihood, shelter, and lives.

Observation of tropical deforestation based on optical systems is limited by optical visibility. Tropical forests are frequently occluded by cloud, fog, and mist. Radar frequencies reliably penetrate these media so that tropical forests are observable from radar satellite perspective. While there is a forestry radar data base (including C-Band), it largely does not include tropical data. More importantly, there exists no coordinated plan to utilize existing and prospective radar remote sensing resources to respond to this problem. And finally, regarding the institutional infrastructure necessary to implement the program, there is a lack of linkages between the leading radar sponsoring nations and the developing countries whose futures may depend on addressing the deforestation issue.

Critical milestones scheduled to occur in the next few years make it appropriate for a major program to be considered. The most important experimental radar mission is SIR-C (1992). Future operational C-Band radar satellite systems include ERS-1 (launch in 1990) and RADARSAT (launch in 1994). Ground stations to receive and process satellite SAR data are being established within the developing regions under consideration. A project should be started now to prepare techniques and institutions to use data from these systems to address the tropical deforestation problem.

3.31 Methodology

The initial phase of basic science and program planning, expected to take two to three years, would lead up to a demonstration period (two to three years), followed by a phase geared to operationalize an ongoing global monitoring system (3 or more years). Steps necessary to implement the first phase would require obtaining in situ and airborne radar scattering data, securing commitment for experimental data from SIR-C for sites (or region) of the project, and obtaining agreement for a minimal set of ongoing data from planned radar satellite systems. It would be necessary for the project team to design a basis for ongoing data supply from the satellite radar systems, probably through the purchase of the data, contingent on its acceptance and proven utility of a sample data set. Also, a local data collection and data reduction campaign must be designed and implemented. This would entail the establishment of a local capability to support and then continue measurement, leading to local expertise. The radar data would be integrated into a transferable data base, since it is hoped that in the operational phase the information could be used locally, as well as internationally.

On the institutional side, potential participants must be identified and contacted, including both technical and financial sponsors, as well as host countries. Institutional linkages must be created, including those with key international committees and agencies concerned with tropical deforestation and global change. An education program should be established in the host nations if the technology is to be integrated into national planning and decision making processes. Finally, a methodology for measuring the effectiveness of the program, including its technology transfer aspects, would be valuable for ongoing assessment and generalizability.

3.32 Results

Results of this program would be tied to the accomplishment of both scientific and institutional objectives. Scientific results should include the creation of a quantitative methodology for tropical forest observation by C-Band radars, including specific forest radar signatures, extension to the tropics of existing (temperate) forest radar data bases, and global qualification of radar sensible parameters in forestry models. Institutional outcomes should include the establishment of expertise in radar tropical forest observation resident in the tropical participating nations, and the establishment of linkages between the participating nations and the radar satellite sponsoring nations for ongoing tropical forest assessment. Successful implementation would require cooperative working relationships among technical and scientific communities, host country government(s), regional organizations, and international bodies. In this way, the results of quantitative observation of the key environmental parameters can be fed into local, national, and international monitoring and planning processes.

4. MEETING THE CHALLENGE

For long range interdisciplinary and perhaps abstract problems, such as those which lie at the heart of ISY initiatives, it is tempting for us to think that someone else will take on the responsibility of dealing with these issues. However, for ISY to be successful, the search for solutions must involve expertise and leadership from the scientific and technical communities, as well as from those in the political sphere.

GRS-S represents a key group of individuals whose skills and talents can be applied to these endeavors. ISY is an open agenda item at GRS Administrative Committee (Ad Com) meetings. A small committee has been formed to study additional areas of involvement. The GRS-S ISY committee looks forward to receiving suggestions from the membership at this early phase of idea generation and discussion. Beyond areas of GRS-S sponsorship, the Society will provide an information network and resource base for projects that extend beyond the scope of any one professional group. Our energy, good will, and drive can contribute to the potential of this important undertaking.

5. REFERENCES

Bondi, Sir H., "Remote sensing: Moving toward the 21st century", Proceedings of IGARSS '88 Symposium, Edinburgh, pp. 5-7, 13-16 September 1988.

McCord, T.B., ed., Report on the Pacific ISY Conference: A Planning Meeting for the International Space Year, Kona, Hawaii, 19-21 August 1987.

McLucas, J. & Meyerson, H., "Planning for International Space Year", Space Policy, pp. 108-111, May 1988.

Meyerson, H., ed., Report of the ISY Mission to Planet Earth Conference: A Planning Meeting for the International Space Year, Durham, New Hampshire, 29 April - 1 May 1988.

Specter, C. & Raney, R.K., "Observation of tropical forests by satellite C-Band radars: Proposed Ten-year project embracing the International Space Year (ISY)", abstract submitted for presentation at the ASPRS/ACSM 1989 Fall Convention, Cleveland, 17-21 September 1989.

TECHNOLOGY TRANSFER - A CANADIAN EXPERIENCE

Helmut Epp

Canada Centre for Remote Sensing
c/o NWT Centre for Remote Sensing
Department of Renewable Resources
P.O. Box 1320, Yellowknife, NWT
Canada, X1A 2L9
Tel. (403) 920-3329
Fax: (403) 873-0221

Jeff Whiting

Saskatchewan Research Council
15 Innovation Blvd.
Saskatoon, Sask., Canada,S7N 2X8
Tel. (306) 933-5423
Fax: (306) 933-7446

ABSTRACT

From September 1984 to July 1987, the Saskatchewan Research Council in collaboration with the Canada Centre for Remote Sensing, introduced remote sensing technology for resource management to potential users within the province of Saskatchewan. The program was established through a Memorandum of Understanding (MOU) between the Government of Saskatchewan and the Canada Centre for Remote Sensing (CCRS), Energy, Mines and Resources.

Under the Saskatchewan Technology Enhancement Program (STEP) ten demonstration projects were undertaken in consultation with CCRS and Saskatchewan provincial departments and agencies. Each was designed to demonstrate the capabilities of remote sensing techniques to address a specific problem of interest to that organization. The objective was to determine whether remotely sensed data could be used operationally to provide more comprehensive, timely and/or less costly information than did traditional methods.

The average mapping accuracy for the ten projects was 88% with benefit/cost ranging from 5 to 400. An average of 160 learning hours were required by each agency to adequately operate the equipment. Projects took approximately 40 hours to complete with an equivalent amount of time for ground verification of the results. The average cost was $5,500 per project. Over the three year period, some 8,000 hours of computer time were consumed.

The remote sensing technology transfer provided benefits to the agencies taking part in advancing their data collection and analysis procedures and making many aware of the additional data bases available to them to be exploited.

INTRODUCTION

For over 17 years, the American Landsat series of satellites have orbited the earth, providing data on the world's natural resources . The NOAA series of weather satellites complements Landsat by providing a synoptic view of a larger part of the earth's surface. In 1986 the French SPOT satellite was launched. The European Space Agency and Japan have launched additional satellites and Canada will soon have its own satellite, RADARSAT, adding the new capability of microwave observation.

The remote sensing application Technology Enhancement Program (TEP) was first formulated by the Canada Centre for Remote Sensing (CCRS) in 1979 and became a full program in 1980. The program was to enhance the economic and social benefits of natural resources through the use of remotely sensed data. Through the TEP the provinces and territories have access to the most advanced and cost-effective resource management methods. Industry also benefits from having a larger and more varied source of revenue through services and products. Universities benefit by being able to expand their remote sensing expertise to further their functions in education and research (Heyland 1986).

The TEP functions through Memoranda of Understanding (MOU) which are signed between CCRS and an agency of the host province or territory with each partner committing certain elements to the program on a best effort basis. A number of projects are jointly conducted using proven methods to solve resource management problems over a two to three year period.

In January 1985 a Memorandum of Understanding between CCRS and the Saskatchewan Research Council (SRC) was signed to formally start the

Saskatchewan Technology Enhancement Program (STEP). The program began with seven demonstration projects with additional projects added on later.

THE NEED

Saskatchewan's diverse resources (natural,cultivated and non-renewable) are the core of the Saskatchewan economy. Of the $5.4 billion worth of products and services exported from the province in 1984, nearly 80% were directly related to three main resource industries: $2.3 billion in minerals, $1.75 billion in agricultural goods, and $231 million in forest products. The long-term stability of Saskatchewan's economy and the continued well-being of Saskatchewan's citizen's are highly dependent upon the effective management of these resources to ensure abundant and continuing supply, efficient use, and profitable and competitive production for world markets.

OBJECTIVES

The objectives of the Saskatchewan Technology Enhancement Program were to:

1. train Saskatchewan resource managers and government personnel in the uses of remotely sensed data.

2. provide a facility, equipment and a remote sensing specialist to support the program.

3. provide a forum for exchange of information.

THE PROGRAM

As its contribution to the STEP, CCRS stationed a remote sensing specialist at the SRC for the duration of the program. In addition CCRS loaned a digital image analysis system to SRC on which data analysis for the projects were performed.

The STEP program had ten demonstration projects which included migratory waterfowl, deer, moose and caribou habitat; bulk area estimation of crops; field boundary assessment; forest fuel mapping; burn and cut-over inventory; salinity in irrigated areas; and depression storage as a measure of flood potential.

The projects and principal investigators were geographically distributed over the province. Digital and visual analysis techniques were used in executing the projects.

Adopting a new technology is not without cost. Some are monetary, but many are related to human behaviour and attitude. Because the electronic equipment for manipulating satellite data was available at low cost to Saskatchewan during STEP, it was a factor not considered. Rather, challenges faced by the principal investigators was considered.

All the principal investigators received basic training in remote sensing and digital image analysis. To some it was a new technology and rather intimidating. In some cases there was a long learning curve, particularly in the use of the image analysis system. Some investigators became very proficient in the use of the computer, while others experienced considerable difficulty. Still others, due to personal schedules or schedules imposed upon them, had long lapses between computer sessions and therefore were forced to relearn some procedures several times over.

Obtaining cloud free imagery within a suitable "biological window" was a challenge. Once obtained it was difficult in some of the studies to discriminate, decisively, between different types of ground cover. For example, trees versus shrub communities, spring crops versus similar dark areas on the landscape and jack pine versus open bogs.

In some instances demands were put on the data which could not be met. For example, land cover is often of such diversity that absolute discrimination of type cannot be obtained from satellite data or perhaps even with aerial photography. A mixed forest is exactly that and complete separation if deciduous and conifers may not be possible even though it may be desirable. As a result, expectations and criteria sometimes had to be lowered. Problems of this nature, however, were resolved by the principal investigators by accepting a lower number of classes and accuracy in a classification. Calculations of the area of cover type, once acceptable criteria were established, were always made with a speed and accuracy that far surpassed any other method of area determination.

The study of saline areas was particularly challenging. Where accuracies of classification for the crop and summer fallow study were 75% and better, accuracies of classification of saline soils fell between 20% to 30%. Non saline soils were misclassified, not through mishandling of the data, but because the differences in spectral reflectance of the soil types were too subtle to separate. This indicates that a different type of data (such as that from radar) may be required, that the

digital data must be visually interpreted, or that a more appropriate biological window needs to be identified.

The study of artificially drained land had a low success evaluation because the criteria for some aspects put demands on the data which could not be met. Small ditches and other man-made structures below the resolution of the satellite sensor were to be mapped. These features, however, were to small to be resolved by the data. Satellite imagery should not be abandoned as a suitable data source, because they, in combination with microwave data may well become useful for other hydrological purposes in the future.

BENEFITS FROM STEP

There are enormous benefits to be gained through the use of space-borne data for monitoring Saskatchewan's natural resources (Table 1) and the joint projects undertaken within the STEP have illustrated many. Most agencies that took part have been able to advance their data collection and analysis procedures.

TABLE 1
ASSESSMENTS OF BENEFITS FROM DEMONSTRATION PROJECTS
(from Whiting, Epp, Heyland, 1988)

Project	Accuracy Assessment (%)	Benefit Cost Ratio	Time(hours)			Program Cost $
			Training	System	Field	
Field Boundaries	82	n.d.	176	18	-	8,900
Forest Cover						
Burns	93	5:1	255	12	70	10,550
Clearcuts	82	5:1				
Roads	73					
Caribou Habitat	n.d.	n.d.b.	0	20	20	-
Crop Area	108	n.d.	-	-	-	-
Deer Habitat	93	4:1	48	63	60	1,858
Forest Fire Fuel	87	400:1	100	72	20	4,170
Moose Habitat	87	35:1	200	40	20	2,200
Salinity in Irrigated Fields	90	n.d.	200	40	n.d.	-
Drained Land	-	n.d.	162	-	-	

NOTE: n.d. : not done
n.d.b. : not done before by any other method
- : not available

CONCLUSIONS

Space-borne data of the earth's resources can be applied to many problems facing resource managers. They do not, however, hold all the answers. Techniques, such as air photo interpretation, are still the best in many areas (e.g. detailed urban planning). However, when a resource manager must address a problem which involves an extensive area, or perhaps one in a remote and inaccessible region, he must turn to other lower cost data sources. Such situations dictate satellite data.

The window of opportunity opened by the STEP indicated to SRC that there was an obligation to continue providing assistance to users as they moved into operational use of remotely sensed data. Choosing not to continue, or being restricted from improving the service, would have denied the first Saskatchewan users access to a technology available to their counterparts elsewhere in Canada. The potential benefits afforded by remote sensing could have been lost to them and to the economy and the people of the province as a whole. Saskatchewan has therefore implemented a follow-on program to provide a remote sensing service to resource managers.

REFERENCES

Heyland, J.D. 1986. The Canada Centre for Remote Sensing's Technology Enhancement Program. Proc. 10th Canadian Symp. on Remote Sensing, May 5-8, 1986, Edmonton, Alberta, 1011-1015.

Whiting, J., H. Epp, J.D. Heyland. 1988. Summary Report of the Saskatchewan Technology Enhancement Program. SRC Tech. Report No. 213, 16 pg.

IS REMOTE SENSING AN ECONOMIC TOOL IN THIRD WORLD COUNTRIES?

B Turner and N M Walters

Earth, Marine and Atmospheric Science and Technology,
Council for Scientific and Industrial Research, Republic of South Africa

ABSTRACT

Remote Sensing is the only effective means currently available for obtaining the information needed by Third World planners. However, in applying remote sensing, certain constraints specific to the third world environment must be considered. These include a lack of trained manpower and established infrastructure.

The results of a survey are discussed and certain conclusions are reached. Remote sensing is compared with conventional techniques and found to have considerable advantages for the Third World situation, especially in relation to the ratio of cost versus amount of information collected.

Keywords: Natural resources, inventories, scale.

1. INTRODUCTION

The increase in population in the region imposes severe demands upon the use of the natural resources. Thus the need for natural resources inventories in Third World countries is both urgent and essential. Remote Sensing technology is perhaps the only method currently available to satisfy the urgent needs of the decision makers. However, to succeed in the Third World environment modification must be made.

First World remote sensing technology cannot be directly transferred to the Third World situation. Things which the First World takes for granted, such as trained manpower, scientific equipment and even mundane items such as up-to-date maps, very often do not exist in the Third World.

If quantitative remote sensing techniques are to be used successfully in the Third World, adaptations must be made to suit local conditions of land-use, agriculture, population etc. and also cultural, organizational and social factors.

In view of the above, a survey was conducted to determine which remote sensing technology could best be used in southern Africa and what main problem areas should be addressed. The results of this survey are described and the conclusions discussed.

2. METHOD

Collection of information

Information was collected from individuals and groups (committees) mainly associated with local and government authorities in agriculture, water, forestry and urban development from the following countries and areas in the southern African subcontinent: Malawi, South West Africa/Namibia, Lesotho, Swaziland, Botswana, Transkei, Ciskei, Bophuthatswana, Kwa Zulu and Venda.

The marketing information and needs were initially obtained by means of a questionnaire. From the 200 questionnaires distributed by post, 195 replies and requests for further information were received. The replies were analysed and follow-up visits and discussions were and are still being arranged.

3. RESULTS

3.1 In almost every case, keen interest was expressed in remote sensing as a means of solving urgent (indeed crisis) situations concerning the management of natural resources especially with respect to providing the primary needs of food, shelter, water and fuel for an increasing population. The utilization of the resources of the region for the creation of work opportunities and for recreation are important secondary requirements.

3.2 The fragility of the natural balance due to the predominantly semi-arid conditions was emphasised. The preservation and management of wild life and ecosystems and the impact of development upon the environment were topics which were highlighted as problems of extreme importance due to the inadequacy of existing methods.

3.3 The problems are very diverse and require a mult-disciplinary approach.

3.4 Present efforts to find solutions to the problems are not satisfactory.

3.5 Many areas within the region lack up-to-date maps. The simultaneous collection of airborne spectral information and conventional aerial photography affords some states an excellent opportunity of acquiring a digital geographic information system upon which all future planning can be based. An example of this is available.

3.6 Most of the problems are characterised by specific local parameters. Thus local examples are needed to demonstrate the potential of remote sensing technology. These examples must be generated by fundamental research supported by field spectroscopy using locally collected data.

3.7 Available trained manpower is usually very limited and often non-existent.

3.8 Funds available for natural resource management are limited.

4. DISCUSSION: THE TECHNICAL ADVANTAGES OF USING MODERN QUANTITATIVE REMOTE SENSING TECHNIQUES TO MEET THE NEEDS OF THE THIRD WORLD.

4.1 The need for a National Natural Resources Inventory

There can be no doubt that many countries in southern Africa need a national natural resources inventory to facilitate the planning, management and monitoring of the national wealth of the country and to monitor the impact of development upon the environment.

Data captured by an airborne scanner offer a practical method of acquiring the geographic information data base for this purpose.

Many developing countries in southern Africa are at present, in an accelerated development phase. Existing economies are very dependent upon agriculture and forestry. Hence the immediate needs are for an inventory of the soils, soil potential, water, soil moisture, land use and land potential. In the near future, decision makers will require a natural resources inventory for the planning of diversification strategies.

4.2 Ideal Data Format

Quantitative remote sensing technology produces data in a digital format. A digital data base can immediately be used in computer modelling, offering extremely powerful advantages over the old-fashioned analogue methods. Developing countries have a unique opportunity of starting with a digital geographic information data base by employing remote sensing rather than conventional methods.

4.3 Effective utilization of trained manpower

The adoption of remote sensing substantially reduces the need for manpower-intensive field surveys and subsequent data analysis and integration. It is therefore very effective in conserving scarce trained-manpower resources.

4.4 Remote Sensing versus Conventional Methods

Data for natural resources management have, to date, been collected by aerial photography or by taking field samples using surveying methods. These methods are suitable for projects over small regions and the data collected by field samples are

- exact
- point measurements
- project related
- collected over a long period of time
- costly for large regions
- costly for large numbers of variables
- labour intensive.

Since the data are collected by means of point samples, often the sampling density of such measurements must be compromised in order to reduce costs and time of data acquisition.

In countries where extensive conventional information has already been archived, most of these data were collected for specific projects and hence are regional, limited in the number of parameters covered and of a fixed scale. Attempts to integrate these data into a regional information system fail due to the fragmented nature of the data. The result of this failure is usually that the survey must be repeated to satisfy the new working criteria.

Furthermore it is even difficult to use the data collected for one project in applications related to other projects in the same region.

By contrast, remotely sensed data have multi-disciplinary value and may be applied to diverse projects over small or large (even national) areas and at differing scales.

4.5 Comprehensive and extensive content of the data

Data captured remotely by satellite or airborne sensors are

- spectral
- digital
- collected for each and every element on the ground being surveyed
- designed to be integrated with existing conventional data into geographic information systems
- suitable for conversion to varying scales
- suitable for multi-disciplinary applications
- suitable for speedy interpretation with the assistance of computers (personal computers to main frame)
- suitable for quantitative monitoring of change
- suitable for environmental impact studies.

Thus these data are ideal for the establishment of a national geographic information system from which national natural resources inventories can be derived.

4.6 Multi-disciplinary nature of Remote Sensed data

The value of an inventory is dependent upon the number of parameters included. If remote sensing technology is used the cost is relatively independent of the number of parameters measured. By contrast, in the case of conventional methods, an escalation in costs is incurred with each additional parameter.

5. SOME IMPORTANT CHARACTERISTICS REQUIRED FOR A NATURAL RESOURCES INVENTORY

5.1 Survey parameters

The greater the number of parameters inventorized, the more useful is the inventory. As noted above, remote sensing techniques provide a comprehensive list of parameters.

5.2 Variability of the Scale

Furthermore, the inventories noted above can be derived from the same data base at any of the appropriate mapping scales, e.g.

National development planning	1:1 000 000
Regional planning	1:250 000
Project identification	1:50 000
Land use (Agriculture and forestry) .	1:10 000
Irrigation planning	1:5 000
Irrigation scheme design	1:2 500

provided that the data in the data base have the required spatial resolution.

5.3 Segmentation of data base into managerial districts

The data base can be segmented into administrative units, such as, district, cadastral units, geographic units defined by latitude and longitude or natural units, such as, river catchments to suit the requirements of the decision makers.

5.4 The timing of the data collection

If expensive mistakes are to be avoided, the information required for effective planning in developing countries is needed immediately. The speedy generation of a natural resources inventory can only be achieved with the assistance of quantitative remote sensing technology. Only remote sensing methods are capable of providing the amount of information needed in the time available.

6. CONCLUSION

The major conclusion reached is that in spite of existing efforts to solve them, certain well-known and serious problems associated with the management of natural resources such as land-use, over grazing, over population etc. still persist. Thus it would appear that the recently developed techniques could be applied advantageously. However, the Third World environment has specific characteristics which make certain techniques more likely to succeed than others. For instance, the results of the survey indicate that airborne data would be more useful in the region than satellite data. Furthermore the term 'cost effective' must acquire a Third World connotation. It is necessary to apply remote sensing technology in such a manner so as to assist in the effective utilization of the available limited trained manpower and funds.

Secondly, the misuse of remote sensing techniques by, for instance, applying general qualitative methods in inappropriate circumstances, can be very detrimental to the future acceptance of quantitative remote sensing in areas where it can give valid results.

Thirdly, and of considerable importance, quantitative remote sensing projects must be designed to be long-term so that a history of quantitative trends can be built up in order to verify theoretical predictions and to provide irrefutable evidence in support of arguments for improved land-use, farming practices, urban developments etc.

Finally, and in general, it is concluded that the potential for the application of quantitative remote sensing in the Third World is very large indeed provided that projects are carefully planned to produce credible results on which subsequent work can be based. The generation of natural resources inventories by means of remote sensing technology is, perhaps, the only economic method currently available to provide the information required by decision makers.

Some specific examples of remote sensing projects presently being planned and executed as a result of the survey are available.

7. ACKNOWLEDGEMENTS

The authors wish to thank the CSIR and Aircraft Operating Company for their support and encouragement which enabled them to undertake the survey and to publish the results. In particular we wish to thank the individuals and groups who completed the questionnaire and Joan Moller for her efficient handling of the dispatch and receipt of the survey questionnaires.

INFRASTRUCTURE

- A NECESSITY FOR OPERATIONAL USE OF REMOTE SENSING

H. Hamnes, E. Henriksen, N. H. Lunde, S. A. Solbakk

FORUT (Foundation of Applied Research, Tromsø)
Addr. : Forhåpningen 21 B, P.O.Box 2806 Elverhøy, N-9001 Tromsø, Norway
Teleph. : +47 83 80150. Telefax : +47 83 82420

ABSTRACT

Tromsø Satellite Station (TSS) is the national data acquisition facility for remote sensing satellites in Norway. For operational use of the data, it is necessary to provide the required infrastructure. This includes facilities for archiving and retrieval of the data, real time as well as off line processing of standard products, and distribution of the products to the users.

The concept of a national data centre at TSS has therefore been introduced as Tromsø Earth Observation Centre (TEOC). The paper describes the concept and overall model of TEOC. A pilot implementation of a near real time distribution facility and an interactive catalogue and ordering facility is presented, and future strategy is outlined.

Key words : *Infrastructure - Data Centre - Operational Remote Sensing Applications*

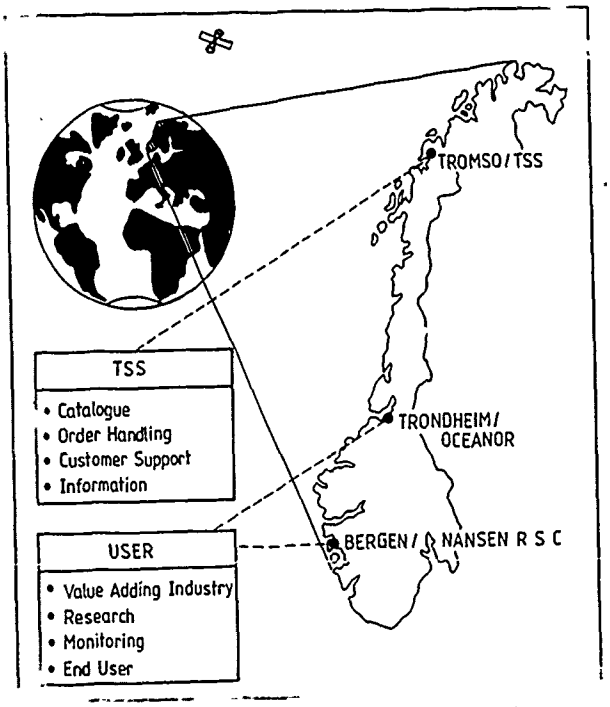


Figure 1: The Concept of Tromsø Earth Observation Centre

TEOC - AN EARTH OBSERVATION DATA CENTRE

Remote sensing is expected to give important contributions to a number of applications (e.g. environmental monitoring, ocean surveillance and ship traffic monitoring, wave forecasting, ice mapping, sea surface temperature studies, etc.). However, due to frequent cloud coverage, the data available from present optical systems is strongly limited. This will be improved by the new generation of radar satellites offering active sensors independent of cloud coverage.

For utilization of the large amount of satellite data in operational systems, the necessary infrastructure must be provided. This includes data centre functions for archiving and retrieval of data, real tir. ., as well as off line processing of standard products to different levels, and distribution of the products to the users.

The primary data acquisition facility for remote sensing satellites in Norway is Tromsø Satellite Station (TSS). TSS receives data mainly from the NOAA satellites and from the MOS satellite. TSS is also the National Point Of Contact (NPOC) for ESA's Earthnet distribution service. The station is presently being upgraded to a national ERS-1 data acquiring and processing centre. The upgrading will be completed at the ERS-1 launch in 1990. Participation in the future EO programs (e.g. Columbus Polar Platform, Radarsat) may further increase the variety of data available from TSS in the future.

Currently, TSS is mainly a data acquisition facility. The data centre functions outlined above are based on manual ad hoc solutions. We are therefore developing a national data centre at TSS, Tromsø Earth Observation Centre (TEOC) [1][2][3], see figure 1. The work is financed mainly by the Norwegian Space Centre (NSC).

The main objectives of an Earth observation centre are to provide Earth observation products to customers, and to do business. The customer must be able to retrieve the desired products in an easy way, and the centre must provide effective processing and distribution facilities. The main function therefore is an effective order handling system, closely integrated with a catalogue and an archive. The customers may be located far away from the acquisition station or data centre. By using computer networks in order handling, catalogue access and product distribution, the effect of distance is minimized. Automation of data centre functions increases the efficiency and reduces the costs.

The first step in the TEOC development is an electronic delivery system combined with a catalogue and order handling system. These systems have to operate in a heterogeneous environment (i.e. different types of computers, terminals, networks, and software), and should be accessible on a 24 hour basis from remote terminals through a wide area network.

PRODUCT DELIVERY VIA DATA NETWORK

During 1989 we are carrying out a pilot experiment with near real time product delivery through a data network. Two national pilot users have been selected. Oceanor in Trondheim represents the value adding industry, while Nansen Remote Sensing Centre in Bergen is a research institute. In the experiment, NOAA AVHRR data is used. We focus on quick-looks, raw data, and sea surface temperature data.

The distance between TSS and most of its users is quite large. Nansen Remote Sensing Centre is located more than 1500 km from TSS. The present method for 'fast delivery' of satellite data products has been by plane and taxi, with delivery times ranging from 4 to 12 hours. Using data network we obtain delivery times down to 2 minutes for 0.5 Mbyte products, and reduced delivery costs. Especially for operational monitoring (sea temperature for fish industry, pollution detection, ship detection, etc.), such a dramatic decrease in delivery time is of significant importance.

TSS will make quick-looks available through the data network immediately after a satellite pass. A user may then fetch the quick-look from TSS (either by electronic mail, or by file transfer). After inspecting the quick-look, the user may decide to order a given product from an area covered by the satellite pass. The order is sent to TSS using electronic mail. Shortly afterwards the user may fetch the generated product from TSS. See also figure 3 The requested product will be at the user site approximately one hour after the data is read out at TSS. This is within the 90-minutes time limit of an ESA fast delivery product.

UNINETT, the Norwegian research data network, is used in the experiment. The network architecture is illustrated in figure 2. UNINETT interconnects the Norwegian universities, research institutions, and colleges. In addition, UNINETT is responsible for gateways to a number of international networks. UNINETT supports electronic mail handling (X.400). In addition, remote login and file transfer is supported for X.25, the DARPA protocols (part of the international internet), and DECnet. The UNINETT backbone consists of 64 Kbps permanent X.25 connections combined with leased 64 Kbps lines. Cisco IP routers are used to handle the DARPA traffic. In the pilot experiment the DARPA protocols are used together with X.400 message handling systems.

To prepare for the large amounts of data from future satellites we will start experiments using 2 Mbps satellite connections later this year. This also includes multicasting of products from TSS to the users.

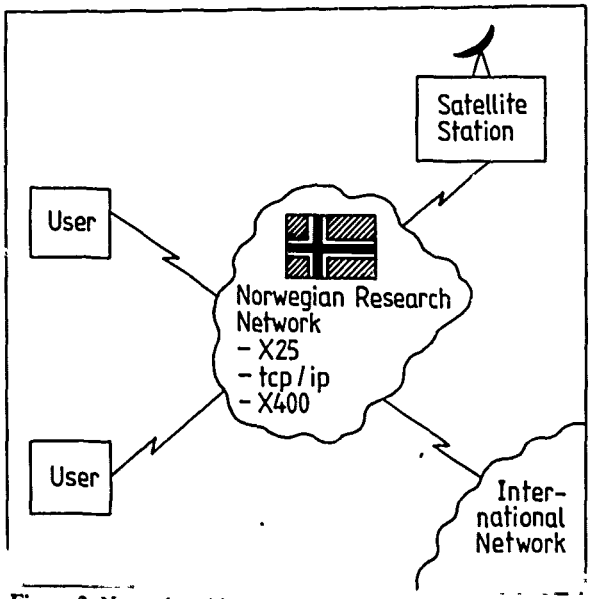


Figure 2: Network architecture for satellite product delivery.

One of the advantages gained by using the existing research data network, is that there is an operation and maintenance organization. In addition, the network is continuously improved using state-of-the-art technology.

We expect that the quality improvement in the distribution service will encourage increased use of satellite data. Existing applications may be improved, and new applications made possible.

The order handling in the pilot experiment is very simple. We concentrate on the last satellite pass available, and the order is created after manual inspection of a quick-look. However, development of an advanced catalogue and order handling system is carried out in parallel with the pilot experiment on data network product delivery.

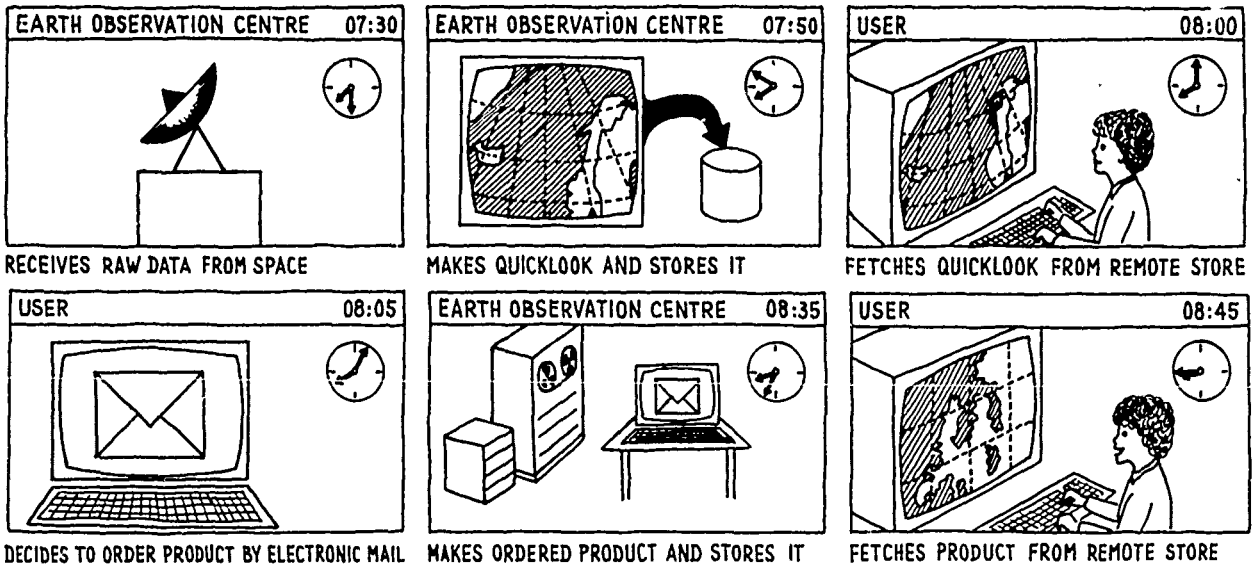


Figure 3: Pilot experiment with near real time product delivery through a data network.

ORDERING FACILITY FOR REMOTE SENSING PRODUCTS

In addition to near real time delivery, some users will need to order products based on archived data. To do this, the user will need access to a catalogue giving information about possible products. A pilot implementation of a catalogue and ordering facility will be developed in 1989.

The catalogue describes products offered by TSS. The archive contains mainly raw data and information needed for further processing of this data. The archive may also include some further processed data. Data is stored on different media (magnetic and optical discs, tape, film, slides, paper), and may be located at different sites.

The user can search the catalogue to identify both existing products, and raw data needed to generate new products. A user will need information like: Satellite, sensor, acquisition date and time, geographical area covered, quality, product type, processing level, etc. The search can be done using any combination of attributes and ranges of attribute values. The result of a search is to be presented as a list of records, each record describing one remote sensing product.

To order, the user must also get information about: Delivery methods, time aspects for processing and delivery, product prices, etc. It should be possible for the user to work on a temporary order, which at any time may be canceled. Before sending the order, the user should get an estimate of the total costs.

In the same way as described above, it should be possible to order products based on future passes.

The ordering system must emphasize user friendly and efficient dialogue, for novice, as well as for expert users. A multi-window system with menus, selection from maps, and presentation of quick-looks, could be one version of this interface, see figure 4. The system must, however, also offer an interface to ordinary terminals.

It should be easy to build applications on top of the ordering system. For instance, the system could provide a procedural interface in addition to the interactive user interface, in order to facilitate automatic use of the system. It should also be possible to integrate the system with other available and future catalogue systems. This would give the user access to several catalogues and archives, but the user would still "see" the ordering facility as one system.

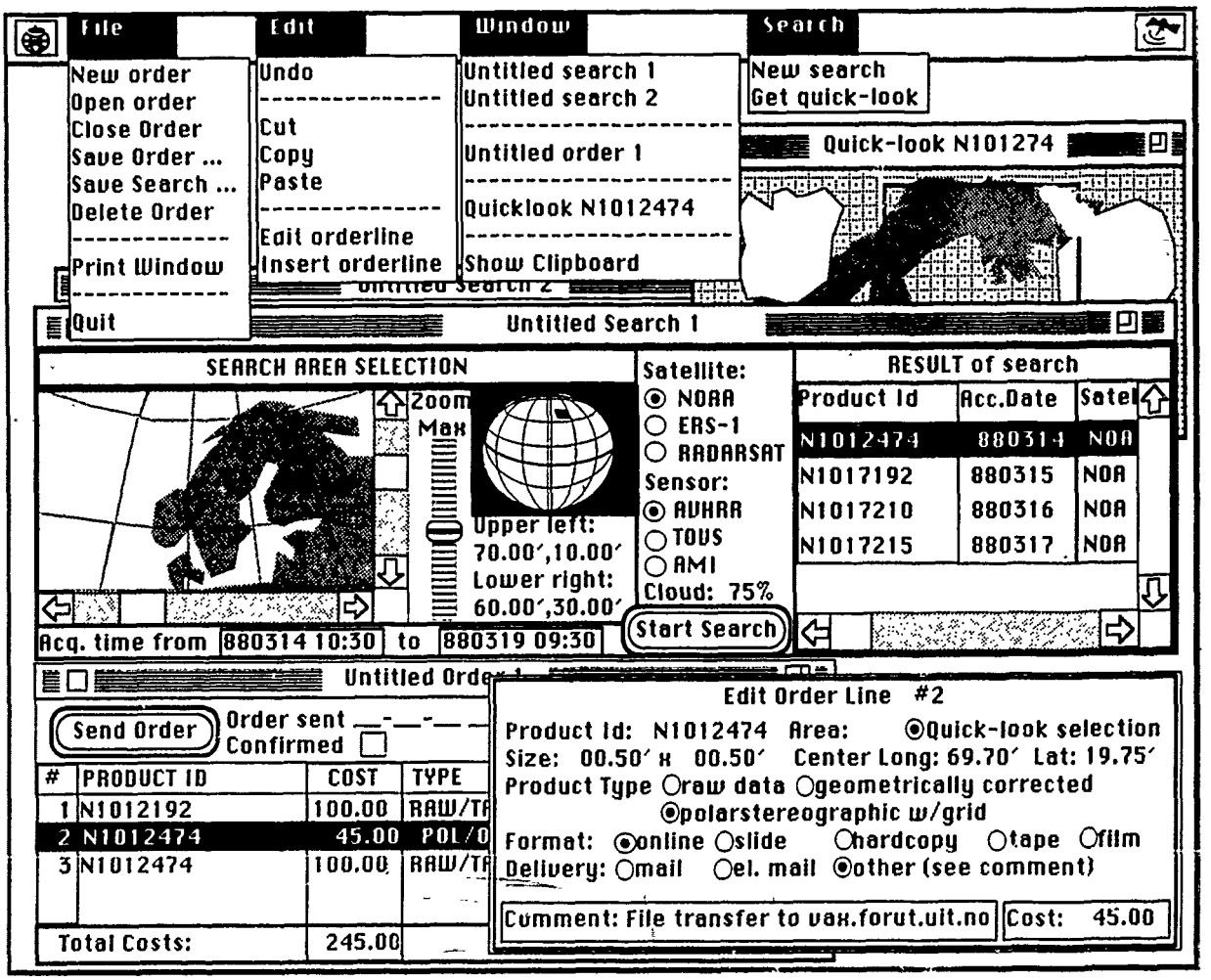


Figure 4: Example of an order handling system providing graphical user interface, with multi-windows and menues.

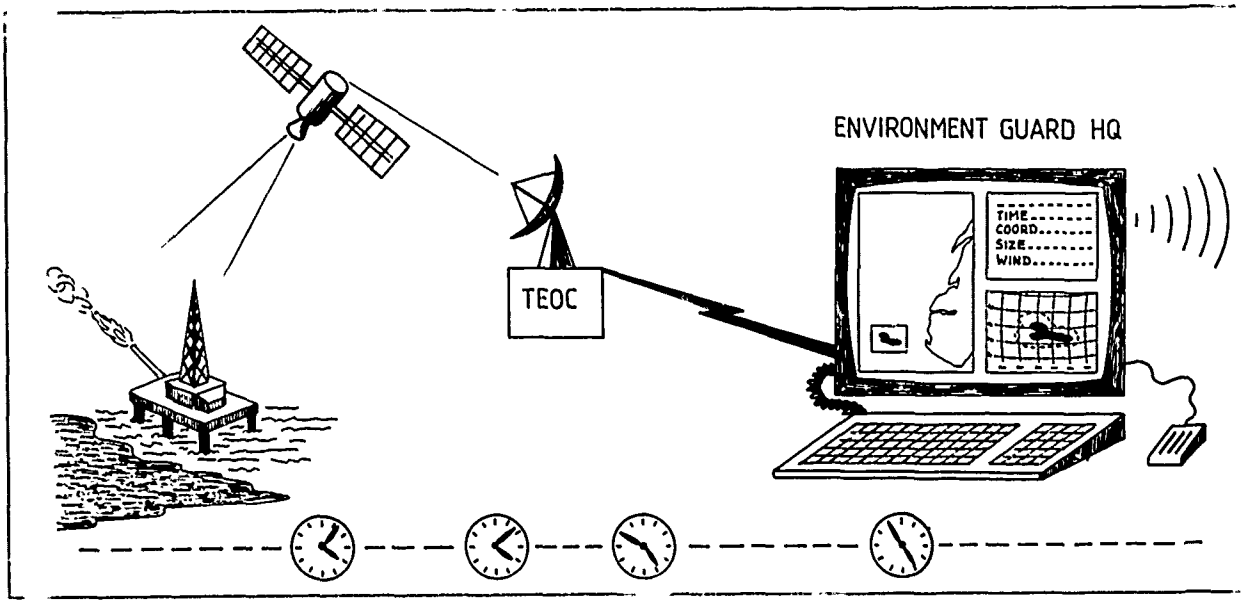


Figure 5: A scenario of a Coast Environment Guard.

AN OPERATIONAL APPLICATION SCENARIO

Environmental problems are among the main challenges to be dealt with in the future, and this is an area where remote sensing products are expected to give important contributions. One example is the problem of ships polluting the oceans, deliberately or not. Valuable and invaluable resources are at stake: Sea life, shores, and income for people depending on the fisheries. The short term and long term effects may not be limited to one country. When the pollution is due to an accident, it will normally be reported from the polluting ship. In other cases, the pollution may evolve undetected for a long time.

To be able to minimize the damages, it is important to detect pollution as soon as possible. Oil spill is an example of such pollution. Effort must be made in cleaning up the oil, estimating the damages, and tracking down the responsible. We present a system to deal with these problems. which may be called the Coast Environment Guard (see figure 5).

The backbone in such a system will be a 24 hour operating monitoring system based on regular satellite data.

FIRST, automatic detecting algorithms discover polluted areas shortly after the satellite has passed. The findings are reported to the Coast Environment Head Quarter, where further actions are co-ordinated.

SECOND, if oil pollution is detected, the clean up operations need *in situ* measurements and additional real time satellite data to describe the extension of the oil spill. Satellite data like ocean current and sea temperature data must be combined with *in situ* measurements in models to estimate how fast the oil will deteriorate, and to give forecasts on the oil spill movements. This in turn makes planning of the clean up operation easier, resulting in less damage done to the nature and better resource utilization.

THIRD, the culprit could be tracked down using ship detection from radar images (historical radar data in the archive) combined with international registers of ship positions. The faster the oil spill is found, the easier it will be to identify the responsible. Oil samples may be used to establish the source of the oil spill.

This application needs both periodic real time data and access to historical data from selected areas. Integration with data from other sources (e.g. *in situ* measurements), and models for combining the data will be needed.

CONCLUSIONS

The future strategy in providing infrastructure for operational use of Earth observation data should be influenced by a few crucial observations:

- Earth observation data will be one of several data sources for application systems [4]. The systems must integrate these data before delivery to the end-user.
- Multi temporal and multi sensor products will require standardized storage and retrieval formats.
- The large amount of data (e.g. from radar satellites) requires centralization of data processing before delivery to users, to reduce communication cost and storage requirements for the user.
- Earth observation infrastructure (distribution networks, order handling, user support, etc.) should be developed in a general way to handle different types of users and products. Providing infrastructure is a continuous process where new technology must be integrated as required by the users and the applications.
- Effort must be made to identify and develop future markets for remote sensing products. Continuous product development will be essential. User friendliness and customer support must be emphasized.

Our work heads in this direction by

- developing a general order handling and catalogue system capable of utilizing existing and future user interfacing, storage, and communication facilities.
- continuous product development
- analyzing, evaluating, and developing current and future markets for existing and future products.
- co-ordinating the development of the Tromsø Earth Observation Centre with the national Earth observation strategy outlined by the Norwegian Space Centre (NSC).

BIBLIOGRAPHY

[1] Fredriksen, Å. et. al. , "TEOC - Concept Services Functions", TEOC Rep. No. 1, FORUT 1987
[2] Pedersen, J.P. et. al., "TEOC - User Communities and Product Specification", TEOC Rep. No. 2, FORUT 1987
[3] Hamnes, H. et. al., "TEOC - Kernel Services and Functions Requirements", TEOC Rep. No. 3, FORUT 1988
[4] "The Norwegian Polar Orbiting Satellites Utilization Study", NSC-report(89)3, March 1989

The Remote Sensing Industry in Canada: A Profile of Capabilities

R.A.Ryerson
Canada Centre for Remote Sensing
Energy Mines and Resources Canada
Ottawa, Ontario
K1A 0Y7

J. Kerr
Technology Development Division
External Affairs
Ottawa, Ontario
K1A 0G2

This paper summarizes the capabilities of the 100 Canadian companies that make up the Remote Sensing Industry in Canada. The paper is based on a study undertaken in early 1989 for an export oriented capability guide.

The paper outlines the major areas of technology and services provided by industry and recent significant developments in the technology. It also provides an indication of the wide geographic dispersion of the industry in terms of the location of the companies involved, their employment levels (aggregated), and their major markets in Canada and abroad.

The final section looks to the future with a short discussion of the future expected market.

This paper is intended for the CRSS Proceedings, and is related to the economic topic under Topic D.

REMOTE SENSING:

A TOOL TO ASSIST FOREST MANAGEMENT DECISION MAKING

W. Boyd Brown, B.S.F., R.P.F.
British Columbia Forest Service
Bag 5000, Smithers, B.C., V0J 2N0
Phone: (604) 847-7508, Fax # 847-7217

Dr. Robin V. Quenet, P.H.D., R.P.F.
Nawitka Resource Consultants
840 Cormorant Street, Victoria, B.C., V8W 1R1
Phone: (604) 384-8992, Fax # 384-1841

Ecosystems are complex. Forest managers who must make integrated resource management decisions need a decision assistance tool that provides current, integrated and reliable information. Comprehensive and timely remote sensing products can provide a reliable and accurate source of data for a local decision assistance tool that integrates planning, inventory, silviculture, protection, harvesting (scaling), recreation and other data.

To provide the support necessary to assist forest managers, a computer program has been developed for micro/mini systems. For a specific area, the Lakes Forest District in central British Columbia, Landsat 5 transparences were purchased annually and the changes in forest cover transferred to a hard copy of a digitized forest cover maps using the Procom II. Annual data was captured on a separate digital graphics level, compared to last year's forest cover graphics level and a resultant level created. From the graphics resultant level a change or transaction file was created. Activity reports were generated from this transaction file relative to the activities planned for that year. In the Lakes Forest District low volume per hectare stands constitute a large percentage of the landbase. Logging operators have agreed to harvest the same percentage of low volume stands annually as is found in the total inventory in exchange for having this increased landbase included in their allowable cut plans. Results of the planned vs actual preformance are included.

Actual harvest volumes (scale returns) were compared to the planned harvest volumes on a per hectare basis Silviculture achievements and fire/pest losses were added or subtracted from the projected (modelled) forest growth. The trend of increasing or decreasing total forest resource was reported. The potential of including status, recreation and range annual transaction were also reviewed.

The conclusion of the project is that bringing forest resource date from various sources together into a decision assistance tool can provide much needed support to local decision making. Remote sensing makes the program possible by providing a cheap, reliable, accurate, comprehensive and timely source of information.

Doppler Centroid Estimation Ambiguity for Synthetic Aperture Radars

C.Y. Chang and J.C. Curlander
Jet Propulsion Laboratory
California Institute of Technology
M/S 300-235
4800 Oak Grove Drive
Pasadena, Ca 91109

ABSTRACT

A technique for estimation of the Doppler centroid of synthetic aperture radar (SAR) in the presence of a large antenna boresight pointing uncertainty is described. Also investigated is the image degradation resulting from data processing using an ambiguous centroid. Two approaches for Doppler centroid estimation (DCE) ambiguity resolution are presented: The range cross-correlation technique and The multiple PRF technique. For the multiple PRF technique, since other design factors control the PRF selection for SAR, a generalized algorithm is derived for PRFs not containing a common divisor. An example using the Shuttle Imaging Radar (SIR-C) parameters illustrates that this algorithm is capable of resolving the C-band DCE ambiguities for antenna pointing uncertainties of $2^\circ \sim 3^\circ$.

I. Introduction

The utility of most pulse Doppler radar systems is dependent on the determination accuracy of the Doppler parameters of the echo data, such as the moving target indicator radar (MTIR), where the target velocity estimate is based on its Doppler shift and the synthetic aperture radar (SAR), where the Doppler centroid of the echo is needed for the generation of the azimuth reference function and the range migration compensation [Skolnik, 1970], [Wu, 1982], [Li, 1985]. Due to discrete sampling of the target spectrum at the pulse repetition frequency (PRF), the interval of unambiguous Doppler spectrum is limited to the PRF. Thus, the Doppler centroid estimation ambiguity (DCE ambiguity) occurs if the pointing uncertainty results in a Doppler shift greater than PRF/2 [Li, 1983]. This is very likely to occur for short wavelength spaceborne SARs (e.g., C-band, X-band and above). Table 1 shows the antenna pointing uncertainties, drift rate, azimuth beamwidth, and expected maximum ambiguity number for the Shuttle Imaging Radar (SIR-C), the X-SAR, and the European Remote Sensor (ERS-1) SAR.

In this paper, two approaches are presented to resolve this DCE ambiguity for the SAR. 1) The range cross-correlation technique [Jin, 1981], [Cumming, 1986] and 2) The multiple PRF technique [Skolnik, 1970], [Taylor, 1972]. The range cross-correlation technique is based on the misregistration error between two independent looks of the same SAR image in the range direction. It requires the existence of high contrast targets to reliably detect the ambiguity. The actual pixel shift may be only a fraction of a resolution cell, making precise determination additionally difficult. The multiple PRF technique has been a well known method to resolve the Doppler ambiguity for MTIR. The unique constraint in utilizing this multiple PRF technique for SAR is that other factors in the radar design determine the selected PRF values (i.e., they are not necessarily multiples of a large common factor). The performance of the multiple PRF technique is limited by the available PRF in addition to the unmodeled platform stability (i.e., attitude drift rate) and Doppler centroid estimation error (DCE error).

II. Effects of the DCE Ambiguity on SAR Imagery

The effects of the DCE ambiguity include: 1) Degradation in the point target response in terms of mainlobe broadening and sidelobe values; 2) Degradation in the image signal-to-noise ratio (SNR) and signal-to-azimuth ambiguity level ratio (SALR); 3) Geometric distortion, primarily azimuth skew; and 4) Misregistration error between looks resulting in error in the multilook overlay process.

A simulation of the 2-dimensional point target response by using the SIR-C C-band parameters was conducted to quantitatively evaluate the effects of DCE ambiguity on the point target response. The relevant radar parameters are: wavelength = .0565 m, sensor altitude = 250 Km, sensor velocity = 7500 m/sec, PRF = 1620 Hz, range bandwidth = 10, 20 MHz, chirp duration = 33.8 μsec, and look angle = 30°. The resolution broadening, integrated sidelobe ratio (ISLR), signal-to-noise ratio, misregistration between look-1 and look-4 data, and squint angle uncertainty error are summarized in Table 2. In the calculation of the ISLR, the mainlobe region was fixed at that of

the point target response without ambiguity ($k = 0$). Fig. 1 shows the two-dimensional point target responses at different ambiguity numbers. The degradation in the point target response results mainly from the residual range walk. The range walk at different ambiguity numbers is also shown in Table 2.

III. Range Cross-Correlation Algorithm

The range cross-correlation algorithm is an approach unique to SAR for ambiguity resolution. It is based on the fact that a range misregistration error between different looks of the SAR image is induced if an incorrect Doppler centroid is used for the range walk compensation. The range misregistration error, ΔR, between two time-separated (or frequency-separated) single-look images centered at t_a and t_b, is

$$\Delta R = |t_a - t_b| \frac{\lambda(k\ PRF)}{2} , \tag{1}$$

where k is the ambiguity number, and $\Delta T = t_a - t_b$ is the time separation between the centers of the two looks. The misregistration error, and hence the DCE ambiguity, can be detected by cross-correlating two different looks of images in range.

The advantages of the range cross-correlation algorithm compared to the multiple PRF algorithm are that its performance is not as sensitive to the available PRFs, the platform stability, and the DCE uncertainty. Its disadvantage is that its performance is highly scene dependent and is degraded by the speckle noise. High contrast targets are required to yield accurate results. For the subpixel misregistration, the ambiguity is difficult to resolve using range cross-correlation due to the speckle noise effect in the 1-look SAR imagery. Hence, the range cross-correlation technique can only reliably approximate the ambiguity number.

IV. Multiple PRF Algorithm

The multiple PRF technique has been a well-known method for ambiguity resolution for MTIR. The unique constraint in utilizing this multiple PRF technique for SAR is that other factors in the radar design determine the selected PRF values (i.e., they are not necessarily multiples of a large common factor). For the SIR-C, the multiple PRF switching is planned to be part of the radar on/off sequences. Three selected PRFs are cycled through at the beginning of each data take. The dwell time for each PRF is designed to be one second long. The third PRF is used throughout the data take (approximately 10 minutes long). An inverted PRF sequence is repeated at the conclusion of the data take which will be used by the processor to verify the Doppler tracking.

In general, the selected PRFs do not contain a large common factor. One obvious way to resolve the ambiguity is a brute force search through all the possible combinations of PRFs. This section presents a fast algorithm to resolve the DCE ambiguity. Since, three PRFs are necessary and sufficient for SIR-C, the case of three PRFs is presented.

Let PRF_1, PRF_2, and PRF_3 be three of the PRFs used to resolve the DCE ambiguity and $\hat{f}_{d1}$, $\hat{f}_{d2}$, and $\hat{f}_{d3}$ be three ambiguous Doppler centroid estimates (the observed Doppler centroids), which are obtained from the Doppler centroid estimation algorithm (DCE algorithm). Due to the repeating frequency spectrum, $0 \le \hat{f}_{d1} < PRF_1$, $0 \le \hat{f}_{d2} < PRF_2$, and $0 \le \hat{f}_{d3} < PRF_3$. Let f_{d1}, f_{d2}, and f_{d3} be the unambiguous Doppler centroid estimates for these three different PRFs. We have

$$f_{d1} = k_1 PRF_1 + \hat{f}_{d1} \tag{2a}$$

$$f_{d2} = k_2 PRF_2 + \hat{f}_{d2} \tag{2b}$$

$$f_{d3} = k_3 PRF_3 + \hat{f}_{d3} , \tag{2c}$$

where k_1, k_2, and k_3 represent the ambiguity number. Ideally $f_{d1} = f_{d2} = f_{d3}$. Due to the DCE error and antenna drift, they are not equal in general. The effect of the antenna drift rate can be partially compensated by utilizing the antenna attitude drift rate information provided in the ephemeris data such that the difference in f_{d1}, f_{d2}, f_{d3} is only subject to the DCE error and undetermined attitude drift, i.e.,

$$\hat{f}_{d2} = \hat{f}_{d2} - \Delta f_{12} \tag{3a}$$

$$\hat{f}_{d3} = \hat{f}_{d3} - \Delta f_{13} , \tag{3b}$$

where Δf_{12} represents the Doppler drift between PRF_1 and PRF_2 and Δf_{13} the drift between PRF_1 and PRF_3. Both drifts are estimated from the ephemeris data.

Doppler Centroid Estimation Ambiguity Resolution Algorithm
[input: PRF_1, $\hat{f}_{d1}$, PRF_2, $\hat{f}_{d2}$, PRF_3, $\hat{f}_{d3}$; output: f_{d1}, f_{d2}, f_{d3}, δ_1, δ_2]
Let

$$\Delta PRF = |PRF_2 - PRF_1| \tag{4a}$$

and

$$m_2 = \frac{PRF_2}{\Delta PRF} , \tag{4b}$$

where m_2 is a real number. The solution of (2a) can be expressed as a function of an integer i as

$$f_{d1} = PRF_1 * \text{rnd}\left[\frac{\hat{f}_{d2} - \hat{f}_{d1}}{\Delta PRF} + i * m_2\right] + \hat{f}_{d1} , \tag{5}$$

where "rnd" represents the integer round off operation. Compared to the brute force search, the complexity is reduced by a

factor of m_2. The optimal $\hat{i}$ can be obtained by computing

$$\delta_1 = |[f_{d1} - (k_2\, PRF_2 + \hat{f}_{d2})] \bmod PRF_2| = |(f_{d1} - \hat{f}_{d2}) \bmod PRF_2| \quad (6a)$$

$$\delta_2 = |[f_{d1} - (k_3\, PRF_3 + \hat{f}_{d3})] \bmod PRF_3| = |(f_{d1} - \hat{f}_{d3}) \bmod PRF_3| \quad (6b)$$

and minimizing

$$\delta_{min} = \sqrt{\delta_1^2 + \delta_2^2}\ , \quad (7)$$

where $\delta_1 \le PRF_2/2$ and $\delta_2 \le PRF_3/2$. Let $\hat{k}_2$ and $\hat{k}_3$ be the solutions that minimize (7). Then,

$$f_{d1} = PRF_1 * \mathrm{rnd}\left[\frac{\hat{f}_{d2} - \hat{f}_{d1}}{\Delta PRF} + \hat{i} * m_2\right] + \hat{f}_{d1} \quad (8a)$$

$$f_{d2} = \hat{k}_2 PRF_2 + \hat{f}_{d2} \quad (8b)$$

$$f_{d3} = \hat{k}_3 PRF_3 + \hat{f}_{d3}\ . \quad (8c)$$

##

If δ_{min} is smaller than some threshold, the ambiguity is claimed to be resolved. If it is greater than the threshold, the ambiguity is not completely removed due to the presence of large DCE error or unmodeled antenna drift. The choice of threshold determines the probability of ambiguity resolution.

V. Performance of the Multiple PRF Algorithm

Let ϵ represent the DCE error and Δt_{12} be the time between the the first and second DCEs. To avoid an ambiguity in (5), the maximum acceptable undetermined antenna drift rate in squint angle (i.e., combination of yaw and pitch) is

$$\phi'_{max} \simeq \frac{\lambda}{2V\Delta t_{12}}\left(\frac{\Delta PRF}{2} - \epsilon\right)\ . \quad (9)$$

Let ϕ be the unknown azimuth squint angle of the antenna beam due to the antenna attitude error and the maximum unambiguous Doppler frequency interval of the multiple PRF technique be denoted by $l.c.m.(PRF)$. Then, the maximum squint angle is

$$\phi_{max} \simeq \sin^{-1}\left(\frac{\lambda * l.c.m.(PRF)}{4V}\right) \quad (10)$$

The DCE ambiguity is resolved when $|\phi| < \phi_{max}$ given by (10) and $|\phi'| < \phi'_{max}$ given by (9).

By using the given PRF set for SIR-C C-band, an example of the maximum acceptable squint angle uncertainties and drift rates at different look angles is shown in **Fig. 2**. It shows that at many of the look angles, the maximum antenna attitude uncertainty in squint angle, ϕ_{max}, can be as high as $2^\circ \sim 3^\circ$ and the maximum unmodeled antenna drift rate in squint angle, ϕ'_{max}, can be as high as $.026^\circ$/sec.

VI. Concluding Remarks

Two different approaches were presented to resolve the DCE ambiguity: range cross-correlation and multiple PRF technique. The range cross-correlation technique is achieved by cross-correlating two single-look images of the same target to detect the cross-track misregistration. High contrast targets in the SAR data are required. The performance of the multiple PRF technique is primarily limited by the system design (i.e., available PRF, unknown antenna drift rate, and DCE error). An example using the Shuttle Imaging Radar (SIR-C) parameters illustrates that this algorithm is capable of resolving the C-band DCE ambiguities for antenna pointing uncertainties of $2^\circ \sim 3^\circ$ and unmodeled antenna drift rate $\sim .026^\circ$/sec.

VII. Acknowledgement

The research described in this paper was performed at the Jet Propulsion Laboratory, California Institute of Technology, under contract with the National Aeronautics and Space Administration.

References

1. M.I. Skolnik, Radar Handbook, McGraw-Hill, New York, 1970.

2. C. Wu, et al., "Modeling and a Correlation Algorithm for Spaceborne SAR Signals," IEEE Trans. on AES, Vol. AES-18, No. 5, Sept. 1982, pp. 563-575.

3. F.K. Li, et al., "Doppler Parameter Estimation for Spaceborne Synthetic Aperture Radars," IEEE Trans. on GARS, Vol. GE-23, No. 1, Jan. 1985, pp. 47-56.

4. F.K. Li, et al., "Ambiguities in Spaceborne Synthetic Aperture Radar Systems," IEEE Trans. on AES, Vol., AES-19, No. 3, May 1983, pp. 389-397.

5. M. Jin, "Doppler Ambiguity on Range Migration Correction in a C or X-band SAR," Internal JPL Memorandum, 1981.

6. I.G. Cumming, et al.,"Resolving the Doppler Ambiguity for Spaceborne Synthetic Aperture Radar," Proc. of IGARSS, Zurich, Switzerland, Sept. 1986, pp. 1639-1643.

7. F.J. Taylor, "Pulse Doppler Ambiguity Resolution," IEEE Trans. on AES, AES-8, No. 1, Jan. 1972, pp. 591-595.

	SIR-C		X-SAR	ERS-1
Pointing Error (3σ)	roll:±1.24° yaw:±1.43° pitch:±1.78°			roll:±0.11° yaw:±0.17° pitch:±0.13°
Drift Rate (3σ)	±0.033°/sec in roll, yaw, pitch			±0.0015°/sec in roll, yaw, pitch
Transmitter Frequencies	L-band 1.26 GHz	C-band 5.31 GHz	X-band 9.60 GHz	C-band 5.31 GHz
Azimuth Beamwidth	1.13°	0.27°	0.15°	0.32°
Maximum Ambiguity Number	±2.0	±8.5	±15.2	±0.7

Table 1: SIR-C, X-SAR, and ERS-1 parameters.

range chirp BW	10 MHz				20 MHz			
k	1	2	3	4	1	2	3	4
2-D resolution (normalized)	1.04	1.25	1.84	4.10	1.22	4.21	13.4	25.5
2-D ISLR (degradation in dB)	0.4	1.4	3.2	5.3	1.5	5.5	10.2	13.6
SNR (loss in dB)	0.4	1.5	3.4	5.8	1.5	5.8	9.7	12.2
ΔR (pixels)	0.48	0.97	1.45	1.94	0.97	1.93	2.90	3.87
ϕ (degrees)	0.35	0.70	1.05	1.40	0.35	0.70	1.05	1.40
range walk (pixels)	0.84	1.49	2.13	2.78	1.69	2.98	4.27	5.56

Table 2: Simulation results of the DCE ambiguity effects for the SIR-C C-band two-dimensional point target response. The relevant parameters are: wavelength = .0565 m, sensor altitude = 250 Km, sensor velocity = 7500 m/sec, PRF = 1620 Hz, range bandwidth = 10, 20 MHz, chirp duration = 33.8 μsec, and look angle = 30°. The range pixel spacing is 13.3 m and 6.7 m, respectively. k: ambiguity number, 2-D resolution: 2-dimensional resolution, 2-D ISLR: 2-dimensional integrated sidelobe ratio; ΔR: misregistration error between look-1 and look-4 images, ϕ: azimuth squint angle uncertainty.

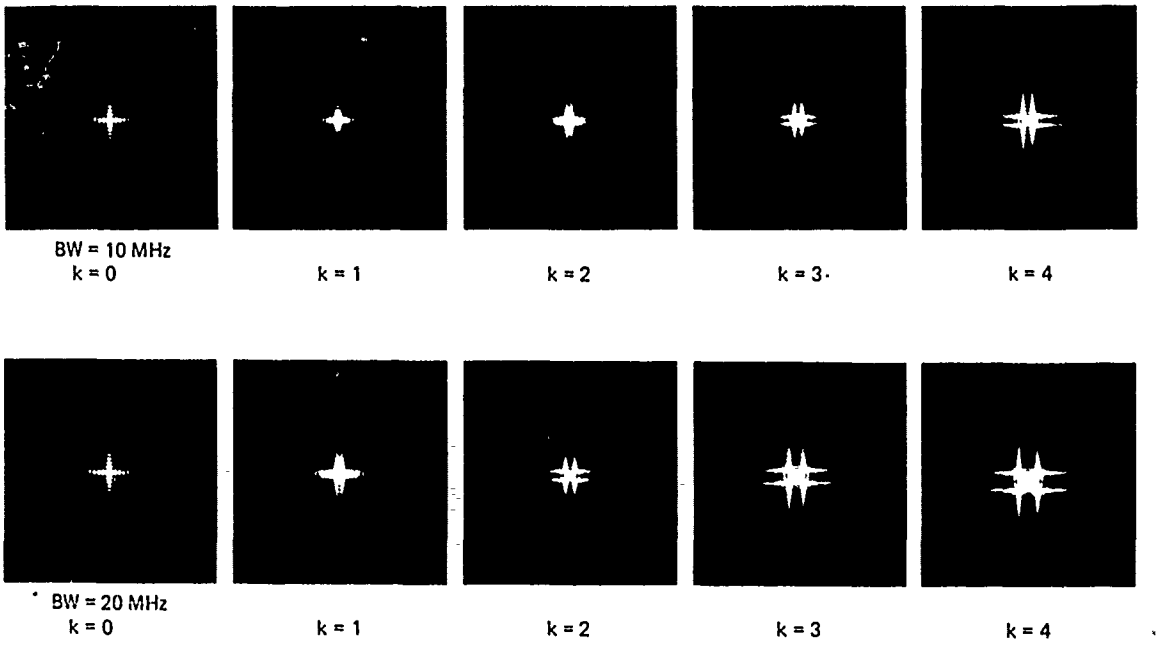


Fig. 1: The SIR-C simulated point target responses as a function of ambiguity numbers. BW: range bandwidth, k: ambiguity number. The results are summarized in Table 2.

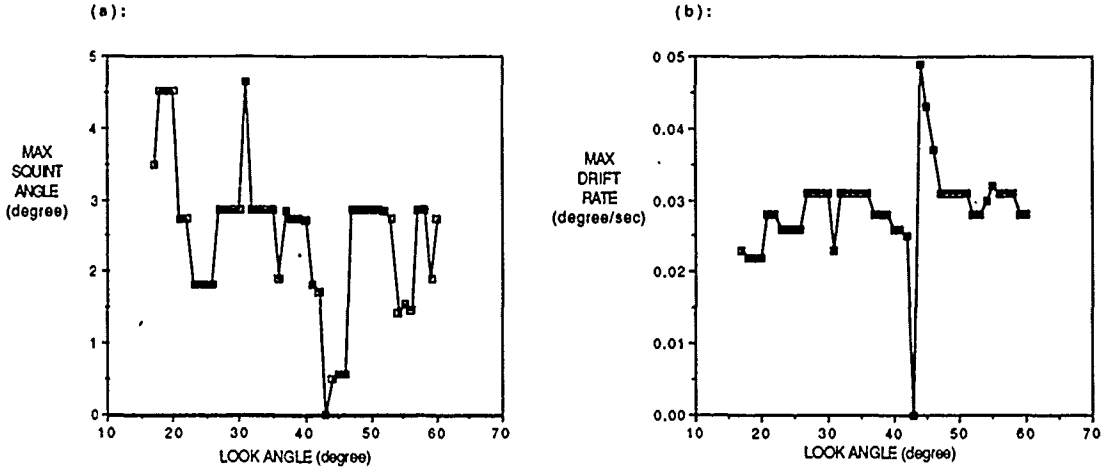


Fig. 2. An example of the maximum squint angle and drift rate by using the multiple PRF switching technique for the given SIR-C PRF set. (a): maximum squint angle. (b): maximum drift rate.

PRF AMBIGUITY RESOLVING FOR SAR

Hartmut Runge
Richard Bamler

DLR
German Aerospace Research Establishment
WT-DA/MV SAR
D-8031 Oberpfaffenhofen
West Germany

Abstract

For high precision SAR (Synthetic Aperture Radar) processing, the determination of the absolute Doppler Centroid is indispensable. The Doppler frequency estimated from azimuth spectra, however, suffers from the fact that the data are sampled with the PRF and an ambiguity about the correct PRF-band remains. Five methods for Ambiguity Resolving are proposed and discussed together with the already known technique of look correlation. None of these methods have a requirement on the mission schedule. It is shown that the following effects can be used to measure the absolute Doppler frequency: the Doppler shift of range spectra, range migration, image geometric misregistration and the use of multifrequency radar data.

Keywords:
SAR Processing, SIR-C, X-SAR, PRF Ambiguity, Doppler Ambiguity

1. Introduction

The processing of SAR (Synthetic Aperture Radar) data requires the knowledge of the absolute Doppler Centroid frequency (f_{DC}) for each data segment. The f_{DC} is defined as the Doppler frequency caused by the relative motion between the sensor and a target in the azimuth beam center line. Due to earth rotation and a sensor pointing angle different from the boresight, the value of f_{DC} can vary considerably. The amount of variation increases with the radar carrier frequency and sensor attitude instabilities and uncertainties.

The Doppler Centroid frequency is usually determined by an azimuth spectral analysis [Li, 1985]. Due to the pulsed nature of SAR systems, the azimuth signal is sampled by the Pulse Repetition Frequency (PRF). Hence an azimuth spectral estimate $f_{DC,base}$ is always confined to the baseband.

$$-PRF/2 \leq f_{DC,base} \leq PRF/2$$

and thus suffers from an ambiguity modulo PRF:

$$f_{DC} = f_{DC,base} + n \cdot PRF$$

An advanced Doppler Centroid Tracker should be able to follow the Doppler frequency variation even across the PRF-band boundaries. Thus it will update the number n relative to its starting point.

The subject of 'Ambiguity Resolving' is the determination of the absolute value of the number n which will be taken as a bias for the Doppler Tracker. This needs to be done only once during a data take. The required absolute accuracy of the Ambiguity Resolver is PRF/2. In a SAR system, the PRF is adjusted to the azimuth bandwidth, but is limited by data rate constraints and the risk of range ambiguities. For an L-band SAR on a properly stabilized platform, like the Seasat satellite, the ambiguity number n in equ. (1) can be calculated from orbit and nominal attitude data. For the X-band radar [Wahl, 1988, Miller, 1989] on the SIR-C shuttle mission [Huneycutt, 1989], however, the uncertainty of the ambiguity number n can be up to $\pm$ 20, because of the unprecise attitude data. The use of a wrong value of n for the SAR image formation process results in a loss of spatial resolution [Li, 1983] and an image misregistration [Curlander, 1982]. Therefore, the resolving of the PRF ambiguity is essential for high precision SAR processing.

In this paper we will discuss methods that use only the SAR raw data itself and have no major implications on the mission time lining like the 'PRF hopping' method [Li, 1983] or the transmission of a pilot tone. The methods described in the following are based on either the Doppler shift of single radar echoes or on effects caused by sensor/target geometry which are not affected by the PRF sampling

2. Doppler Shift of the Range Spectrum

Each received radar return is subject to the unknown absolute Doppler shift. Hence the range power spectra are shifted by f_{DC} relative to the nominal chirp spectrum. In contrast to the PRF, the range sampling rate is much higher than the maximum Doppler shift. It should thus be possible to measure f_{DC} from the range spectra without ambiguity. The shift between the spectra may be determined by a frequency discrimination technique or methods originally developped for the azimuth Doppler Centroid estimation like the crosscorrelation or energy balancing method described by [Li, 1985] and [Madsen, 1986]. The resolution requirements of such a spectral estimator are however enormous: the typical range chirp bandwidth of a spaceborne SAR is in the order of 20 MHz, while the shift must be measured with an accuracy of at least 700 Hz, assuming a PRF of 1400 Hz. This requirement can only be met by averaging a vast amount of range spectra in order to suppress the speckle noise (see Fig. 1).

It can be shown both experimentally and theoretically that at least 50 000 range spectra must be averaged to achieve the required resolution. This need not necessarily lead to a high computation time, because this procedure might be incorporated into the spectral domain range compression. In a high frequency SAR the f_{DC} can vary in the range direction by several PRF bands. This fact makes a range segmentation necessary for the Ambiguity Resolving. If the size of the segments comes close to the length of a range chirp, a range compression prior to the segmentation is required. Care must be taken not to alter the envelope of the range spectrum which contains the information about the Doppler shift. The requirement for the range spectrum preservation must be met by the whole data aquisition chain, especially by the filters in the the RF and IF electronics. We assume that bandpass filters not designed according to this requirement prevented us so far from a successful demonstration of this method with Seasat data. However, the theoretical performance of this method is carrier frequency independent.

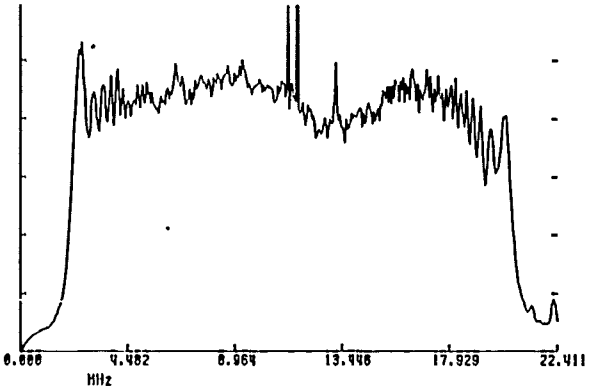


Figure 1: Average of 1000 Seasat range power spectra. Note the still high amount of speckle noise.

3. Methods based on Range Migration Measurements

Another phenomenon caused by relative sensor/target motion is the range migration, i.e. the variation of the slant range r(t) of a target passing through the antenna beam. The linear range migration is proportional to the unknown absolute Doppler frequency f_{DC} and is not subject to PRF undersampling. Thus the slope of the target trace in azimuth direction is a measure for f_{DC} (Fig. 2a):

$$\frac{dr(t)}{dt} = -\frac{1}{\vartheta} \cdot f_{DC} = \tan\alpha \qquad (2)$$

With r(t) given in units of echo time and ϑ being the radar carrier frequency The slope can be determined in three different ways:

The Look Correlation Technique
As mentioned by [Luscombe, 1982] and shown by [Cumming, 1986] the amount of range walk can be measured by crosscorrelating the two outer looks. This method works best on scenes with high contrast. Since range walk depends on both range migration and antenna integration time, the accuracy of the look correlation technique is proportional to λ^2 .

2D-Spectrum Analysis
The slope of the azimuth trace corresponds to a skew of its two-dimensional power spectrum (see Fig. 2b). The Doppler frequency f_{DC} can be determined by calculating the inclination of the principal axis of the spectrum. Since only the range migration and not the antenna integration time determines the sensitivity of this method its accuracy is proportional to the radar wavelength λ.

2D-Autocorrelation Method
The time domain counterpart to the preceeding method exploits the fact that the two dimensional autocorrelation function (ACF) of the azimuth chirp trace is also skewed proportional to the range migration (Fig. 2c). A measure of the skewing is the offset Δ shown in Fig. 3. In order to measure Δ it is sufficient to calculate the 2-D ACF only for an azimuth lag of one sample. This single column of the 2-D ACF can be computed by crosscorrelating each range line with the consecutive one and complex summation of the results. These individual crosscorrelations can be efficiently carried out in frequency domain. This allows a combination of this Ambiguity Resolving technique with the range compression task. It should be noted that the amount of the shift Δ for an f_{DC} which equals the PRF is only the reciprocal of the radar carrier frequency. The accuracy which can be gained by the 2-D ACF method is proportional to λ.

4. Use of a Multifrequency SAR

In a multifrequency SAR experiment like the SIR-C Radar Lab the lower frequency channels (L- and C-band) may be used to resolve the Doppler ambiguity in the X-band. Due to the different antenna opening angles one PRF band in the L-channel (1,256 GHz) corresponds to 4,24 PRF bands in the C-channel (5,36 GHz) and to 7,68 PRF bands in the X-channel (9,6 GHz). Therefore, a Doppler Centroid measurement in a lower frequency band can be used to determine the correct PRF band in the higher frequency channel. This method requires the knowledge of a possible azimuth misalignment of the three antennas with respect to each other. Such a misalignment can be measured by a geometric comparison between the processed images of the same scene from different radar channels.

5. The Tiepointing Method

The use of a wrong PRF band in the geolocation process results in a misregistration of the image pixels by an integer multiple of the antenna azimuth footprint. By matching the tiepoints between a map and the SAR image pixels the Doppler ambiguity number can be directly obtained. Remember that this somewhat time consuming process has to be carried out only once per data take. The accuracy of this method is proportional to λ.

6. Conclusion

It was shown that several different techniques are applicable for PRF Ambiguity Resolving. Each of them has its own merrits and drawbacks. Since the Doppler Ambiguity

problem arises mainly in high frequency radars, a comparison of the different algorithms with respect to their wavelength dependance is of interest.

We found that only the technique which uses the Doppler shift of the range spectrum is independant of the carrier frequency. All other methods, especially the look correlation technique, lose accuracy with increasing radar frequency. The tiepointing method and the look correlation technique work on detected SAR images and hence require a good scene contrast. The other raw data based methods work best with homogeneous scenes like sea surfaces. We want to emphazise that a reliable Doppler Tracker is indispensable since an Ambiguity Resolver can only provide a single ambiguity number for a scene or a whole data take. The summary of different methods provided in this paper is intended to be a base for further detailed theoretical and experimental investigations towards an operational PRF Ambiguity Resolver.

7. Acknowledgement

The authors like to thank James Weymes of Sherrington's Software Services for the excellent programming job

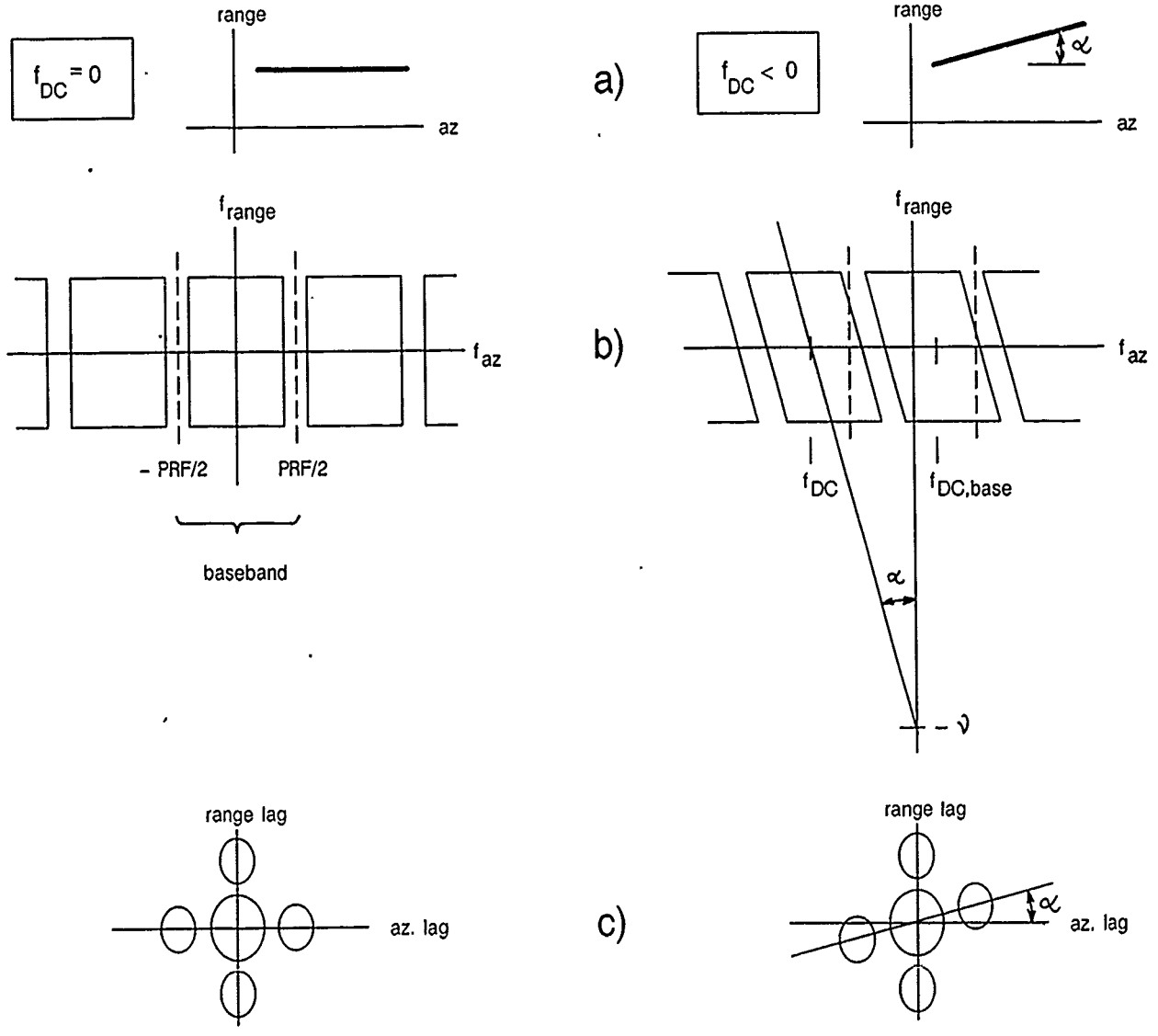


Figure 2: Azimuth chirp trace (a), two dimensional power spectrum (b), and autocorrelation function (c) for $f_{DC} = 0$ (left) and $f_{DC} < 0$ (right). Slope angle α is exaggerated for illustration.

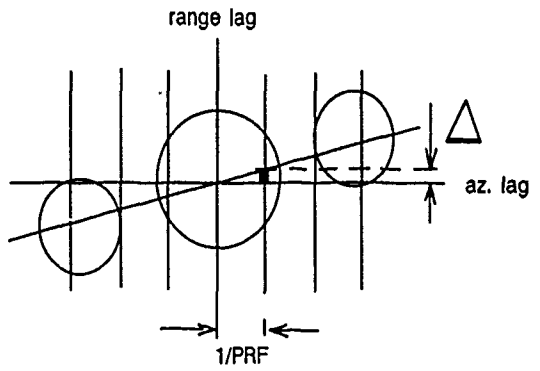


Figure 3: Detailed sketch of autocorrelation function of SAR data with linear range migration.

8. Bibliography

Li, F.K., Held, D.N., Curlander, J. and Wu, C., "Doppler Parameter Estimation For Synthetic Aperture Radars", IEEE Transactions On Geoscience and Remote Sensing, Vol. GE-23, No. 1, pp 47-56, 1985.

Curlander, J.C., "Location Of Spaceborne SAR Imagery", IEEE Transactions On Geoscience And Remote Sensing, Vol. GE-20, No. 3, pp. 359-364, 1982.

Wahl, M., Ammendola, P., "Main Features Of The X-SAR Project", Proceedings of IGARSS '88 Symposium, Edinburgh, pp. 1018 - 1020, Ref. ESA SP - 284, 1988.

Miller, D., "X-SAR Specification, Design and Performance Modeling", IEEE Transactions on Geoscience and Remote Sensing, Vol. 27, No. 2, 1989.

Huneycutt, B.L., "Spaceborne Imaging Radar-C Instrument", IEEE Transactions on Geoscience and Remote Sensing, Vol. 27, No. 2, 1989.

Li, F.K., Johnson, W.T.K., "Ambiguities in Spaceborne Synthetic Aperture Radar Systems", IEEE Trans. on Aerospace and Electronic Systems, Vol. AES-19, No. 3, 1983.

Madsen, S.N., "Speckle Theory", Ph.D. Thesis LD62, Technical University of Denmark, 1986.

Luscombe, A.P., "Auxiliary Data Networks for Satellite Synthetic Aperture Radar", Marconi Review, Vol. XLV, No. 225, 1982.

Cumming, F.G., Kavanagh, P.F., Ito, M.R., "Resolving The Doppler Ambiguity For Spaceborne Synthetic Aperture Radar", Proceedings of IGARSS'86 Symposium, pp. 1639 - 1643, Zürich, Ref. ESA SP-254, 1986.

THEORY AND QUANTITATIVE COMPARISON OF DOPPLER CENTROID ESTIMATION METHODS

J. Siewerth

German Aerospace Research Establishment (DLR)
WT-DA/MV SAR
D-8031 Oberpfaffenhofen
West Germany

Abstract

The purpose of this paper is to describe the theory and implementation of three different Doppler centroid estimation methods and to present the first results,of currently performed quantitative investigations.

The Doppler centroid shift caused by the relative velocity between the sensor platform and the targets is derived by analysing the received SAR data. In contrast to the conventionally used ΔE-method (also called 'energy balancing'), which is a frequency approach, the two other methods, the Correlation Doppler Estimator (CDE) and the Sign Doppler Estimator (SDE), are both performed in the time domain.

Keywords: Doppler centroid estimation, ERS-1, Intelligent SAR Processor

Introduction

The quality of a processed SAR image strongly depends on the accuracy of the estimation of the expected phase histories of the illuminated targets. Estimation errors lead to errors in the generation of the azimuth processing matched filter and thus to a degradation of image quality. More precisely, particularly a linear phase term estimation error (i.e. Doppler centroid error) leads to a significant degradation of the signal-to-noise ratio, the signal-to-azimuth ambiguity level and to an azimuth shift of the pixel location (Chang et al., 1988). In general, the Doppler centroid estimation must be accurate enough to meet specific SAR image quality requirements.

As the sensor ephemeris data and sensor attitude data are not sufficiently precise to provide accurate Doppler centroid estimation, the Doppler information has to be derived by analysing the coherent radar return.

Within the scope of the ERS-1 ground segment project, DLR (German Aerospace Research Establishment) is going to develop an Intelligent SAR (ISAR) processor. Specifically, in this German ISAR processor, as it is designed as an operational processor, an automatic method for estimating the Doppler centroid, using the radar echo return, must be employed. This automatic module also has to fit the three basic requirements of the overal ISAR processor system well, i.e. 'high quality', 'high throughput' and 'high flexibility'. These three 'major operational requirements' lead to some specific requirements due to the Doppler centroid estimation.

High quality images can only be achieved if the azimuth reference function is a precise matched filter of the radar echo return and the azimuth reference function is updated frequently enough, both in range and azimuth dimensions. In order to perform SAR processing within a minimum time span, the Doppler centroid estimation must also be performed within the shortest possible duration. Thirdly, the processing of radar echo data from various sensor types acquired over different terrain types must be possible.

In a first step toward the selection of an appropriate algorithm, three different Doppler centroid estimation methods have been implemented. The algorithms were taken from (Madsen, 1986) but have been adapted to the specific usage within the ISAR processor. In addition, these prototypes have been written using the programming language ADA, in order to obtain initial experience in the language in which the whole ISAR processor will be built up around.

At the moment, first investigations are being performed using only simulated ERS-1 point targets and some SEASAT scenes. The ISAR raw data simulator was used to generate ERS-1 sensor related raw data from single and multiple point targets. All simulations were based on the ERS-1 3 day reference orbit, assuming nominal operation (yaw-steering) mode. For a final assessment, the algorithms have still to be tested on a greater number of SEASAT scenes.

In the following, the implemented Doppler centroid estimation algorithm will be described and subsequent first results of the investigations will be presented.

The ΔE-Algorithm

The ΔE-algorithm is a well known frequency domain approach, which has already been described by (Curlander et al., 1982), by (Li et al., 1985) and by (Madsen, 1986).

The algorithm makes use of one of the basic SAR properties in azimuth dimension, that the frequency at beam centre, the Doppler centroid, corresponds to the maximum power of the antenna azimuth pattern in the frequency domain.

To detect this frequency which corresponds to the maximum power, the discrete azimuth signal a_r, at a specific range r, is transformed in the frequency domain via a Discrete Fourier Transformation

$$A_r\left(\frac{n}{N_a T}\right) = \sum_{k=0}^{N_a-1} a_r(kT)\,\exp\left[-j\,\frac{2\pi}{N_a}\,nk\right] \qquad n = 0, 1, \ldots, N_a - 1\ ;$$

N_a number of azimuth samples
T pulse repetition interval

and the periodogram P_r of the spectrum is calculated.

$$P_r\left(\frac{n}{N_a T}\right) = \left| A_r\left(\frac{n}{N_a T}\right) \right|^2 ;$$

Figure 1 shows the azimuth periodogram of 11 simulated point targets.

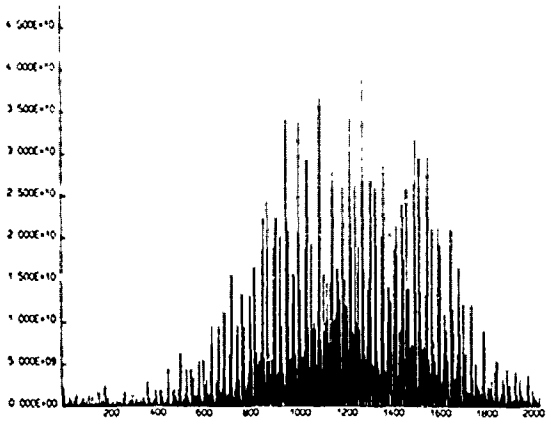


Figure 1: Periodogram of 11 simulated point targets

For reducing the variance to mean square ratio of the periodogram, the accumulated periodogram P_{ac} for several azimuth periodograms at range $n \leq r \leq r_u$ is calculated (Figure 2).

$$P_{ac}\left(\frac{n}{N_a T}\right) = \frac{1}{N_r} \sum_{r=r_1}^{r_u} P_r\left(\frac{n}{N_a T}\right) ;$$

N_r number of accumulated azimuth periodograms
n lower range value boundary
r_u upper range value boundary

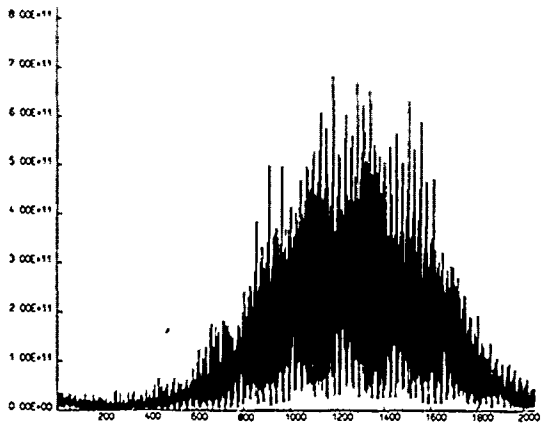


Figure 2: Accumulated power spectrum of 11 simulated point targets (N_r = 10)

This accumulated periodogram is now considered as a function, where the slowly varying antenna pattern function is multiplicatively superimposed by the amount of high frequent target backscatter information. Thus in addition to the Madsen approach, an appropriate low-pass filter is applied to the accumulated periodogramm. The result of low-pass filtering is a modified azimuth periodogram, where the high frequencies, corresponding to the target information, are discarded and almost only the slowly varying antenna pattern function remains (Figure 3).

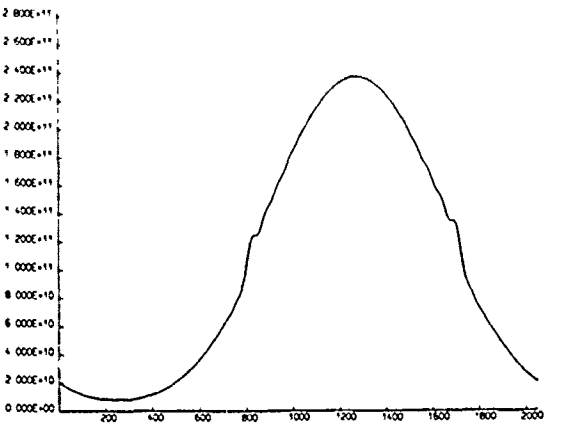


Figure 3: Accumulated power spectrum of 11 simulated point targets after low-pass filtering

The frequency, corresponding to the maximum power of this modified azimuth periodogram and therefore, the Doppler centroid frequency (modulo f_{PRF}), is related to the bin belonging to the curve maximum of Figure 3. The usual energy comparison (Madsen, 1986) is unnecessary if an appropriate low-pass filter is applied to smooth the periodogram.

The CDE-Algorithm

Although the CDE-algorithm takes place in the time domain of the azimuth signal, the approach is quite similar to the ΔE-method. In constrast to the frequency domain algorithm, the respective processing steps are performed using the corresponding time domain equations.

The shift of the periodogram of the azimuth signal, due to the Doppler centroid, is hereby measured by calculating the phase shift of the corresponding time domain function.

First the autocorrelation function R_r of the discrete azimuth signal a_r at a specific range r, is calculated

$$R_r(kT) = \sum_{i=0}^{N_a-1} a_r((k+i)T) \cdot a_r(iT) ;$$

N_a number of azimuth samples
T pulse repetition interval

R_r corresponds to the azimuth periodogram P_r via the Discrete Fourier Transformation.

$$R_r(kT) \circ\!\!-\!\!\bullet\, \left| A_r\left(\frac{n}{N_a T}\right) \right|^2 = R_r\left(\frac{n}{N_a T}\right) ;$$

Analogous to the ΔE-Algorithm, several autocorrelation functions at range values $n \leq r \leq r_u$ are summed.

$$R_{ac}(kT) = \frac{1}{N_r} \sum_{r=1}^{r_u} R_r(kT) ;$$

N_r number of range values
n lower range value boundary
r_u upper range value boundary

From the so called shifting property of the Fourier Transformation, it is known that the inverse Fourier Transformation of a shifted periodogram is simply the autocorrelation function of the unshifted function, multiplied by an exponential factor, having a linear phase (Gaskill, 1978).

$$P_r\left(\frac{n-i}{N_a T}\right) \bullet\!\!-\!\!\circ\, R_r(kT) \cdot \exp\left[j \frac{2\pi}{N_a} i k\right] ;$$

Thus the phase difference between the autocorrelation function R_{ac} and the nominal (zero-Doppler) autocorrelation function R_o is proportional to the Doppler shift.

As the SAR azimuth signal is sampled with a rate which is not much higher than the Nyquist rate, the phases of the respective autocorrelation functions are calculated at the correlation coefficient j = 1, to get a high unambiguous Doppler range (modulo PRF). After calculating the phase difference Δarg,

$$\Delta arg = arg\,\{R_{ac}(T)\} - arg\,\{R_o(T)\}\ ;$$

the Doppler Centroid estimation value is then given by

$$f_D = \left[\frac{1}{2\pi T} \cdot \Delta arg\right] \pm n \cdot f_{PRF}, \qquad n = 0, \pm 1, \pm 2, \ldots\ ;$$

Note that the selection of the correlation coefficient j = 1 corresponds to the first harmonic of the azimuth periodogram. Thus, comparing the phases at j = 1 is analogous to the use of a very restrictive low-pass filter in the frequency domain approach. Figure 4 shows the first harmonic of Figure 3, which corresponds to the correlation coefficient of R_{ac} at j = 1.

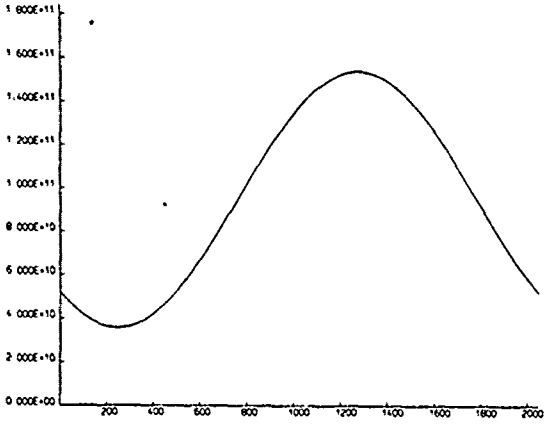


Figure 4: First harmonic of 11 simulated point targets

The SDE-Algorithm

Just as the CDE-Algorithm, so the SDE-Algorithm also uses correlation coefficients to estimate the Doppler centroid.

The difference between both methods is only the way in which the autocorrelation function R_{ac} is calculated. While the CDE- algorithm uses the well-known standard equation for the calculation of the auto-correlation function of the azimuth signal, the basic idea of the SDE-Method is to use the so called 'Arcsine Law' of Gaussian processes (Papoulis, 1965) to calculate the autocorrelation function.

The central statement of the Arcsine law is, that if the real part and the imaginary part of a complex digital signal are nearly Gaussian processes (which is fulfilled for SAR azimuth signals), the autocorrelation function can be calculated only by examining their signs.

Results

The first investigations of the Doppler centroid estimation algorithm were performed using only simulated ERS-1 point targets. As expected all three different methods yield the exact frequency, which has been computed by the ERS-1 simulator. The results varied only by 0.1Hz around the exact Doppler frequency. After these tests performed on idealised SAR data, the algorithms were applied to the well known SEASAT 'Goldstone scene'. The results are shown in Figure 5. The solid curve represents the accumulated and low-pass filtered azimuth periodogram, calculated by the ΔE-algorithm. The broken curve depicts the azimuth periodogram which corresponds, via the Fourier Transformation, to the correlation coefficient of the autocorrelation function, calculated by the CDE- and SDE-algorithms.

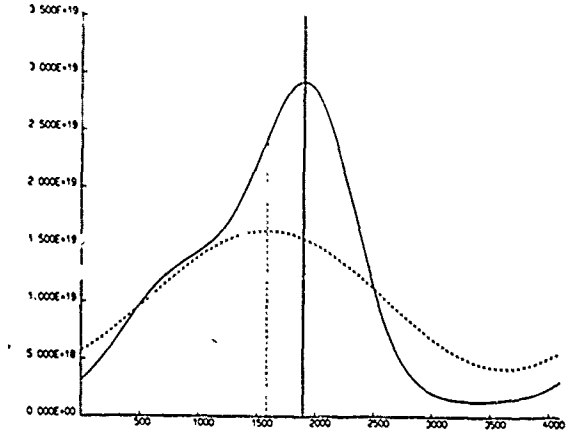


Figure 5: Azimuth periodograms calculated from the Goldstone scene

The difference between the frequency corresponding to the peak of the solid curve and the frequency corresponding to the peak of the broken curve is approximately 25 Hz. The comparison of the estimation results to the Doppler centroid frequency, which has been computed by the GSAR (Generalized Synthetic Aparture Radar Processor, developed by MacDonald Dettwiler and Associates Ltd), has shown that the frequency corresponding to the solid peak curve is in agreement with the frequency calculated by the GSAR. This leads to the preliminary result that the Doppler centroid frequency cannot be determined sufficiently precisely if only the first harmonic of the azimuth periodogram is examined.

Subject to further investigation, the Doppler centroid estimation method used for the German ISAR processor should be based on the ΔE-algorithm, using an appropriate low-pass filter.

References

1. Madsen, S.N., "Speckle Theory Modelling, analysis, and applications related to Synthetic Aperture Radar data", Ph.D. thesis, 1986.

2. Li, F.K., Held, D.N., Curlander, J.C., Wu, C., "Doppler Parameter Estimation for Spaceborne Synthetic-Aperture Radars", IEEE Transactions on Geoscience and Remote Sensing, Vol. GE-23, No. 1, pp. 47-56, 1985.

3. Chang, C.Y., Curlander, J.C., "Data Products and Image Quality Specifications for the SIR-C/X-SAR Mission", JPL, 1988.

4. Papoulis, A., "Probability, Random Variables, and Stochastic Processes", McGraw-Hill Book Company, 1965.

5. Gaskill, J.D., "Linear Systems, Fourier Transforms, and Optics", John Wiley & Sons, Inc., 1978.

6. Brigham, E.O., "FFT Schnelle Fourier-Transformation", R. Oldenburg Verlag, 1982.

FAST DIGITAL SIGNAL PROCESSING TECHNIQUES FOR SPACEBORNE SYNTHETIC APERTURE RADARS

R.Albrizio(*), G.Aloisio(#), C.Marzo(@), A.Mazzone(*)
G.Milillo(@), F.Pasquali(^), N.Veneziani(*)

(#) Dip. di Elettrotecnica ed Elettronica (Univ. Bari,Italy)
(*) Istituto Elaborazione Segnali ed Immagini (Bari,Italy)
(^) Centro di Geodesia Spaziale (Matera,Italy)
(@) Agenzia Spaziale Italiana (Italy)

ABSTRACT

A modular digital SAR processor, exploiting the range-doppler approach, has been developed and tested both on real and simulated SAR raw-data.
The processor includes modules for range migration correction and Doppler parameters estimation, as key points for high quality image generation. The output from the processor consists of a complex data matrix, at full geometrical resolution; when more radiometric quality is required, a multi-look technique allows to produce detected images at user defined resolutions.
The results of two experiments are reported in the paper: the first regards the "Doppler centroid" estimation based on a non standard use of the "Clutterlock Algorithm", directly applied on sensor's raw-data; the second on the improvement obtained in the processor throughtput, combining the performance of the Prime Factor Algorithm (PFA) with the "overlap & save" optimization technique for faster convolutions.
An example of image focalized by this processor on DEC-VAX and the related CPU time are reported in the paper, referring to SIR-B data processing. Potential applications of the processor are also considered, referring to the operative conditions of future spaceborne SAR missions.

Key-words: SAR-processor, fast convolution, Doppler parameters estimation.

1. INTRODUCTION

The growing use of spaceborne Synthetic Aperture Radars in remote sensing applications induced the development of several research activities, aimed to the study of interaction between ground targets and electromagnetic wave and to the improvement of the hw/sw technologies related to the whole SAR system. In this context, the studies on digital processing techniques for the SAR image focusing and on computers architectures suitable for this kind of processing are strictly related to the operative modes of the sensor, determining at the same time its evolution.

Owing to the different sensor's specifications and their different operative conditions, a flexible raw-data processing is required to fastly produce high quality images and several efforts are needed in this field, both to investigate new signal processing algorithms and to evaluate the efficiency of new computer architectures.
The Italian Space Agency (ASI), with the aim to develop autonomous know-how in this field, promoted the growth of a scientific cooperation between the IESI-CNR (Images & Signal Processing Inst.-Bari), the DEE (Electrotechnic & Electronic Dept.-University of Bari) and CGS (Center for Spatial Geodesy-Matera).
One of the outputs of this cooperation was a flexible digital SAR processor, allowing to easily manage the raw data produced by the future remote sensing space missions. The processor has been tested on SIR-B raw data provided by Jet Propulsion Laboratory (JPL-Pasadena) and on simulated data, [1].
In digital SAR-processing the main computational load occurs in the point target responce compression, for which several techniques can be used.
To have a precision processing and, at the same time, to match the high computational requirements, the range-doppler approach has been in our development preferred to the SPECAN algorithm, which does not allow to easily perform the range migration correction in the frequency domain, and to the "time domain" approach too, due to the high computational load involved.

2. SAR SENSOR'S CHARACTERISTICS

The Italian Space Agency is strongly involved in remote sensing activities jointly with ESA and DLR. Two missions are scheduled to fly in the next future, one by ERS-1 satellite and one by Space Shuttle. During these missions, an Active Microwave Instrumentation operating in C-band as a Synthetic Aperture Radar (SAR) also, and an X-band SAR will be respectively used.
Some specifications of these sensors and the related mission parameters are reported in Tab.1, compared with those of SIR B. A nominal altitude value and a single off-nadir angle have been considered, and the lengths of the azimuth reference function have been

	X-SAR	ERS-1	SIR-B
scheduled to fly on:	1992	1991	1984
nominal altitude [Km]	230	785	230
antenna size [m]	12.1*0.33	10*1	10.7*2.2
carrier freq. [MHz]	9375	5300	1274
pulse length [us]	40	37.1	30.4
" bandwidth [MHz]	9.5/19	15.55	12
sampl. freq. [MHz]	11.25/22.5	18.96	30.353
quant. [bits/sample]	4I,4Q/6I,6Q	5I,5Q	5
PRF [Hz]	1240-1820	1720	1539.8
attitude rate [deg/s]	0.03	0.0017	0.03
reference functions:			
range [samples]	451-902	705	922
azimuth ["]	119-175	1260	1221
at off-nadir [deg]	30	23	28

Tab.1 Sensor and mission parameters

	Y-SAR	ERS-1	SIR-B
look angle [deg]	30	23	28
Fdc [Hz]	12165	6408	954
Fdr [Hz/s]	-12974	-2041	-1789
Fdc [Hz]	124-182	172	47
Fdr [Hz/s]	108	2	0.45
azimuth update of			
Fdc [s]	0.25-0.37	10	3.28
Fdr [s]	27	/	3.50
MFD [RLs]	310-674	17200	5050
azimuth shift [m]	70-102	300	624
range update of			
Fdc [m]	1000-1400	4400	13600
Fdr [m]	2400	870	256
MSD [ALs]	75-108	110	25

Tab.2 Update rates (either across or along track direction)

evaluated, for X-SAR, according to the min/max values of PRFs.
Due to the attitudinal rates foreseen, changes are induced in the antenna pointing and the doppler parameters must be updated at a proper frequency, both in the along track and across track directions, in order to guarantee the designed image quality specifications.
The update rates for X-SAR and ERS-1, evaluated referring to a max error equal to 10% of PRF for doppler centroid and equal to 1/T**2 for doppler rate, are reported in Tab.2. The requirements used are tentative specifications for the X-SAR survey image.
At the operative conditions considered, a maximum frame dimension of "MFD" range lines, according to the PRF used, and a maximum subswath dimension of "MSD" azimuth lines can be processed using the same reference function. The efficiency in the azimuth processing is strictly dependent on these parameters and on the azimuth reference function length, so non standard DFT techniques for fast convolution have been explored and compared with the traditional FFT Cooley-Tuckey algorithm. As we will show, the Prime Factor Algorithm is most flexible with regard the sequence lengths allowable and faster than CTA on sequential machine, without vector processor.

3. THROUGHPUT IMPROVEMENT

Several fast Discrete Fourier Transform algorithms have been developed, to minimize the number of multiplications at the expense of memory accesses and control complexity.
To improve the throughput of our SAR processor, the Prime Factor Algorithm (PFA) has been used to efficiently compute the convolution operations. Like in the traditional Cooley-Tuckey Algoritm (CTA), the efficiency of PFA is obtained transforming the one-dimensional DFT into a multi-dimensional one, by proper mapping of the input and output data indices [2,3].
The efficiency ot PFA algorithm comes from joining the Rader results, to convert DFT into circular convolution in some specific conditions, with the Winograd technique for computing short convolutions in the minimum number of multiplications.
In CTA algorithm, the computation of one-dimensional transforms on N-length sequences, so that N=R**m, is performed in m-stages, each composed of N/R elementary modules (butterflies), involving arithmetic operations on a number of elements equal to the radix R (generally 2 or 4). Only one of these butterflies has to be programmed.
In PFA algorithm, used if the sequence lengths can be expressed as a product of mutually prime factors ($N=\prod_i Mi$), the DFT computation is made in a number of stages equal to the number "m" of prime factors. At each stage, N/Mi DFTs of order equal to the factor Mi considered can be computed independently. Different modules need to be programmed, one for each Mi factor, using the Rader and the Winograd algorithms.
In Tab.3, the number of operations required for CTA and PFA is shown, referring to nearly equal sequence lengths.
On sequential machines, without vector processor, PFA is about 30% faster than CTA, due to the time saved in less multiplications [4].

N	factors	PFA mult.	PFA add.	CTA mult.	CTA add.
252	4*9*7	1136	5952		
256	2**8			1800	5896
840	16*9*7	5704	29100		
1024	2**10			10248	30728
2048	2**11			23560	68616
2520	8*9*7*5	17660	82956		

Tab.3 Number of operations for PFA and Cooley-Tukey DFT algorithms

Since the azimuth line (AL) can be thought ideally unlimited, the "overlap & save" technique has been used more efficiently to compute the convolution operation with the azimuth reference function. So the AL can be divided in a number of overlapped sections, constituting overlapped frames.

The "overlap & save" optimization for PFA algorithm, in the case of infinite sequence length, has been derived and reported in Tab.4.

Q	PFA
20÷25	120
26÷33	144
34÷60	240
61÷74	360
75÷191	720
192÷303	1680
304÷376	2520
377÷1200	5040

Tab.4 Overlap & Save optimization for PFA (infinite sequence length)

For each reference function length (Q), the PFA optimized for the segmentation has been selected according to the minimum number of operations per useful data in output.
If a "multi-look" technique [5] is adopted for the azimuth compression, this table must be used taking into account that the overlap margin must be equal to the length of the azimuth reference function, divided by the number of independent "looks" that will be processed in the azimuth compression, minus one. It should be noted that, for a 4-looks processing, the optimization in the convolution computation is compatible with the updating frequency of the doppler parameters reported in Tab.2, also in the critical case of X-SAR sensor.
In the case of the range compression, the same optimization technique can be exploited, referring to the finite length of RLs.
The results for equal length sectioning and for the whole set of the RLs available for X-SAR, in the coarse resolution mode (1), are reported in [6]. The values have been evaluated dividing each RL length in a number of overlapped sections, as a function of the possible PFA lengths, then minimizing the global number of arithmetic operations per useful data in output.
For the fine resolution mode, similar tables must be evaluated considering the relative reference function length.
A further improvement can be achieved in several cases, if the last data segment is only partially filled, using for it the PFA length immediately smaller, when this is possible.

4. DIGITAL SAR PROCESSOR DESCRIPTION

The flow diagram of our digital SAR processor is shown in Fig.1.
For the range and azimuth compressions, the "matching filter" technique in the frequency domain was adopted.
To compute the Discrete Fourier Transform, the Prime Factor Algorithm, recently introduced instead of the usual Cooley-Tuckey FFT algorithm, and the "overlap & save" optimization technique allowed to save about 40% in the processing time.
In the following, the principal characteristics of each software module developed are described.

(1) In the coarse resolution mode the X-SAR operates with the chirp band of 9.5 MHz and the sampling frequency of 11.25 MHz.

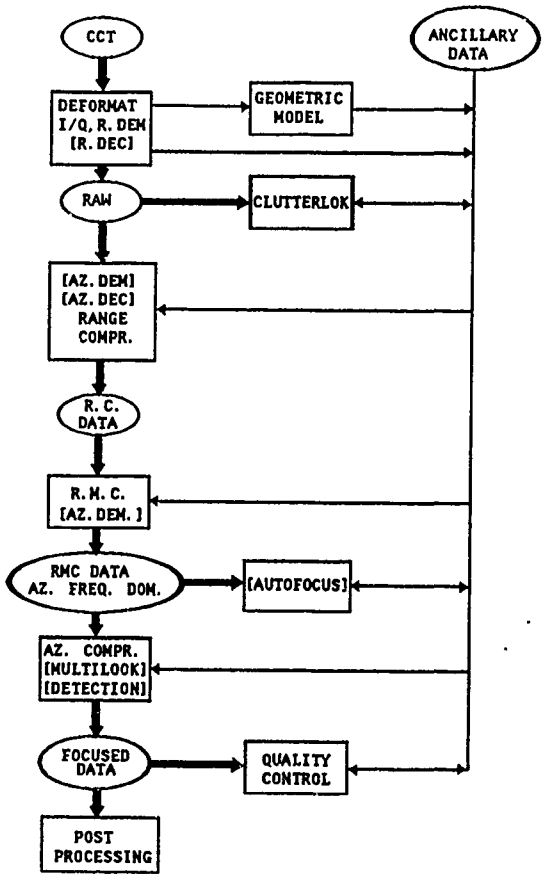


Fig.1 SAR processor block diagram

4.1 Raw-Data Deformatting

In order to avoid the corner turning operation at different phases of processing, i.e. each time the data matrix access is changed from range to azimuth directions and viceversa, the raw data files, obtained from CCT format, are organized so that each record contains a subframe of original data.
When the input data are real, the conversion to a complex signal (analytic transform) and the resampling can be performed. Moreover, the demodulation to base band of the complex signal, along the range direction, is achieved in this phase by a frequency shift performed in the time domain. At this point a decimation can be also applied.

4.2 Geometrical Model

The sensor and mission parameters obtained from the header are introduced in an accurate geometrical model to determine the nominal values of all parameters required for the processing.

4.3 Doppler Centroid Estimation

The Doppler Centroids are determined by the "Clutterlock procedure", as a cross-correlation between the azimuth line and the theoretic pattern of antenna, [7,8]. It should be noted that, in our approach, this procedure has been tested not only on the range compressed data, but also directly on the raw-data since the backscattered energy is better distributed on the antenna pattern before any type of data focusing. The results obtained in these two cases are shown in Fig.2 and 3, where the Fdc behaviour, as evaluated for each azimuth line, is reported. The fit to the experimental values gives the Fdc linear behaviour between near and far range.
Due to the lower noise level found in the Clutterlock output, when applied directly on the raw-data, the first approach was extensively adopted in our SAR processor

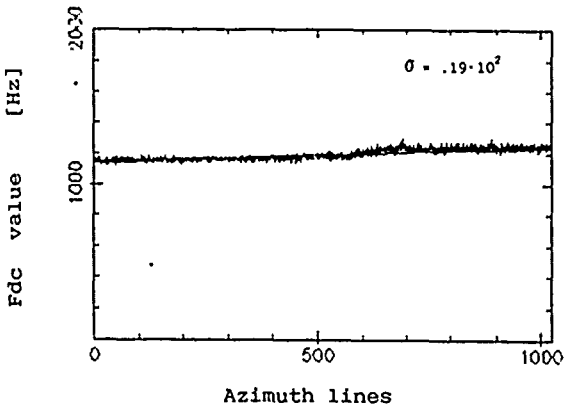


Fig.2 Result of Clutterlock applied on raw-data: the standard deviation of difference between the fitted and experimental values is also reported

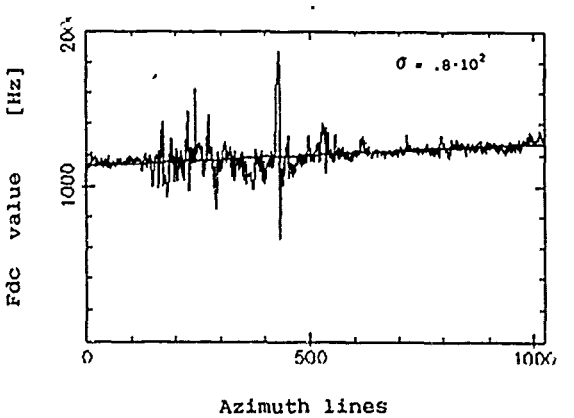


Fig.3 Result of Clutterlock applied after range compression: the standard deviation of difference between the fitted and experimental values is also reported

4.4 Azimuth Demodulation and Decimation

Optionally, also in the azimuth direction a demodulation to base band, followed by decimation, can be performed to reduce the data set dimension to be processed.

4.5 Range Compression

The reference function (matching filter) is properly generated by the informations deduced from the raw-data's header.
A complex chirp is used and the sidelobe reduction is obtained by a symmetric weighting, using a raised (0.3) cosine squared function.

4.6 Autofocus Procedure

Substantially the algorithm is based on a cross-correlation between two "partial looks". These are obtained compressing a properly selected azimuth line with two partial reference functions, [7].
The algorithm's basic idea is the following: when the right value of Fdr is used, only a relative time shift is generated between the outputs of the partial looks, so that, if these are now cross-correlated, the correlation peak will appear in the position of the first sequence's sample. On the contrary if an improper Fdr value is used, an additional displacement between the partial looks will be produced so, the correlation peak will appear displaced with respect to the first sample.
The optimal rate is searched by an iterative procedure starting from the nominal Doppler parameters.

4.7 Range Migration Correction (RMC)

The RMC operation is made by a proper shift in the range direction for each frequency component of the azimuth lines. The shifts are determined, for each frequency, depending on its position in the spectrum and on the behaviour of the radar-target distance along the swath.. The RMC algorithm is strictly dependent on the accuracy with which the Doppler parameters are determined. As known, the range migration is composed by a linear component (Range Walk), depending on the Doppler Centroid, and by a quadratic component (Range Curvature), depending on the Doppler Rate.
In order to correct at the same time these two effects, the range migration correction algorithm has been performed in the frequency domain [9], as it is faster than the time domain algorithm.
In fact, in the frequency domain, the correction is made simultaneously for all the points belonging to the same azimuth line. An high quality image is achieved by the cubic splines data interpolation in the range direction.
At this point, if the signal is not in base band, a demodulation in the frequency domain is performed. This operation, made at this stage of the processing, simplifies the addressing of the samples to be shifted.

4.8 Azimuth Compression

The compression in the azimuth direction can be performed updating, at each azimuth line, the Doppler parameters used in the reference chirp. The window used is a raised (0.3) cosine squared function.
Independent images of the same scene (looks) are produced by processing subapertures of the whole synthetic antenna. Then the multilook-image is produced by incoherently adding different views of the scene so obtained [10].
In this way, the geometrical resolution is lower but an higher radiometric resolution can be achieved, due to the speckle reduction.
A full-resolution complex image (bulk image) can be also produced for non standard applications.
An example of SIR-B image, focused by our SAR processor is shown in Fig.4.

Fig.4 SIR-B image of Mt.Shasta (CA)

4.9 Quality Measurements

To evaluate the image quality, suitable algorithms characterizing both point and extended targets have been developed. They allow to measure geometrical resolutions ISLR (Integrated Side Lobe Ratio), PSLR (Peak Side Lobe Ratio) on the point-targets and the radiometric resolution (Rrad) on the distributed one.
The results obtained both on full resolution and multilook images are reported in Tab.5. Due to the absence of very shining point targets in the SIR-B image considered, the geometrical resolutions, in slant range (Rsr) and azimuth (Raz), the PSLRs and the two dimensional ISRL have been calculated on SIR-B simulated data. Considering the weighting used for the sidelobe reduction, the results obtained are in agreement with the expected values [5,10].

N.of look	Rsr [m]	Raz [m]	PSLR [dB] Ra/Az	ISLR [dB]	Rrad [dB]
N=1	14.94	6.37	27.6/28.1	15.7	2.98
N=4	14.98	25.57	27.4/27.7	15.5	1.87

Tab.5 Results of quality measuremnets.

5. CONCLUSIONS

A modular digital SAR processor, allowing to focus raw-data coming from different sensors, has been developed.
The processor was tested on SIR-B raw data and its suitability for other sensors is currently evaluated, taking into account the parameters of Tab.2 and optimal sectioning rules.
To test the image quality, a point target responce simulator has been developed. The use of this simulator and of a package for the quality measurements allowed to test the functionality of the processor.
An high throughput has been also achieved by using the PFA algorithm to efficiently compute the convolution operations and by means a proper data structure.
To focus a 5040*1260 complex data matrix, on a VAX 11/780, about 106 min of CPU time are needed. A version of the processor running on a microVAX 3600 is also available, taking about 1/3 of this time to focus the same data set, with I/O time less than 10% of the CPU time.
Our next goal is to further increase the processor efficiency, by its parallel implementation on a multiprocessor machine.

6. REFERENCES

[1] Marzo,C., Milillo,G., "Simulazione di dati prodotti da un sistema SAR, approccio sistemistico" (italian), PSN/CNR, CGS-Matera (Italy), 1988.

[2] Nussbaumer,H.J., "Fast Fourier Transform and Convolution Algorithms", Springer, 1980.

[3] Burrus,C.S., Parks,T.W., "DFT/FFT and Convolution Algorithms", J.Wiley & Sons, 1985.

[4] Mehalic,M.A., Rustan,P.L., Route,G.P., "Effects of Architecture Implementation on DFT Algorithm Performance", IEEE on ASSP, vol.33, Jun 1985.

[5] Li,F.K., Held,D.N., Croft.,G, "Comparison of several techniques to obtain multiple-look SAR imagery", IEEE Trans. on Geo. and Remote Sensing, vol.21, J-21, n.3, 1983.

[6] Aloisio,G., Fox,G.C., Milillo,G., Veneziani,N., "X-SAR Digital Processor: computational load for real-time", JPL/Caltech, C3P-740, 1989.

[7] Li,F.K., Held,D.N., Curlander,J.C., Wu,C. "Doppler parameter estimation for spaceborne synthetic aperture radars", IEEE Trans.on Geo. and Remote Sensing, vol.GE-23, n.1, 1985.

[8] Ulaby,F.T., Fung,R.K., "Microwave Remote Sensing", vol. II, Addison Wesley, Reading (Mass.), 1981.

[9] Barber,B.C., "Theory of digital imaging from orbital syintetic-aperture radar" Int. J. of Remote Sensing, vol.6, n.7, 1985.

[10] Miller,P.F., "Measurements of radiometric criteria", ESA SP-172, Frascati (I), 1980.

ERROR SENSITIVITIES OF A SECONDARY RANGE COMPRESSION ALGORITHM FOR PROCESSING SQUINTED SATELLITE SAR DATA

F.H. Wong/I.G. Cumming
MacDonald Dettwiler
13800 Commerce Parkway
Richmond, B.C., Canada V6V 2J3
Tel: (604)278-3411 - Telex: 04-355599 - Fax: (604)278-0531

ABSTRACT

A commonly employed SAR processing algorithm is the range/Doppler in which frequency domain range correlation is followed by frequency domain azimuth correlation, and range cell migration correction is performed in the range time/azimuth frequency or range/Doppler domain. When data is obtained with the radar antenna squinted away from zero Doppler, this algorithm tends to broaden the impulse response of a compressed point target in the range direction. The broadening is caused by a reduced time bandwidth product in the azimuth FFT of a point target in a range cell. A secondary range compression (SRC) algorithm, first developed by Jin and Wu at JPL, can be used to correct for the range broadening.

This paper analyzes the factors affecting the accuracy of the algorithm. Experiments were performed on simulated C-band and L-band satellite SAR data by assuming a squint angle of about 6°. The SRC algorithm successfully restored the resolution to an acceptable level.

Keywords: SAR data processing, squint processing, Doppler centroid, impulse response width.

1. INTRODUCTION

This paper investigates an SRC algorithm to process squinted SAR data. The algorithm was first developed by Jin and Wu [1], and extended by Schmidt [2] who investigated methods of implementation and multi-look processing.

A range Doppler algorithm [3] is often chosen in a SAR processor because of its ability to compress a point target accurately, and because it can be implemented as a two one-dimensional matched filtering process. This allows efficient computer implementation (including SRC), and provides adequate accuracy control over key operations such as range cell migration correction (RCMC). The typical processing steps are range compression, azimuth FFT, range cell migration correction and finally azimuth compression.

The range/Doppler algorithm introduces range impulse response broadening in processing squinted data. The broadening is due to the fact that a target traverses a range cell in a shorter time for a higher squint, resulting in a smaller time bandwidth product (TBP) in that cell. The principal of stationary phase which is assumed to be valid in performing the azimuth FFT, does not hold when the TBP is too small, typically less than 1. Range broadening occurs in the azimuth FFT as a result of violating the principal of stationary phase.

For a given squint angle, the broadening is more severe for a longer target exposure time (such as L-band). The SRC algorithm is to designed to correct for this broadening. In this paper, the algorithm is tested for squint angles of up to 6°, and it is found that the algorithm's ability to recover the lost resolution is acceptable. The algorithm is now incorporated in the MacDonald Dettwiler Generalized SAR processor.

The purpose of this paper is to analyze the algorithm's sensitivity with respect to various parameters, perform experiments on simulated C-band and L-band data and evaluate the results.

2. THEORETICAL BACKGROUND

The SRC is a range filtering operation and can be incorporated in the range compression step (range domain) or range cell migration correction step (Doppler domain). While the Doppler domain SRC provides more control over errors, the range domain SRC is considerably more efficient and is preferred. For this reason, error sensitivities of the range domain SRC are investigated here.

In the range domain SRC algorithm, the SRC filter can be cascaded into the range compression matched filter. Because the SRC filter, similar to the matched filter, is also linear frequency modulation (FM) encoded, the cascading of these two filters can be effected by a combined filter with FM rate given by [2]:

$$K_c = \frac{K_r K_{src}}{K_r + K_{src}} \tag{1}$$

where K_r is the range chirp FM rate and K_{src} is the FM rate of the SRC filter, and K_{src} is given by:

$$K_{src} = \frac{-2 V^2}{\lambda R} \left(\frac{\lambda \eta_c}{c} \right)^2 \left(1 - \frac{V^2 \eta_c^2}{R^2} \right) \tag{2}$$

where λ is the radar wavelength, R the target slant range, V^2 is the square of the relative velocity between the target and the radar, c the speed of light, and η_c the azimuth slow time of Doppler centroid offset from zero Doppler. Usually, $K_{src} >> K_r$ so that K_c involves a relatively small FM rate adjustment to the range matched filter. Typically the value of K_{src} is at least 100 times that of K_r for C-band and 30 times for L-band. For a zero squint angle ($\eta_c = 0$) K_c is the same as K_r as anticipated.

Examples of the C-band and L-band parameters for radar satellites are given in Table 1, taken from ERS-1 and SEASAT respectively. In the table, T_r is the pulse duration, T_a is the maximum processing aperture time and R_{sw} is the slant range swath width. These parameters will be used in the error analysis below.

TABLE 1 C-BAND AND L-BAND PARAMETERS

Parameters	C-band	C-band
λ, m	0.0567	0.235
K_r, GHz/sec	417.8	562.0
T_r, μsec	37.1	33.9
T_a, sec	0.7	2.0
v^2, Km^2/sec^2	50.0	50.0
R_o, Km	854	877
R_{sw}, Km	40	40

3. RANGE IMPULSE RESPONSE WIDTH BROADENING

An error in K_c (Equation 1) will tend to broaden the range impulse response width (IRW). An experiment was performed to measure the broadening as a function of the error in K_c. The broadening was measured and was observed to have the same form as a quadratic phase error in the matched filtering of a linear FM signal.

Graph (a) in Figure 1 shows the well known form of the IRW broadening as a function of the quadratic phase error $|\Delta\phi|$ at the ends of the pulse duration $\pm T_r/2$. The phase error is related to the K_c error ΔK_c by:

$$\Delta\phi = \pi \Delta K_c (T_r/2)^2 \qquad (4)$$

For example, the graph shows that for a broadening of less than 5%, $|\Delta\phi|$ has to be less than 0.4π rad. The graph, together with Equation 4, can be used to deduce the range IRW broadening if the SRC filter is not tuned properly.

Another cause of broadening is that the SRC filter is kept constant along a target trajectory. Therefore, even though $\Delta\phi$ is zero at the Doppler centroid, it is non-zero at other points along the trajectory. Hence, after azimuth compression, an overall range IRW broadening occurs. This cause of broadening is discussed further in the next section.

4. SENSITIVITY ANALYSIS

Two types of errors exist in applying the SRC algorithm; namely, by holding the SRC filter constant along a target trajectory, and by holding it constant within a filter range invariance region. These errors do exist even in the absence of beam squint, and they increase with squint. They are discussed separately below.

4.1 Constant SRC Filter in Azimuth Time

Ideally the SRC filter should vary along the target trajectory since the azimuth time η changes along the trajectory. This can only be done in the Doppler domain where the algorithm can be adjusted for each Doppler value or look. As have mentioned, this approach is considerably less efficient and is not considered here.

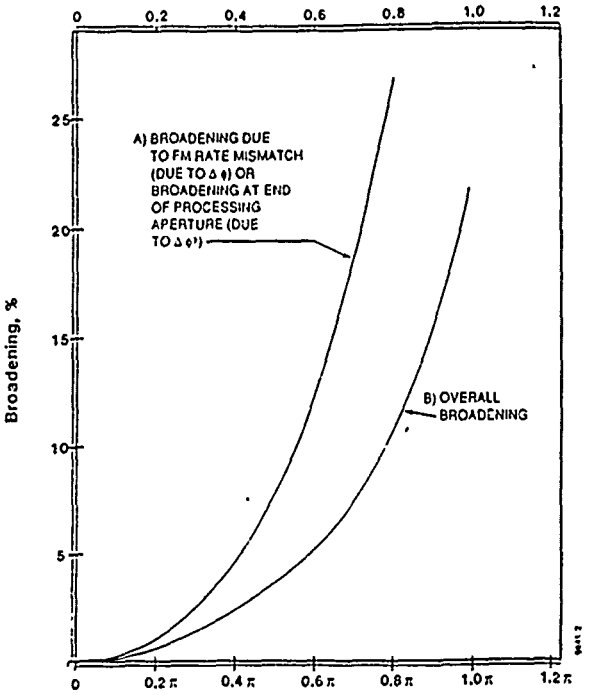


FIGURE 1 IRW BROADENING VERSUS QUADRATIC PHASE ERROR

In Equation 2, the SRC filter is evaluated only at the Doppler centroid and this approximation prevents the SRC filter from recovering the full range resolution. An interpretation is that prior to azimuth compression, no broadening of range impulse response occurs (provided that the SRC filter is correctly computed) at the trajectory centre and the broadening gradually fans out to either end of the processing aperture. The broadening at a specific trajectory point is due to a quadratic phase error. From Equations 1,2 and 4, it can be shown that this error, $\Delta\phi'$ (the prime notation is used here to distinguish it from a similar error defined in the next subsection), is given by:

$$\Delta\phi' = \pi \left[\frac{K_r T_r}{2 (K_r + K_{src})} \right]^2 \left\{ \frac{\lambda R}{v^2 \eta_c} \left[\frac{c}{\lambda \eta_c} \right]^2 \Delta\eta_c' \right\} \qquad (5)$$

where $\Delta\eta_c'$ is the azimuth time difference between the point and its Doppler centroid η_c. In the above equation it is assumed that for the squint angle considered, $R^2 >> v^2\eta_c^2$ is still valid. The effect of azimuth compression is to integrate coherently the range impulse responses over the processing aperture, thus introducing an overall range IRW broadening in the compressed point target.

The overall broadening was simulated by assuming the range impulse response to be a sinc function with an azimuth time dependent broadening factor governed by the quadratic phase error, and a $sinc^2$ antenna pattern. The processing aperture was taken to be within the ±3 dB point of the antenna pattern. Also, the SRC filter was computed correctly at the trajectory centre. Graph (a) in Figure 1 shows the broadening at the end of the processing aperture, and this has the same form as FM rate error in matched filtering. Graph (b) in Figure 1 shows the overall broadening as a function of the quadratic phase error at either end of the processing aperture, after azimuth compression has integrated the effects of the varying amount of broadening along the target trajectory.

For example, using the L-band parameters and a Doppler centroid of 14 sec, $\Delta\phi'$ is 0.8π rad ($\Delta\eta_c'$=1 sec in Equation 5 since T_a=2 sec), giving a range broadening of 11% for the zero squint case. Of course, this overall broadening depends upon the duration of the processing aperture time T_a. A similar analysis with the C-band parameters shows that the overall broadening is negligible.

For a zero squint, $\Delta\phi'$ is 0.06π rad for L-band and 0.003π rad for C-band. In each case the overall broadening is negligible.

4.2 Constant SRC Filter in Invariance Region

Equations 1 and 2 show that the combined range matched filter FM rate K_c is a function of slant range R, relative velocity squared V^2, Doppler centroid η_c and range chirp FM rate K_r. An error in each, except K_r, is introduced by the use of an SRC filter invariance region. In this region the SRC filter is kept constant within this region for computational efficiency purposes, whereas R, V^2 and η_c all vary with range.

Again, from Equations 1,2 and 4, it can be shown that:

$$\Delta\phi = \pi \left[\frac{K_r T_r}{2 (K_r + K_{src})} \right]^2 \left\{ \frac{-\lambda}{2 V^2} \left[\frac{c}{\lambda \eta_c} \right]^2 \Delta R + \frac{\lambda R}{2 (V^2)^2} \left[\frac{c}{\lambda \eta_c} \right]^2 \Delta V^2 + \frac{\lambda R}{V^2 \eta_c} \left[\frac{c}{\lambda \eta_c} \right]^2 \Delta \eta_c \right\} \qquad (6)$$

This is the key equation to determine the expected range IRW broadening, as a function of $\Delta\phi$, due to squint for given errors in R, V^2 and η_c.

Table 2 shows the relative contributions of these errors to $\Delta\phi$ for the given C-band and L-band parameters, and at a squint angle of 6°. In the table, the value of K_{src} is assumed to be constant across the entire slant range swath width of 20 km, so that ΔR is 20 km. The value of V^2 is assumed to vary by at most 0.4% from near range to far range, so that ΔV^2 is 0.2% of V^2. This is increased to 0.5% by including orbit data errors in computing the value of V^2. Finally, for a squint angle of 6°, $\Delta\eta_c$ is 0.3 sec at the swath edge. The table shows that the contribution to $\Delta\phi$ by ΔV^2 is relatively insignificant, and that by $\Delta\eta_c$ is most significant of the three. However, the broadening caused by these errors for the L-band case is small compared to that due to $\Delta\phi'$ (discussed in the last subsection). Also, the broadening for the C-band case is insignificant.

TABLE 2 CONTRIBUTIONS TO $\Delta\phi$ DUE TO SRC FILTER PARAMETER ERRORS AT 6° SQUINT

Contribution	C-Band	L-Band
η_c, sec	17	14
$\Delta\phi$ (ΔR=20 km), rad	0.03 π	0.12 π
$\Delta\phi$ (ΔV^2=0.25 km^2/sec^2), rad	0.01 π	0.05 π
$\Delta\phi$ ($\Delta\eta_c$=0.3 sec), rad	0.05 π	0.23 π
Total $\Delta\phi$, rad	0.09 π	0.40 π

5. SIMULATION OF SRC ALGORITHM

Experiments were performed to verify the performance of the SRC algorithm and the effects of perturbations in the SRC filter parameters. Both C-band and L-band parameters shown in Table 1 were used. Four SAR data sets were simulated: a zero squint for C- band, a zero squint for L- band, a 17 sec squint for C-band and 14 sec for L-band (about 6° squint in each case).

The following experiments were performed to examine each source of error separately:

- The zero squint data sets were processed without SRC and the correct values of V^2 and R were used. The azimuth and range IRWs were measured. The range broadening caused by keeping the filter constant throughout a target trajectory has shown to be negligible in both C-band and L-band. These IRWs were used as references for computing broadening in later trials.
- The squinted data sets were processed without SRC and then with SRC and the IRW broadenings in range and azimuth were measured.
- The squinted data sets were processed with the slant range offset by half the swath width in computing the SRC filter. This was to evaluate the effect of error in slant range.
- The squinted data sets were processed with the Doppler centroid offset by 0.3 sec (error of η_c at swath edge) in computing the SRC filter. This was to evaluate the effect of a Doppler centroid error.

In each experiment, both 1-look and 4-look processing were performed. For C-band, the total processed bandwidth was about 2 sec for 4-look and 1 sec for 1-look processing. For L-band, it was about 0.7 sec and 0.35 sec for 4-look and 1-look respectively. Thus, the processed aperture time for 4 looks was about twice as long as that for 1 look.

Errors in ΔV^2 were not simulated here since it was determined that they would not contribute significantly to the range IRW broadening.

Tables 3 and 4 summarize the range IRW broadening for the various experiments.

Table 3 shows that, with SRC, the range IRW broadening for 4-look L-band data was 11%. This was a result of using a correct SRC FM rate only at the Doppler centroid of a target trajectory but not elsewhere along the trajectory, and this result agreed with the prediction discussed earlier. The broadening was worse with 4-look processing than with 1-look due to the difference in their total processing apertures. This table also shows that the broadening for C-band was insignificant. Overall, the residual broadening was worse for L-band than C-band simply because of the longer target exposure time in the L-band case. Figure 2 is a plot of the range impulse responses with and without using the SRC filter for the L-band 1-look case with 6° squint.

Table 4 shows the broadening due to an error in R (ΔR=20km) acting alone. The broadening was a little more than that at the swath centre (in Table 3). This experiment therefore indicates that, for the value of ΔR considered, the range broadening was relatively small compared to that due to keeping the SRC filter constant along the target trajectory ($\Delta\phi'$ effect).

Table 4 also shows that the broadening due to an η_c error acting alone. Again, the bruadening in the C-band case was insignificant. For the L-band 4-look case, the broadening was about 13%, which was a combined effect of both the η_c error and $\Delta\phi'$ at the end of the processing aperture. Again, comparing to Table 3, the latter was the dominant effect; however, the former could not be ignored.

When ΔR and $\Delta\eta_c$ were combined at the swath edge, the L-band 4-look broadening increased to 15%.

Other parameters such as azimuth IRW broadening, peak power, maximum sidelobe power, integrated sidelobe ratio were also measured. With SRC processing, these parameters did not change significantly when compared to the zero squint case.

TABLE 3 RANGE IRW BROADENING FOR 6° SQUINT (AT SWATH CENTRE), WITH AND WITHOUT SRC

Band	Number of Looks	Processed Aperture Time	Without SRC	With SRC
C	1	0.35 sec	180.5%	1.3%
C	4	0.7 sec	174.9%	1.3%
L	1	1.0 sec	988.7%	4.4%
L	4	2.0 sec	1201.8%	10.9%

TABLE 4 RANGE IRW BROADENING FOR 6° SQUINT (AT SWATH EDGE), $\Delta R/\Delta\eta_c$ ERRORS

Band	Number of Looks	Processed Aperture Time	Broadening ΔR=20 km	Broadening $\Delta\eta_c$=0.3 sec
C	1	0.35 sec	1.6%	1.7%
C	4	0.7 sec	1.7%	1.7%
L	1	1.0 sec	6.7%	9.7%
L	4	2.0 sec	11.4%	12.9%

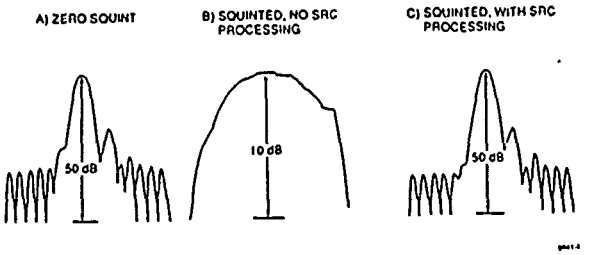


FIGURE 2 RANGE IMPULSE RESPONSE FOR L-BAND 1-LOOK PROCESSING

6. CONCLUSIONS

Performance of the SRC algorithm has been analyzed theoretically and experimentally for C-band and L-band. The analysis has included perturbations in parameters such as range, relative velocity and the effect of holding the SRC filter constant throughout the target trajectory ($\Delta\phi$' effect). The algorithm should be employed to prevent excessive range IRW broadening, especially in the case of long target exposure time.

For C-band data, it is found that the SRC algorithm provides accurate compression of point targets of up to at least 6°. Also for the amount of squint considered, the SRC filter can be kept constant over the entire swath width when processing blocks of up to 40 km in slant range.

For L-band data, for the same squint angle and same slant range swath width, the worst range broadening occurs at the swath edge where the dominant factors are the $\Delta\phi$' and $\Delta\eta_c$ effects in this order. However the results are still acceptable when compared to the case of no SRC filter applied. One way to reduce the broadening further in L-band data is to reduce the SRC filter invariance region size. Another way is to apply the SRC algorithm in the azimuth frequency domain and vary the SRC filter as a function of Doppler centroid frequency.

ACKNOWLEDGEMENT

The authors to thank J. Brace, of MacDonald Dettwiler, for performing the experiments. The authors are equally grateful to W. Chong, also of MacDonald Dettwiler, for his helpful advice on the MacDonald Dettwiler Generalized SAR processor.

REFERENCES

[1] Jin, M.Y. and Wu, C. A SAR Correlation Algorithm which Accommodates Large-Range Migration. IEEE Transactions on Geoscience and Remote Sensing, Vol. GE-22, No. 6, November 1984.

[2] Schmidt, A.R. Secondary Range Compression for Improved Range/Doppler Processing of SAR Data with High Squint. M.Sc. Thesis, Department of Electrical Engineering, U.B.C, September 1986.

[3] Cumming I.G. and J.R. Bennett. Digital Processing of SEASAT Synthetic Aperture Radar Data. IEEE International Conference on Acoustics, Speech and Signal Processing, Washington, D.C., April 1979.

A Phase Preserving SAR Processor

R. Keith Raney

Canada Centre for Remote Sensing
and RADARSAT
2464 Sheffield Road
Ottawa, Ontario K1A 0Y7

Paris W. Vachon

Data Acquisition Division
Canada Centre for Remote Sensing
2464 Sheffield Road
Ottawa, Ontario K1A 0Y7

Abstract

Synthetic aperture radar (SAR) image phase information is necessary to support many advanced SAR applications. The phase information in the complex image for conventional range-Doppler processors is not a robust estimate of scene phase. A SAR processor specifically designed to preserve phase information is being developed at the Canada Centre for Remote Sensing (CCRS). In addition to preserving vital phase information, this processor can support large degrees of range curvature and range migration. Therefore, it is possible, in principle, to use this processor for satellite SAR data, high resolution airborne SAR data, and for both squint mode and spotlight mode SAR data. This paper summarizes the theory and presents early results.

Key Words: SAR, processing, phase

1 Introduction

Phase information is needed to support advanced SAR applications such as interferometry, coherent quadrature polarimetry, precision point scatterer calibration, and extraction of ocean wave information. For such applications, conventional range-Doppler SAR processors, designed simply to satisfy image resolution and magnitude requirements, achieve their goals but do not preserve correct phase information.

We are developing a new phase preserving SAR processor. The innovative portion of our approach is that the bulk of the range walk correction, range curvature correction, and azimuthal compression is performed in the two-dimensional frequency (wavenumber) domain, rather than in the range-Doppler domain. The key to this procedure working correctly is that for most radar systems, a small adjustment in the range phase achieves the necessary range curvature correction without significantly disturbing the range or azimuth focus.

This phase preserving SAR processor methodology was inspired by the SAR processor developed by F. Rocca (Rocca, 1987, Caffario *et al.*, 1988). Rocca's processor is centred around the Stolt change of variables (Stolt, 1978) which is well-known in downward continuation seismic signal processing. However, the phase preserving SAR processor at CCRS invokes SAR-specific considerations to improve the processing efficiency and ease of implementation.

2 Processing Methodology

We begin the development of this phase preserving SAR processor by considering the received signal from a single point scatterer at range R. For platform centred coordinates (x, y), along-track and slant-range respectively, the scatterer trajectory is illustrated in Fig. 1a. For radar memory coordinates $(x, \hat{y})$, the received signal trajectory is illustrated in Fig. 1b. Assuming an antenna squinted by angle α (positive for forward squint), the received signal data s in the $(\zeta, \hat{y})$ domain (Fig. 1b) may be written

$$\begin{aligned} s(\zeta, \hat{y}) &= p[\hat{y} - R(\zeta - X_0)]\, w(\zeta - X_0) \\ &\quad \cdot \exp\{-j\, 2kR(\zeta - X_0)\}\ , \end{aligned} \tag{1}$$

where $p(y)$ is the transmitted range pulse waveform, $w(x)$ is the azimuth antenna weighting function, k is the radar wavenumber, $R(x) = \sqrt{R^2 + x^2}$ is the scatterer range, and $X_0 = R \tan\alpha$ is a reference azimuth location, as illustrated in Fig. 1b. Note that the amplitude of $s(\zeta, \hat{y})$ is scaled by the radar cross-section of the scatterer in question (tempered by radar equation considerations) and that superposition applies for the case of multiple scatterers. In equation (1), we have retained the full hyperbolic range dependence. Hence, the expression accurately represents range curvature and squint-mode SAR. Note the explicit coupling between the ζ and $\hat{y}$ coordinate systems.

We use six processing steps to achieve a complex SAR "image" with correctly preserved phase. A block diagram illustrating the processing steps is given in Fig. 2.

2.1 Two-Dimensional Fourier Transform

We first transform into the two-dimensional spatial wavenumber domain $(\omega_\zeta, \omega_{\hat{y}})$ via two sets of one-dimensional Fourier transforms. The range transform may be explicitly calculated, and the azimuth transform benefits from application of the principle of stationary phase. The result is

$$\tilde{s}(\omega_\zeta, \omega_{\hat{y}}) \approx C_1\, \tilde{p}(\omega_{\hat{y}})\, w(\zeta^* - X_0)\, \exp\{j\, \phi(\zeta^*)\}\ , \tag{2}$$

where C_1 is a complex constant and $\phi(\zeta) = -(2k + \omega_{\hat{y}})R(\zeta - X_0) - \omega_\zeta \zeta$. The value of the stationary point ζ^* is found from $d\phi(\zeta^*)/dx = 0$. We calculate

$$\zeta^* = -\frac{\omega_\zeta R}{\sqrt{(2k + \omega_{\hat{y}})^2 - \omega_\zeta^2}} + X_0\ . \tag{3}$$

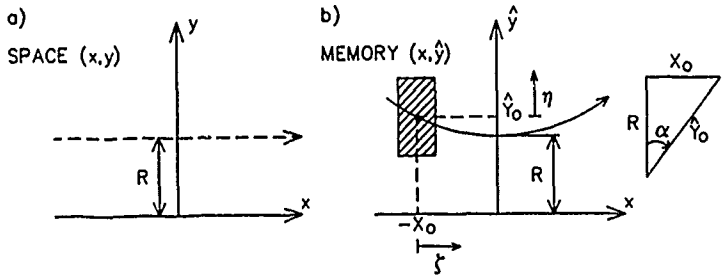


Figure 1: a) Scatterer trajectory through space and b) through radar memory. R is the slant range of closest approach, $\hat{Y}_0$ is the reference range, X_0 is the corresponding reference azimuth offset, and α is the antenna squint angle.

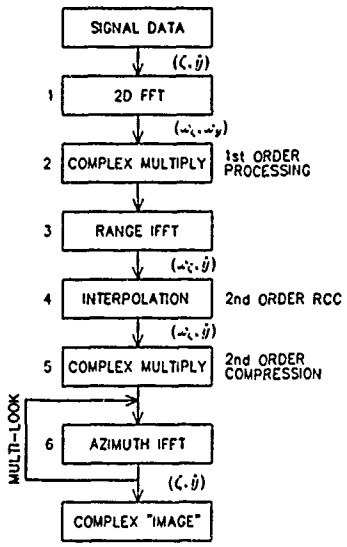


Figure 2: Flow chart of principal processor steps.

Substituting, we find that

$$\tilde{s}(\omega_\zeta,\omega_{\hat{y}}) \approx C_1\,\tilde{p}(\omega_{\hat{y}})\,w\left(-\frac{\omega_\zeta R}{\sqrt{(2k+\omega_{\hat{y}})^2-\omega_\zeta^2}}\right) \cdot \exp\left\{-j\left[R\sqrt{(2k+\omega_{\hat{y}})^2-\omega_\zeta^2}+\omega_\zeta X_0\right]\right\}. \tag{4}$$

The result of equation (4) may be directly used as the basis for a SAR processor of this class. However, it is advantageous at this stage to introduce three SAR-specific simplifications. First, we expand the terms under the $\surd$ about the centre Doppler frequency

$$\Omega = 2k\sin\alpha\,. \tag{5}$$

We define the azimuth Doppler frequencies as

$$\omega_\zeta = \Omega + \omega_b\,, \tag{6}$$

where ω_b represents frequency excursions about the centre Doppler frequency. Such expansion about an offset frequency is sensible for SAR systems that have azimuthal beamwidths much narrower than analogous widths found in other signal processing fields, such as seismics (for example, Gazdag and Sguazzero, 1984) and tomography (for example, DiCenzo, 1986). The approximation that results is excellent over typical beamwidths of a few degrees, and for large offset squint angles, as noted below.

Second, we note that for a typical SAR, the resolution ρ (of order metres) is much larger than the radar wavelength λ (of order centimetres). Thus, we may use, as a good approximation,

$$k = \frac{2\pi}{\lambda} \gg \omega_b, \omega_{\hat{y}}\,. \tag{7}$$

Third, we restrict the domain of the squint angle on the basis of equation (7) such that

$$|\alpha| \ll \tan^{-1}\sqrt{\frac{\rho}{\lambda}}\,. \tag{8}$$

For a typical airborne SAR ($\rho = 1$ m, $\lambda = 0.056$ m), the squint angle domain is $|\alpha| \ll 77°$, which is a rather generous restraint.

Assume now that we wish to process a portion of signal data such as that in the shaded region of Fig. 1b. This region is centred at $(-X_0, \hat{Y}_0)$, where in general $\hat{Y}_0$ is known and X_0 may be calculated from geometry via

$$X_0 = \hat{Y}_0 \sin\alpha\,. \tag{9}$$

We may substitute for the nominal scatterer range parameter

$$R = (\hat{Y}_0 + \eta)\cos\alpha\,, \tag{10}$$

which references our memory coordinate system to the non-skewed spatial coordinate system. Using the above definitions we find

$$\begin{aligned}\tilde{s}(\omega_\zeta,\omega_{\hat{y}}) \approx\; & C_1\,\tilde{p}(\omega_{\hat{y}})\,w\left(-\frac{\omega_\zeta R}{2k}\right) \\ & \cdot\exp\left\{-j\,2k\,(\hat{Y}_0+\eta\cos^2\alpha)\right\} \\ & \cdot\exp\left\{-j\,\omega_{\hat{y}}\hat{Y}_0\left[1+\frac{\tan\alpha}{\cos\alpha}\frac{\omega_b}{2k}+\frac{(1+3\tan^2\alpha)}{\cos^2\alpha}\frac{\omega_b^2}{8k^2}\right]\right\} \\ & \cdot\exp\left\{j\,\frac{\hat{Y}_0}{\cos^2\alpha}\frac{\omega_b^2}{4k}\right\} \\ & \cdot\exp\left\{-j\,\omega_{\hat{y}}\eta\left[1+\frac{\tan\alpha}{\cos\alpha}\frac{\omega_b}{2k}+\frac{(1+3\tan^2\alpha)}{\cos^2\alpha}\frac{\omega_b^2}{8k^2}\right]\right\} \\ & \cdot\exp\left\{j\,\eta\left(\sin\alpha\,\omega_b+\frac{1}{\cos^2\alpha}\frac{\omega_b^2}{4k}\right)\right\}.\end{aligned} \tag{11}$$

Each of these phase terms has an intuitively satisfying interpretation as given in Table 1. Our subsequent processing steps are based upon this generic signal data form.

Table 1: Interpretation of phase terms.

term	interpretation
1	"DC" phase
$\omega_{\hat{y}}$	range position $\hat{Y}_0$ along $\hat{y}$
$\omega_{\hat{y}}\omega_b$	range walk
$\omega_{\hat{y}}\omega_b^2$	range curvature
ω_b	azimuth position $-X_0$ along x
ω_b^2	linear FM

2.2 Multiplication: 1st Order Processing

The second processing step involves the removal of certain exponential (phase) terms in equation (11) via complex multiplication by their conjugate:

$$\exp\left\{j\,\omega_{\hat{y}}\hat{Y}_0\left[\frac{\tan\alpha}{\cos\alpha}\frac{\omega_b}{2k}+\frac{(1+3\tan^2\alpha)}{\cos^2\alpha}\frac{\omega_b^2}{8k^2}\right]\right\}$$

$$\cdot\exp\left\{-j\,\frac{\hat{Y}_0}{\cos^2\alpha}\frac{\omega_b^2}{4k}\right\}. \qquad (12)$$

This multiplier is single-valued and known at all locations in the $(\omega_b,\omega_{\hat{y}})$ domain based on the chosen reference point in the signal data space.

The function $\tilde{s}_1(\omega_\zeta,\omega_{\hat{y}})$ resulting from this step is a first-order range walk and range curvature correction (RCC) as well as a first order azimuth compression (if a range coded pulse is used, the complex multiplication required for range compression should also be done in this step). The signal data is essentially corrected for range curvature and azimuthally compressed in accordance with the data centre location $(-X_0,\hat{Y}_0)$. Subsequent processing steps perform the residual RCC and azimuthal compression for other points relative to $(-X_0,\hat{Y}_0)$.

2.3 Rangeward Inverse Fourier Transform

The third processing step is the calculation of the inverse Fourier transform in the range dimension to move into the $(\omega_\zeta,\hat{y})$ (range-Doppler) domain. The result after inverse Fourier transformation may be written down directly as

$$\tilde{s}_1(\omega_\zeta,\hat{y})\approx \qquad (13)$$
$$C_1\,p\left\{\hat{y}-\hat{Y}_0-\eta\left[1+\frac{\tan\alpha}{\cos\alpha}\frac{\omega_b}{2k}+\frac{(1+3\tan^2\alpha)}{\cos^2\alpha}\frac{\omega_b^2}{8k^2}\right]\right\}$$
$$w\left(-\frac{\omega_\zeta R}{2k}\right)\cdot\exp\left\{-2k\left(\hat{Y}_0+\eta\cos^2\alpha\right)\right\}$$
$$\cdot\exp\left\{j\,\eta\left(\sin\alpha\,\omega_b+\frac{1}{\cos^2\alpha}\frac{\omega_b^2}{4k}\right)\right\}.$$

2.4 Change of Variables: 2nd Order RCC

The fourth processing step requires the change of variables

$$\hat{y}=\hat{y}'\cdot\left[1+\frac{\tan\alpha}{\cos\alpha}\frac{\omega_b}{2k}+\frac{(1+3\tan^2\alpha)}{\cos^2\alpha}\frac{\omega_b^2}{8k^2}\right], \qquad (14)$$

to effect the remaining RCC. This change of variables amounts to a single-valued perturbation implemented as an interpolation step in $(\omega_\zeta,\hat{y})$ space for digital data. This perturbation will have negligible impact upon the ability to fully focus the SAR data in azimuth. Correct positioning of the range pulse locations is ensured.

2.5 Multiplication: 2nd Order Compression

The fifth processing step requires complex multiplication by the conjugate of the remaining phase terms which are now known and constant at each point in range-Doppler space. The multiplier is:

$$\exp\left\{-j\,\eta\left[\omega_b\sin\alpha+\frac{1}{\cos^2\alpha}\frac{\omega_b^2}{4k}\right]\right\}. \qquad (15)$$

The term $(\omega_b\sin\alpha)$ achieves correct geometric positioning of the resulting impulse response in the image domain. The term proportional to ω_b^2 provides the azimuth matched filter for compression. The required function of equation (15) is single-valued on $(\omega_b,\hat{y}')$. The result of this complex multiplication is a new function which we denote $\tilde{s}_2(\omega_\zeta,\hat{y}')$.

2.6 Azimuthal Inverse Fourier Transform

The final processing step involves calculation of the inverse Fourier transform in the azimuth dimension to move back into the $(\zeta,\hat{y}')$ (spatial image) domain. The explicit dependence on η may be removed by considering a scatterer at range R, as defined in Fig. 1. We find that

$$s_2(\zeta,\hat{y}')\approx C_3\,p\left(\hat{y}'-\frac{R}{\cos\alpha}\right)w\left(\frac{2k\zeta}{R}\right)\cdot\exp\left\{-j\frac{2kR}{\cos\alpha}\right\}, \qquad (16)$$

where C_3 is a new complex constant.

A full complex image results from convolution of the impulse response of equation (16) over the complex scattering field. This result represents a complex SAR image which is free from range invariance regions, hence, phase discontinuities in range. Detection (preferably a magnitude-squared operation) yields the usual SAR image product.

3 Simulation Results

The phase preserving SAR processor methodology is being implemented and tested on a digital computer using a squint mode point scatterer complex signal data simulation model. The model is based upon a Gaussian weighting function for the azimuthal antenna pattern (in which β is the azimuthal antenna beamwidth) and a compressed Gaussian range impulse (in which ρ_r is the range resolution). The simulated data is based upon calculating the instantaneous range to the target and inserting this result into a phase term in the signal generator.

The required signal data simulation parameters are outlined in Table 2, with specific values for an airborne SAR case indicated. The parameters are based upon design equations with 20% guard bands and were chosen to emulate the CCRS CV-580 C-band SAR system (Livingstone *et al.*, 1987).

In Fig. 3, we present detected impulse response results of the two data sets considered. The impulse response is one-quarter swath further out in slant range than the reference point, thus exercising all steps in the algorithm. We present the impulse response both with and without the range-Doppler interpolation step. For the squinted case, we see a substantial improvement

in the impulse response we invoke this interpolation step. The azimuthal tail on the squinted impulse response is a consequence of approaching the design constraint of equation (8).

These results demonstrate the capability of this processing approach to produce a conventional SAR image for the case of a high resolution airborne SAR and a high resolution squint mode airborne SAR. We have yet to analyze the phase performance. Actual airborne SAR data has been compressed with this processor.

4 Conclusions

The phase preserving SAR processor which we have outlined presents the potential to support advanced SAR applications by providing uncorrupted image phase information. The extension of our simulation results to a full-fledged SAR processing capability is ongoing. Issues being considered include the development of a Doppler centroid estimator and a Doppler ambiguity resolver in order to accurately estimate the squint angle α, the development of a software-based platform motion compensation capability, and the study of the phase preserving characteristics of this processor.

An important part of the latter issue is the development of a robust capability to perform complex interpolation. This is being investigated in terms of a new amplitude/phase unwrap algorithm. If the data is substantially over sampled, interpolating on the in-phase and quadrature components is a reasonable approximation. However, actual SAR data is generally only adequately sampled, in which case such an interpolation approach often creates undesirable image artifacts. Accurate complex interpolation is at the heart of many modern SAR applications, and is an interesting topic in its own right.

5 Acknowledgements

The authors wish to thank the following individuals for their technical contributions to this project. C. Livingstone (CCRS), P. McConnell (MDA), P. McLean (CCRS), J. Princz (CCRS), and J. Wolfe (CCRS).

Table 2: Parameter sets used in the simulations.

parameter			Case 1	Case 2
azimuth samples	N_{az}		2048	2048
azimuth sample spacing	δx	(m)	0.39	0.39
range samples	N_{ra}		2048	2048
range sample spacing	δy	(m)	4.0	4.0
centre target range	R	(m)	8000	8000
radar wavelength	λ	(m)	0.056	0.056
range resolution	ρ_r	(m)	6.0	6.0
squint angle	α	(°)	0.0	45.0
azimuth beamwidth	β	(°)	3.6	3.6

References

1. Calfforio, C., C. Prati, and F. Rocca "Full resolution focusing of SEASAT SAR images in the frequency-wave number domain", EARSeL Workshop, Capri (Naples), 15-20 May, 1988.

2 DiCenzo, A., "A comparison of resolution for spotlight synthetic-aperture Radar and computer-aided tomography", *Proc. IEEE*, Vol. 74, No. 8, pp 1165-1166, 1986.

3. Gazdag, J., and P. Sguazzero, "Migration of seismic data", *Proc. IEEE*, Vol. 72, No. 10, pp 1302-1315, 1984.

4 Livingstone, C.E., A.L. Gray, R.K. Hawkins, R.B. Olsen, R A Deane, and J G. Halbertsma, "CCRS C-band airborne radar—system description and test results," to appear in *Proc. 11th Canadian Symposium on Remote Sensing*, Waterloo, Ontario, June 22-25, 1987.

5. Rocca, F. "Synthetic aperture radar: a new application for wave equation techniques", *SEP-56*, pp 167-189, 1987.

6. Stolt, R.H. "Migration by Fourier transform", *Geophysics*, Vol. 43, No. 1, pp 23-48, 1978.

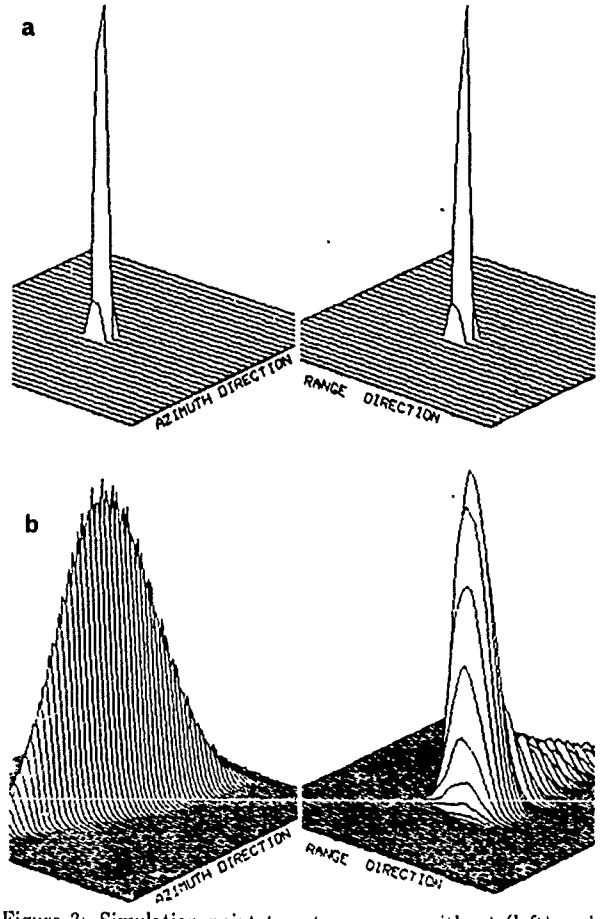


Figure 3: Simulation point target responses without (left) and with (right) the range-Doppler interpolation step for the input parameters of Table 1. a) High resolution airborne SAR and b) high resolution airborne SAR with 45° squint.

SAR Super-resolution using a Perturbed Point Spread Function

D.Blacknell

GEC-Marconi Research Centre
West Hanningfield Road
Great Baddow
Chelmsford
Essex
CM2 8HN
United Kingdom

S.Quegan

The Department of Applied and
Computational Mathematics
The University of Sheffield
Western Bank
Sheffield
S10 2TN
United Kingdom

Abstract

Super-resolution techniques attempt to improve upon the theoretically optimum resolution of a coherent imaging system by using prior knowledge about the imaged scene. Ideal operation requires the exact point spread function (PSF) of the imaging system to be known. For airborne SAR systems, even after motion compensation, residual across-track motions of the SAR platform will cause unknown perturbations of the PSF. In this paper, the effects of such perturbations on the performance of a particular super-resolution technique are illustrated. The consequences for the use of super-resolution with airborne SAR systems are discussed.

Keywords

synthetic aperture radar, SAR, motion compensation, super-resolution

1 Introduction

For coherent imaging systems, such as synthetic aperture radar (SAR), there is a theoretical limit to the resolution which can be achieved when, prior to imaging, nothing is known about the object being imaged. Super-resolution techniques attempt to achieve better resolution than this theoretical limit by partially inverting the imaging process using prior knowledge about the object.

Much of the motivation for super-resolution work comes from the result that, for a band-limiting imaging system mapping from a continuous object space to a continuous image space, if it is known that the object is of limited spatial extent (the prior knowledge) then the object can be reconstructed exactly given any finite portion of the image, i e infinitely narrow resolution can be achieved. Of course, in reality, practical considerations such as additive noise on the image will degrade the performance of super-resolution techniques. Methods have been developed to limit the detrimental effects of additive noise and theoretical studies have shown resolution enhancement by factors of two to four to be possible However, a further practical consideration which has not been considered previously is the fact that the point spread function (PSF) used in formulating the super-resolution reconstruction may not exactly match the PSF which characterizes the imaging process. This is especially true in the case of airborne SAR for which uncorrected across-track motions of the SAR platform cause unknown perturbations of the PSF from its ideal form. Motion compensation schemes can be used to remove the gross effects of these motions, but some residual perturbation will still remain. This will produce a further degradation in the performance of super-resolution techniques It is intended to characterize this degradation in performance and discuss the consequences for the use of super-resolution with airborne SAR systems

The super-resolution technique to be considered is the stochastic inverse method (Luttrell, 1985 and Delves et al, 1988) which is described in Section 2. In Section 3, a simulation is used to establish a baseline performance of the technique in the absence of PSF perturbations. In Section 4, the PSF is perturbed to a degree which is consistent with the performance of a typical airborne SAR system (an X-band SAR operated by the Royal Signals and Radar Establishment) after optimum motion compensation has been applied (Blacknell, 1989) and the consequent degradation in performance is assessed. Conclusions are given in Section 5.

2 The Stochastic Inverse Method

Let the imaging process be described as,

$$g = Tf + n, \tag{1}$$

where,

$g \equiv$ the image,
$f \equiv$ the object,
$T \equiv$ the imaging operator,
$n \equiv$ the additive noise.

The best linear estimate of the object in a minimum mean square error sense is,

$$f_r = WT^\dagger[TWT^\dagger + N]^{-1}g, \tag{2}$$

where,

$f_r \equiv$ the reconstructed object,
$W \equiv < ff^\dagger >$,
$N \equiv < nn^\dagger >$.

The covariance matrix, W, contains the prior knowledge and must be estimated before a reconstruction can be obtained. Typically, the covariance matrix will be assumed to be diagonal which means that the prior knowledge can be specified by estimating the expected power of the object at each pixel position required for the object reconstuction. If a previous reconstruction is available, then the prior knowledge can be obtained by choosing some suitable threshold and setting the expected power to the mean background value at those positions where the pixel value is less than the threshold and setting the expected power to the actual power of the reconstructed pixel at those positions where the pixel value is greater than the threshold. This choice of prior knowledge reflects the belief that objects of interest will be seen as bright targets against a dark background To produce the first reconstruction, the super-resolution technique is applied using no prior knowledge ($W \equiv$ the identity matrix). Thereafter the process is iterated with each new reconstruction providing the prior knowledge for the next reconstruction until the procedure converges to a consistent reconstruction.

3 Baseline Performance

For this investigation, objects specified on a 32×32 pixel grid are imaged to produce images specified on an 8×8 pixel grid The objects and images have complex pixel values but the intensity values are used for display purposes The point spread function (PSF) is that which would result if the SAR platform was free from across-track motions (It is not the perfect $sinc(x)$ function since a non-uniform antenna illumination has been considered) The image is sampled at the Nyquist rate appropriate for the bandwidth of the system being modelled Thus an object reconstruction onto the 32×32 pixel grid represents an attempt to improve the resolution by a factor of four The object to be considered is an arrangement of two point targets which have equal amplitudes and are 90° out of phase (in quadrature) and which are at identical range positions but are separated by three quarters of

a resolution cell in azimuth The closeness of the targets provides a suitable challenge for the super-resolution technique This object is imaged using the perfect PSF and the resulting image is corrupted with 20dB additive noise The image is then resampled onto a 32×32 pixel grid Forty-nine additive noise realizations were used and the resulting forty-nine resampled images are shown in Figure 1. (For all four figures presented in this paper, the highest valued pixel in each component image has been set to black and pixels of zero intensity have been set to white in order to show features of interest more clearly. The azimuth direction is from left to right and the range direction is from bottom to top The individual component images within the array will be referred to using (x,y) coordinates where x runs from 1 to 7 and from left to right and where y runs from 1 to 7 and from bottom to top) The stochastic inverse super-resolution technique was applied to each of these images using a maximum of fifteen iterations. The resulting forty-nine reconstructed objects are shown in Figure 2.

It can be seen from Figure 1 that each image appears as a featureless 'blob' surrounded by clutter. Although the method of display which has been used tends to emphasize the structure of the clutter (the lower intensities) rather than the structure of the blob (the higher intensities) it is, in fact, the case that the blob has no internal structure to indicate the nature of the original object. This is to be expected given the sub-resolution cell spacing of the point targets. Figure 2 shows immediately that the super-resolution technique has been very successful in discerning the structure of the original object. There are, of course, some misleading reconstructions

For instance, in some cases more than two point targets have been resolved. One of the more severe examples of this is reconstruction (3,6) which shows three point targets each at different range positions One of the targets is correctly positioned but the other two are completely incorrect Less severe examples have two point targets of approximately equal strength at the correct positions but also have spurious lower intensity point targets at other positions, for example reconstructions (3,7), (1,4) and (4,4) amongst others. Referring to Figure 1, it can be seen how these spurious point targets have arisen at positions where the background clutter is quite high. Once again employing our belief that objects of interest will be of high intensity, we feel justified in thresholding the reconstructions at the 3dB or half power level This leaves reconstructions of three types. those which show only one point target, those which show two point targets either or both of which may be incorrectly positioned and those which show two point targets at the correct positions

Assessing the quality of a reconstruction is inevitably subjective since the assessment will depend on what properties of the reconstruction the observer feels to be important It is our opinion that the phase of the scatterers is not of great importance when attempting to identify the target which is composed of these scatterers and so the accuracy of the reconstructed phase has not been addressed For this reason, the performance assessment to be used is the proportion of reconstructions which show two point targets of approximately equal intensity at the correct positions. It transpires that thirty-six of the forty-nine reconstructions satisfy this criterion giving an (admittedly ad hoc) baseline performance assessment of 72%.

4 The Effect of a Perturbed Point Spread Function

A further forty-nine resampled images were generated as before except that this time the PSF in the azimuth direction was a realization of a perturbed PSF resulting from a typical airborne SAR system after optimum motion compensation had been applied These are shown in Figure 3 and the corresponding reconstructions are shown in Figure 4. It can be seen that the performance of the super-resolution technique has been significantly degraded. There are many more examples of spurious point targets being resolved. A typical example is reconstruction (7,7) where two point targets have been resolved of approximately equal intensity and the correct distance apart (although shifted in azimuth – see below) but a further point target has also been resolved. Referring to image (7,7), it can be seen that there is a distinct blob at the position where the spurious point target has been detected This is the result not only of the clutter but also of the high sidelobes which may occur for the perturbed PSF. A further characteristic of perturbed PSFs is a shift in the position of the main lobe. This may cause a problem when trying to accurately locate targets on the ground but is not of importance for super-resolution. It is only the relative positioning which is of importance.

Again employing a 3dB threshold on the reconstructions removes some of the spurious point targets. Counting the number of reconstructions which have resolved two point targets of approximately the same intensity and the correct distance apart gives a performance assessment of twenty-four out of forty-nine, i.e. about 48%.

5 Conclusions

Much of the theoretical work on super-resolution has forged ahead happily disregarding the problems involved in accurately characterizing the PSF of the imaging system In many cases, where the imaging system being considered is not specified, this is probably justified since the results will apply to optical systems for which the PSF can be characterized very accurately. However, when attempting to apply super-resolution to SAR, especially airborne SAR, the effects of perturbations in the PSF must be taken into account The main purpose of this paper has been to illustarate the degradation in performance which can occur Our 'performance measure' is no more than an ad hoc indicator of what can be interpreted visually from the figures and should not be taken too seriously.

On looking at Figures 3 and 4, it may be felt that, if this is the best that can be done on simulated images, the super-resolution technique must be improved before it can be applied to actual SAR images. There are three approaches to achieving this.

1. Use better motion compensation.
   The motion compensation scheme underlying this work (Blacknell, 1989) operates without inertial measurement unit (IMU) measurements and cannot, therefore, perform as well as a motion compensation scheme which uses accurate IMU measurements. It may be that, by using a better motion compensation scheme, sufficiently good image quality can be achieved to allow the super-resolution technique to be applied as it stands.

2. Use a better estimate of the PSF
   In this paper, it has been assumed that no attempt has been made to estimate the PSF from the image itself. If this were to be done in the locality of the object of interest, an improved performance may be possible. However the problem of estimating the PSF from the image is not trivial.

3 Improve the super-resolution technique
   As a last resort, it may be necessary to attempt to generalize the super-resolution technique to take into account that the PSF is not deterministic but is a statistical entity whose properties depend on the typical SAR platform motions which occur and the motion compensation which has been used This is quite a difficult problem.

The use of super-resolution with airborne SAR systems may still prove to be profitable but much work has yet to be done.

6 Acknowledgements

The authors would like to thank Prof. C.J.Oliver, Dr. S.P.Luttrell and Dr. R.G White of the Royal Signals and Radar Establishment, U.K., for useful discussions during the course of this work. Thanks are also extended to Dr P Muller of the University College, London who provided the image display software for the SUN workstation and Dr.P.J.Harley who took the photographs This paper was prepared with LaTeX which is a special version of Donald Knuth's TeX typesetting system. The work was supported by the GEC-Marconi Research Centre, U.K.

7 References

1 Luttrell,S P , 'Prior Knowledge and Object Reconstruction using the Best Linear Estimate Technique', Optica Acta, Vol 32, No 6, pp.703-716, 1985.

2 Delves,L.M ,Pryde,G.C. and Luttrell,S.P., 'A Super-resolution Algorithm for SAR Images', Inverse Problems, No.4, pp.681-703, 1988.

3 Blacknell,D and Quegan,S , 'SAR Motion Compensation using Autofocus', submitted to the International Journal of Remote Sensing.

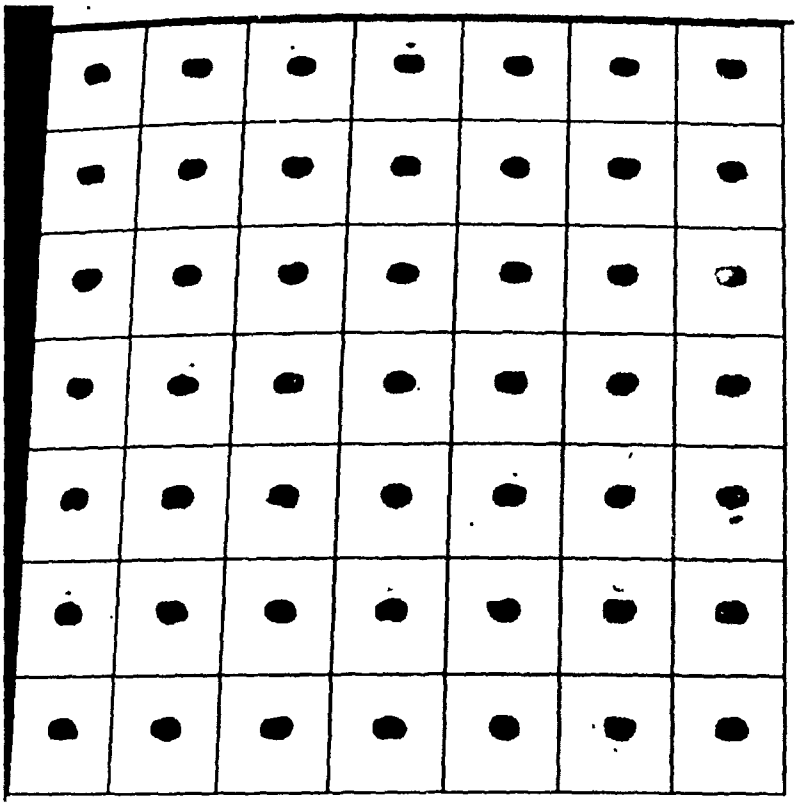


Figure 1

Resampled images of quadrature dipole using forty-nine additive noise realizations.

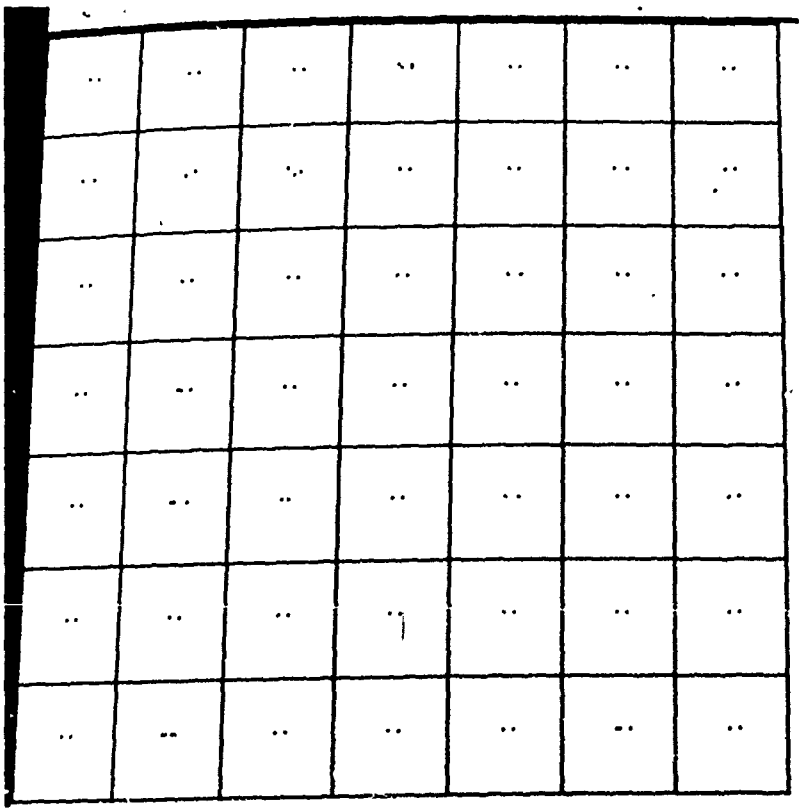


Figure 2

Reconstructions of images shown in Figure 1 using the stochastic inverse super-resolution technique.

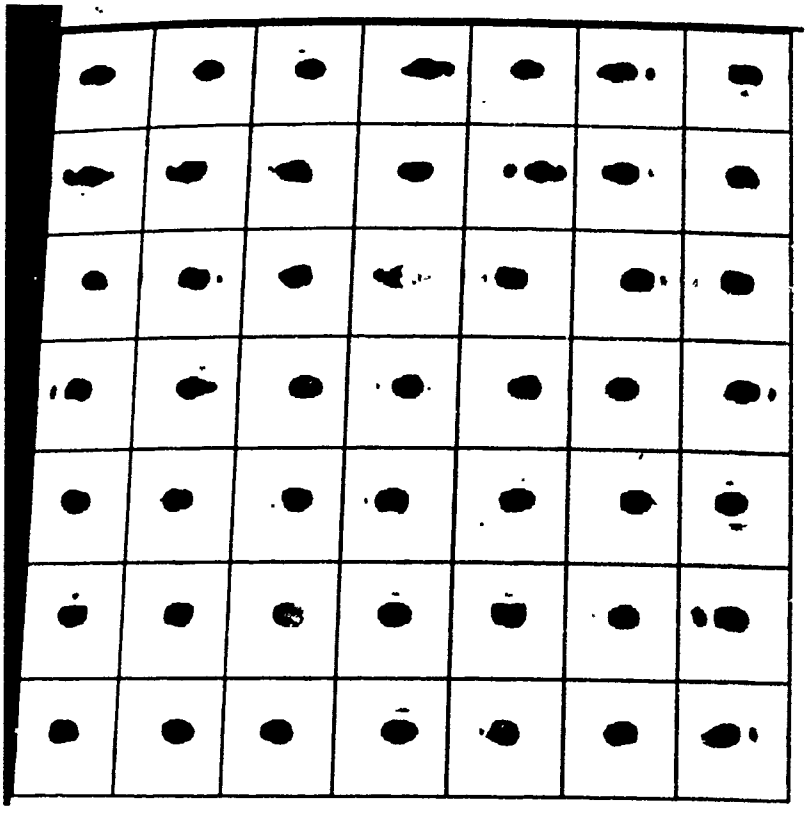


Figure 3

Resampled images of quadrature dipole using forty-nine additive noise realizations and forty-nine perturbed azimuth PSF realizations.

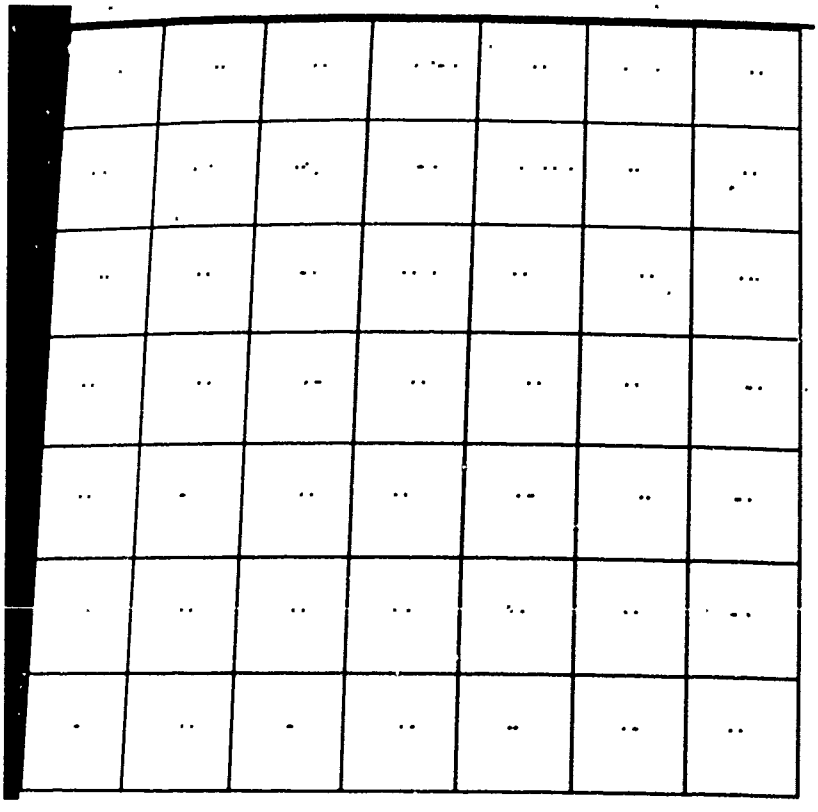


Figure 4

Reconstructions of images shown in Figure 3 using the stochastic inverse super-resolution technique.

FAST SAR PROCESSING WITH THE FERMAT NUMBER TRANSFORM

Jørgen Dall

Electromagnetics Institute
Technical University of Denmark
DK-2800 Lyngby, Denmark

Abstract
Most high resolution SAR processors are based on matched filtering implemented via the Discrete Fourier Transform (DFT). The DFT in turn is computed with the fast algorithm known as the FFT. However, a much lower computational complexity is attainable with the Fermat Number Transform (FNT) which accordingly have the potential of offering faster and/or simpler processors.

Two potential problems must be taken into account when applying the FNT for SAR processing. Firstly, the FNT offers a limited dynamic range and secondly the FNT does not provide the Fourier spectrum as a spin-off of the filtering. The SAR images presented in this paper verify that in spite of these deficiencies the FNT is indeed applicable to SAR processing.

Keywords: Synthetic Aperture Radar, digital processing, Fermat Number Transform.

1. Introduction

Processing raw Synthetic Aperture Radar (SAR) data into high resolution images requires comprehensive computations to be carried out on vast amounts of data. Therefore, in order to do the processing in real time fast algorithms must be applied.

By the SAR processing a spatially dispersed raw image is focused so that the energy from the individual point scatterers is compressed into single points. In some SAR processors the pulse compression is accomplished with a technique based on mixing and spectral analysis, but most processors apply matched filtering. A digital matched filter is a Finite Impulse Response (FIR) filter matched to the input signal. Mathematically filtering is described by an operation called convolution, and according to its definition a convolution can be computed as a sum of products. However, in case of long filters this straightforward approach is inefficient and faster algorithms must be applied. Traditionally long filters have been implemented via the Discrete Fourier Transform by using the Fast Fourier Transform (FFT). However, several other algorithms support fast convolution (Blahut, 1985) and although less well-known, some of them are superior to the FFT. From among thirteen different fast transform and convolution algorithms studied with the emphasis on SAR processing applications the Fermat Number Transform (FNT) proved most attractive (Dall, 1988), (Dall, 1989).

The subject of this paper is the application of the FNT for SAR processing. In Section 2 the FNT is defined, and its advantages are discussed in Section 3. Section 4 compares two off-line SAR processors implemented with the FFT and the FNT, respectively. Likewise two SEASAT SAR processors are compared in Section 5.

2. The Fermat Number Transform

The Fermat Number Transform, G_m, of a signal, g_n, is defined by

$$G_m = \sum_{n=0}^{N-1} g_n \omega^{nm} \pmod{F_t} \qquad (1)$$

and the inverse transform is given by

$$g_n = \frac{1}{N} \sum_{n=0}^{N-1} G_m \omega^{-nm} \pmod{F_t} \qquad (2)$$

n is the time index, m is the index in the transform domain, N is the block length and F_t is one of the Fermat numbers

$$F_t = 2^b + 1 \qquad (3)$$

$$b = 2^t \quad , \quad t = 1, 2, \ldots$$

Is is seen that the definition of the FNT is quite similar to the definition of the DFT and yet there are important differences. All numbers involved in the FNT, including the kernel ω, are integers in the range $[0; F_t-1]$ and all arithmetic operations are modulo operations. This means that in principle the result of an addition, multiplication etc. is divided by F_t and only the remainder is retained.

The transform G_m in Eq. (1) has nothing to do with the Fourier spectrum and in fact no physical interpretation of G_m is known at present. Nevertheless, the FNT is valuable because it supports cyclic convolution in almost the same way as the DFT. That is, the cyclic convolution $g \otimes h$ is obtained by inverse transforming the product of the transforms of the two input signals, g and h.

$$\mathrm{FNT}(g \otimes h) = \mathrm{FNT}(g) \cdot \mathrm{FNT}(h) \qquad (4)$$

What is computed in this way is actually the convolution modulo the Fermat number F_t, but if it is known a priori that the samples of the desired convolution are in the range $[0; F_t-1]$, the modulo arithmetic has no importance at all. Also, if it is known that the desired convolution is confined

to any other interval of length $F_t = 2^b + 1$, e.g. $[-2^{b-1}; 2^{b-1}]$, it is simple to correct for the modulo arithmetic. F_t is simply subtracted from all samples exceeding 2^{b-1}.

The FNT differs from the FFT in that no scaling is possible at intermediate steps during the transformation, so the modulo arithmetic effectively limits the dynamic range allowed for g⊗h. Therefore it is desirable to choose a large F_t, $F_5 = 2^{32}+1$ or $F_6 = 2^{64}+1$ for instance. On the other hand an F_t in excess of F_4 means increased hardware costs or slower computation due to the larger word length. It also reduces the maximum transform length because unlike the DFT the FNT is not defined for every value of N. The possible transform lengths are powers of two and they are constrained by N_{max} which is listed in Table 1. N_{max} reflects the fact that F_t is a prime for $t \leq 4$ but not for $t > 4$.

F_t	N_{max}	$N_{\omega=2}$
$F_3 = 2^8+1$	256	16
$F_4 = 2^{16}+1$	65536	32
$F_5 = 2^{32}+1$	128	64
$F_6 = 2^{64}+1$	256	128

Table 1. The maximum transform length for various FNTs.

With a straightforward approach four integer convolutions are required to compute one complex convolution. However, using the FNT, complex signals can be convolved very elegantly with only two integer convolutions (Nussbaumer, 1981).

3. FNT merits

Unlike the DFT kernel, the FNT kernel, ω, is not uniquely specified for a given transform length, N (and Fermat number, F_t). The principal advantage of the FNT is that it is possible to choose ω so that the FNT can be computed with simple operations like additions and shifts while multiplications are almost completely avoided.

Obviously, if ω equals 2 in Eq. (1), all multiplications degenerate into shift operations. Unfortunately, the transform lengths associated with $\omega = 2$ (indicated by $N_{\omega=2}$ in Table 1) are much less than the lengths of the matched filters required for SAR processing. Recently, the Double Level Decimation (DLD) algorithm has been proposed as a means of circumventing the transform length problem (Dall, 1988). Several other techniques have been proposed, but at least for SAR processing these techniques are inferior to the DLD.

Table 2 compares the number of arithmetic operations required to compute a convolution with the FNT (DLD) and the DFT (FFT), respectively. The difference is a factor of 4.7 to 7.3 for the multiplications and a factor of about 1.5 for the total number of operations. It should also be mentioned that the FNT (DLD) is just as modular and repetitive as the DFT (FFT) and the memory requirements are the same too.

Another advantage of the FNT is that it is an exact transform, not only in theory but also in practice. FFT implementations introduce computation noise because the complex twiddle factors are represented with a finite accuracy, and because the results of the multiplications and additions are rounded or truncated. However, the FNT twiddle factors are integers representable with a finite number of bits, and due to the integer modulo arithmetic, no rounding or truncation is needed. The price paid for this desirable feature is the limited dynamic range discussed in Section 2.

Block length = 2×filter length	Mults		Adds/subs		Shifts		Total	
	FNT	FFT	FNT	FFT	FNT	FFT	FNT	FFT
64	12	56	52	76	26	0	90	132
128	12	64	60	88	30	0	102	152
256	12	72	68	100	34	0	114	172
512	12	80	76	112	38	0	126	192
1024	12	88	84	124	42	0	138	212
2048	20	96	92	136	46	0	158	232
4096	20	104	100	148	50	0	170	252
8192	20	112	108	160	54	0	182	272
16384	20	120	116	172	58	0	194	292
32768	20	128	124	184	62	0	206	312
65536	28	136	132	196	66	0	226	332

Table 2. The number of arithmetic operations per valid output sample for a complex convolution. The overlap-save technique is applied with an overlap of 50%.

4. SAR-580 processing

Two potential problems must be taken into account when applying the FNT for SAR processing. Firstly the FNT offers a limited dynamic range and secondly the FNT does not provide the Fourier spectrum as a spin-off of the filtering. To examine these issues the FNT defined modulo F_4 has been implemented on an HP1000/A900 minicomputer and incorporated in two existing SAR processors developed at the Technical University of Denmark (Dall, 1988).

Modulo arithmetic is very cumbersome on the HP1000/A900 computer, especially when coded in FORTRAN. A factor of three is gained by coding the FNT in assembler language which facilitates bit manipulations etc. Still, the FNT is only marginally faster than the FFT, as seen from Fig. 1. On another general-purpose computer, the IBM 360/370, the FNT has proved substantially faster than the FFT (Agarwal and Burrus, 1974), but in both cases the conclusion is that a special purpose hardware implementation is required to make the merits of the FNT manifest themselves completely. A pipeline FNT chip has already been designed (Truong et al., 1985), but the current VLSI technology allows still faster chips with more parallelism to be fabricated. Recently, a chip adopting a systolic array architecture has been proposed (Dall, 1988). It is expected to be two to five times faster than existing state-of-the-art FFT chips.

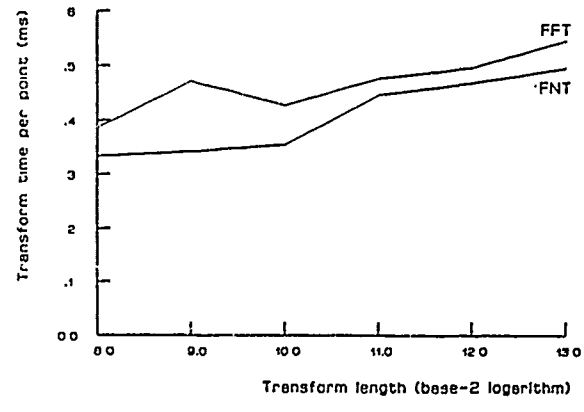


Fig. 1. Execution time of FNT and FFT programs.

In its original version the processor for the airborne SAR-580 uses the FFT for the range and azimuth compressions. Since the SAR-580 does not suffer from range curvature at X-band, both the range filtering and the azimuth filtering are one-dimensional. This means that it is quite straightforward to substitute the FNT for the FFT. The only non-trivial issue is the scaling of the input signals required to keep the amplitude of the processed signals within the dynamic range of the FNT. The scaling should be as moderate as possible, because it introduces additional quantization noise. A worst case scaling is much too conservative and therefore a kind of block adaptive technique is applied. The scale factor is controlled by a running observation of the mean amplitude of the input block to be filtered, g_{mean}. When each sample of an input block is multiplied by the scale factor

$$s = \frac{2^{15}}{r\,g_{mean}} \tag{5}$$

the output samples will fit into the allowed range $[-2^{15}; 2^{15}]$. The factor r is obtained empirically. After the filtering the output block is divided by s to make sure, that the individual blocks are not out of proportion.

Fig. 2 compares two SAR-580 images processed with the FFT and the FNT, respectively. They look almost exactly the same. The FFT used is a 32 bit floating point routine and hence its processing noise is negligible. This means that the quantization noise introduced by the FNT processing can be found by a complex subtraction of the two not yet detected images. The intensity of this noise is shown in Fig. 3. (Fig. 3 is not directly comparable with Fig. 2, because due to the photographic processing applied, Fig. 3 is too light). The crosses are the convolution of the filter quantization noise by the high intensity parts of the signal. The bright spots are due to overflow caused by an insufficient scaling in the azimuth processing. With a proper scaling a signal-to-processing-noise ratio of 20.5 dB is obtained. While this is adequate for most applications, 32 dB is attainable using a more comprehensive processing.

5. SEASAT SAR processing

It is much more complicated to integrate the FNT in the SEASAT SAR processor than in the SAR-580 processor. This is due to

- the real to complex signal conversion
- the subsampling
- the down-conversion to base band and
- the range curvature correction.

which usually are taking place in a SEASAT SAR processor (Bennett et al., 1980).

Unlike the raw SAR-580 signals the SEASAT signals are frequency offset and sampled in a single channel. These real signals must be converted to their equivalent analytic (complex) counterparts in order to meet the sampling theorem in the azimuth direction. When the range filtering is implemented with the FFT this is elegantly accomplished by setting one half of the spectrum equal to zero. The subsampling is then implemented by halving the block length which is equal to the period of the spectrum. More exactly these two operations are combined by not computing the part of the spectrum anyhow rejected. This in turn means that two real input signals can be transformed simultaneously with a single FFT.

When the filter is implemented with the FNT the above-mentioned frequency domain operations cannot be perform-

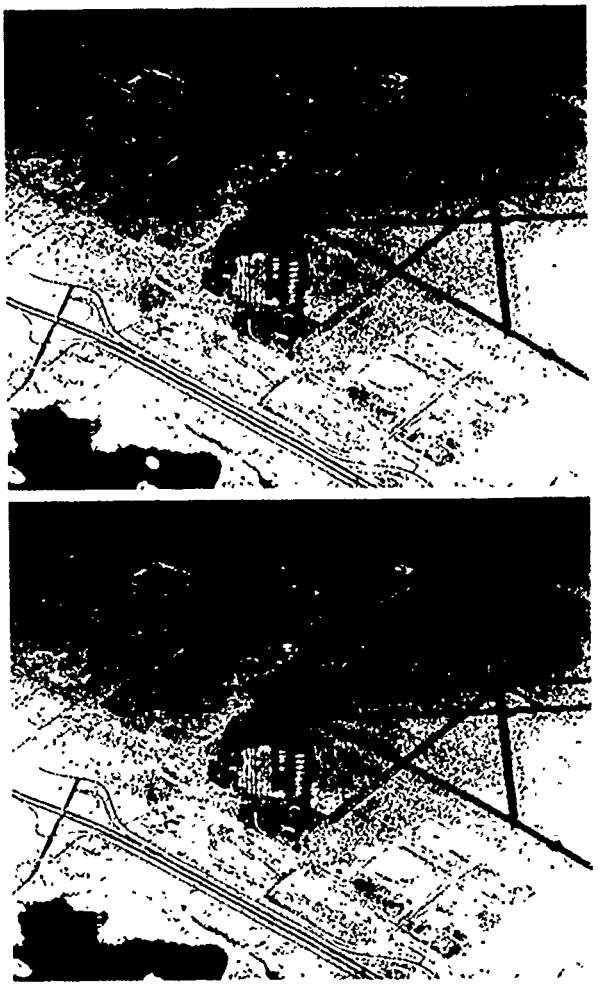


Fig. 2. SAR-580 image of the Willow Run Airport, Michigan. Top image: FFT processing. Bottom image: FNT processing.

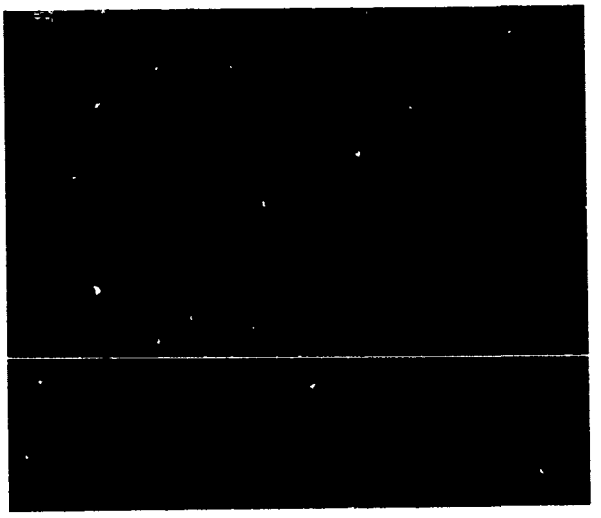


Fig. 3. SAR-580 difference image.

ed. Instead the real to complex signal conversion is accomplished by convolving the input signals by an analytic filter. The subsampling must take place in the time domain and hence be postponed to after the completion of the filtering. This means that the inverse FNT does not profit by the subsampling in the same way as the inverse FFT.

In the original FFT based SEASAT processor the down-conversion is implemented as a cyclic rotation of the range spectrum. This is not possible using the FNT. Fortunately, the down-conversion is not needed at all because a simple modification of the azimuth filter can compensate for the frequency offset.

When SEASAT data are processed to full resolution images the range curvature amounts to about eight range cells. Therefore a two-dimensional azimuth filter is required. Nevertheless, a one-dimensional filter suffices in the original SEASAT processor because it adopts the SAR processing scheme known as the hybrid algorithm (Wu et al., 1982). The word "hybrid" refers to the fact, that the algorithm operates in a mixed time-frequency domain. Range shifts in the range-Doppler domain straightens the target trajectories so that a one-dimensional filter can be applied. The hybrid algorithm requires fewer arithmetic operations than the true two-dimensional filtering, but it introduces more approximations. It also calls for a more complicated control.

Since the FNT does not provide the Doppler spectrum as a part of the azimuth filtering it is not compatible with the hybrid algorithm. Consequently, in order to substitute the FNT for the FFT, the SEASAT SAR processor must be modi fied to apply a two-dimensional azimuth filter. (Like the DFT the FNT can be generaiized to two dimensions).

The SEASAT images in Fig. 4 are processed with the FFT and the FNT, respectively, but they can hardly be distinguished from one another. In the SEASAT case the noise properties cannot be studied by means of difference images because also the substitution of the two-dimensional azimuth filtering for the hybrid algorithm makes the FNT image deviate from the FFT image. Applying FNT for the SAR processing implies the more noisy convolution algorithm but the better SAR algorithm.

6. Conclusions

When computed with the Double Level Decimation algorithm the FNT constitutes an attractive alternative to the FFT as far as filter implementations are concerned. With the FNT the number of multiplications is reduced by a factor of 5 to 7 and the total number of operations by a factor of 1.5 approximately. Special-purpose FNT chips are expected to be 2 to 5 times faster than existing state-of-the-art FFT chips, but with general-purpose computers the advantage may be much less.

The FNT does not provide the Fourier spectrum as a spin-off of the filtering. This is inconvenient in case of significant range curvature, but the main problem with the FNT is its limited dynamic range. A signal-to-processing-noise ratio of 20.5 to 32 dB has been measured. However, judging from the SAR-580 and SEASAT SAR images processed so far, the FNT seems to be adequate for most applications.

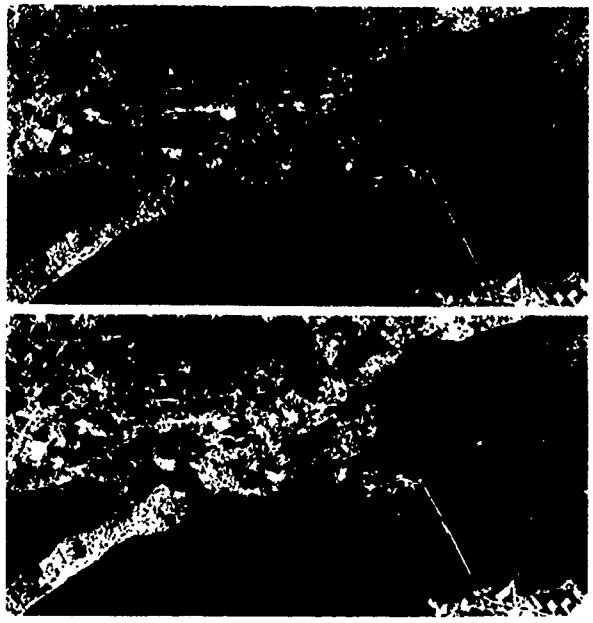


Fig. 4. SEASAT image of the southern part of Zeeland, Denmark. Top image: FFT processing. Bottom image: FNT processing.

References

1. Agarwal, R.C. and Burrus, C.S., "Fast Convolution Using Fermat Number Transforms with Applications to Digital Filtering", IEEE Trans. on Acoustics, Speech and Signal Processing, Vol. ASSP-22, No. 2, pp. 87-97, April 1974.

2. Bennettt, J.R., Cumming, I.G., Deane, R.A. "The Digital Processing of SEASAT Synthetic Aperture Radar Data", III International Radar Conference 1980, pp. 168-175, 1980.

3. Blahut, R.E., "Fast Algorithms for Digital Signal Processing", Addison Wesley, 1985.

4. Dall, J., "Fast Transform and Convolution Algorithms for Synthetic Aperture Radar Processing", Ph.D. dissertation, LD 70, Electromagnetics Institute, Technical University of Denmark, August 1988.

5. Dall, J., "Fast Convolution Algorithms for SAR Processing", to be printed in Proc. of Int. Conf. on Radar, Paris, April 1989.

6. Nussbaumer, H.J., "Fast Fourier Transform and Convolution Algorithms", Springer-Verlag, 1981.

7. Truong, T.K., Reed, I.S., Yeh, C.-S., Chang, J.J. and Shao, H.M., "A Parallel VLSI Architecture for a Digital Filter Using a Number-Theoretic Transform", In VLSI and Modern Signal Processing, by S.Y Kung, H.J Whitehouse and T. Keilath (Eds.), Prentice Hall, pp. 360-374, 1985.

8. Wu, C., Liu, K.Y. and Jin, M., "Modeling and a Correlation Algorithm for Spaceborne SAR Signals", IEEE Trans. on Aerospace and Electronic Systems, Vol. AES-18, No. 5, pp. 563-574, September 1982.

EFFECT OF REFERENCE FUNCTION WEIGHTING AND SELECTION OF OPTIMUM WEIGHTING COEFFICIENT FOR SYNTHETIC APERTURE RADAR

Archana Majmudar , A.M. Jha and N.S. Pillai

Microwave Sensors System Division
Microwave Sensors Development Group
Space Applications Centre
AHMEDABAD INDIA - 380053

Abstract

The impact of reference function weighting on Synthetic Aperture Radar (SAR) image quality parameters is presented. The Hanning weighting , $\eta + (1- \eta)\cos \omega t$, has been considered for the present analysis. The energy distribution of the compressed pulse alongwith the parameters like integrated side-lobe ratio (ISLR), mismatch loss, side-lobe reduction, etc. are considered for the evaluation of the preferred range of Hanning weighting coefficients, 'η'.

Key Words : Synthetic Aperture Radar (SAR), image impulse response, Hanning Weighting, ISLR, energy distribution, image resolution

1.0 Introduction

The image information in case of SAR is quantified in terms of the peak side-lobe level, main lobe resolution width, ISLR, energy distribution, mismatch loss, etc. The image is specified by the 2-D impulse response function which for SAR can suitably be characterised by two independent 1-D impulse responses; one in range and the other in the azimuth direction.

In order to have an increased system resolving capability in range and reduced peak power requirement for the transmitted signal, pulse compression technique is implemented by transmitting a linearly frequency modulated (LFM) signal. Further, the return in azimuth direction also has similar characteristics. The range and azimuth resolutions are obtained by carrying out appropriate matched filtering in the respective directions.

An ideal LFM matched filtered waveform gives a Sinc x impulse response, with a peak side-lobe level of -13.2 dB only. Such a low peak side-lobe will cause significant interfering signals for SAR where the imaging, due to the multitarget environment,is usually over a large dynamic range. Hence, image quality specifications usually call for a much better side-lobe level, which can be achieved by deliberately mismatching the received signal by amplitude weighting of the reference function during pulse compression [ref.1].

A detailed analysis of the relevant image quality parameters, which decide the selection of appropriate Hanning weighting coefficients, is presented in this paper. As the returned signal is LFM in both range and azimuth directions, the present analysis deals basically with the 1-D range impulse response, but the results hold good in general, for the azimuth impulse response also.

2.0 SAR Image Quality Variables

As stated earlier, a SAR using pulse compression will have a sinc x type of impulse response in both range and azimuth directions. The significant aspect of this impulse response is the existence of main lobe where most of the signal energy is contained. The main lobe is associated with side-lobes which extend infinitely in each direction. Due to the large dynamic range associated with the SAR imaging geometry, these side-lobes have significant interfering impact on image. Hence, the main problem in characterising the information content of SAR impulse response, is to associate some image quality variables. The three main parameters are the main lobe width related to the resolution, peak side-lobe level and the energy distribution in the overall image impulse response.

The most common choice for the definition of image resolution is the main lobe width corresponding to the intensity point 3 - dB below the main lobe peak. This definition approximates the Rayleigh criterion for the resolution of a non-coherent imaging system [ref.2]. As SAR is coherent in nature, the range and azimuth resolutions cannot be defined all the time by this 3 - dB point. While defining the resolution, other parameters like peak side-lobe level, energy distribution etc. are also to be specified alongwith. In present discussion, the resolution will be defined by the main lobe width of the image impulse response corresponding to some specified intensity level (e.g. -3 dB) below the main lobe peak in either direction (range and azimuth).

The sinc x type of impulse response corresponds to a 3 - dB resolution width of 0.88/B, where B is the transmitted signal (LFM) bandwidth in range direction or Doppler

bandwidth in azimuth direction. The side-lobe is ~13.2 dB down the main lobe peak. As analysis is similar in both the directions (range and azimuth), range only (as mentioned earlier) is considered presently. Hence, if B is the bandwidth of the transmitted LFM, the range resolution R_r is given as before by 0.88/B .

Two variables will be defined for describing the energy distribution in the image impulse response. Let E_a denote the ratio of the energy content in the main lobe width corresponding to R_r, to the energy content beyond R_r on either side. Let E_b denote the ratio of the energy content in image impulse beyond $2R_r$, to the energy lying within $2R_r$, i.e.

$$E_a = \frac{\text{Energy content in } R_r}{\text{Energy content beyond } R_r} \text{ , and}$$

$$E_b = \frac{\text{Energy content beyond } 2R_r}{\text{Energy content within } 2R_r}$$

Here, E_b is also defined as integrated side-lobe ratio (ISLR) . E_a will be expressed in percentage and E_b in dB. For ideal matched filtering, in case of pulse compression, E_a is ~73% ,whereas E_b is ~ -10.0 dB. The sidelobe obtained (-13.2 dB) is not at all sufficient and is going to camouflage the weaker targets. Hence, some means of reducing the sidelobes is required for SAR imaging. The most widely used approach is to employ the reference signal weighting. The amplitude of the reference function is weighted appropriately to carry out side-lobe suppression at the cost of main lobe widening. Various weightings like Taylor, raised cosine, Dolph-Chebyshev etc. are suggested and their impact on various pulse compresion output parameters are analysed much in detail in literature [ref. 1,3,4]. Type of weighting selected is generally specific to an application requirement. An optimum weighting in case of SAR can be chosen as a trade-off amongst the peak side-lobe improvement, main lobe widening, mismatch loss, E_a and ISLR of the compressed impulse response. For the present analysis, Hanning weighting, w(t), also known as weighted cosine on a pedestal,has been considered. It is expressed by

$$w(t) = \eta + (1-\eta) \cos (2\pi t/T) \quad ; \quad -T/2 < t < T/2$$

where,

η is the Hanning weighting coefficient, and
T is the transmitted pulse duration

The weighting coefficient η , varies from 0 to 1. Smaller the value of η, heavier is the weighting. Further, as is obvious, η equal to 1 corresponds to an unweighted case.

In next section, the effect of weighting on various SAR image quality variables is analysed and results are presented with an aim to arrive at an optimum range of Hanning weighting coefficient, η.

3.0 Selection of Hanning weighting coefficient

One of the major disadvantages of weighting is the broadening of the mainlobe. The variation of the percentage broadening of the 3-dB width with weighting coefficient η, is shown in fig. 1. As the 3- dB width broadens

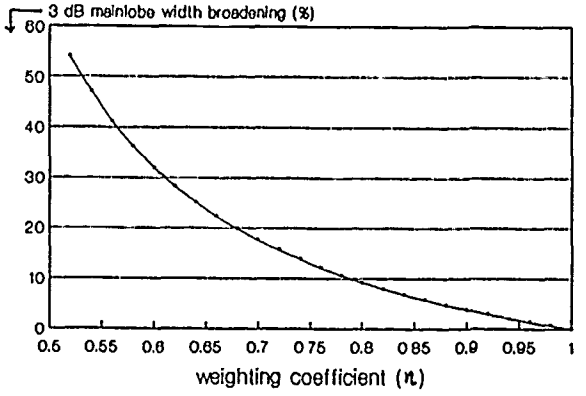


Fig.1 Variation of 3 - dB main lobe width (in %) with η

by more than 50 % for η around 0.5, the effect of Hanning weighting on various image quality parameters has been analysed for $0.5 < \eta < 1.0$ only. The peak side-lobe level variation with η is plotted in fig.2. If a side-lobe level of better than -20 dB is required , η has to be less than 0.77. As is shown in figure, the best side-lobe reduction is obtained for η = 0.54, which results in so called Hamming weighting but is not suitable for SAR as it

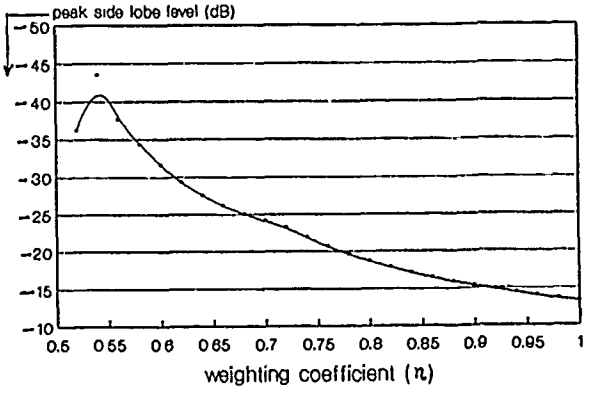


Fig.2 Variation of peak side-lobe level (in dB) with η

causes a main lobe broadening of ~50 % .

As weighting corresponds to a mismatched condition, SNR is going to be degraded. This SNR degradation is termed as mismatch loss and its variation with η is illustrated in fig. 3. As image SNR is very significant, more mismatch loss can not be tolerated. An overall mismatch loss of maximum 1 dB (including both range and azimuth compression) will result in $\eta > 0.67$.

As the side-lobe is suppressed and main lobe is broadened, the overall energy distribution of the image impulse response is going to be significantly affected. Let -m dB resolution correspond to the mainlobe width specified by intensity level m dB below the mainlobe peak. Fig. 4 shows the variation of

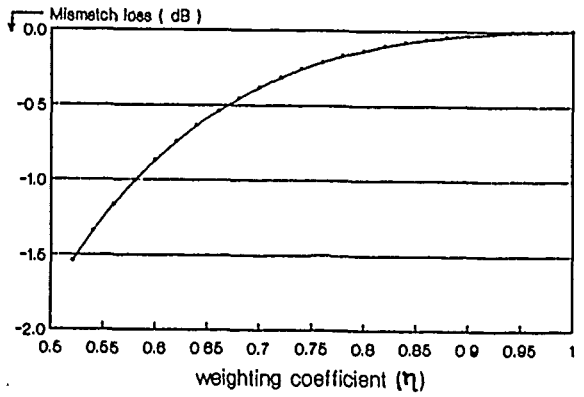


Fig.3 Variation of mismatch loss with η

energy distribution for various -m dB resolutions for η = 0.7, 0.8 and 1.0. A generalised energy distribution parameter E_η (m db) is defined as

$$E_\eta(\text{mdB}) = \frac{\text{Energy in -m dB resln. width}}{\text{Energy outside this resln. width}}$$

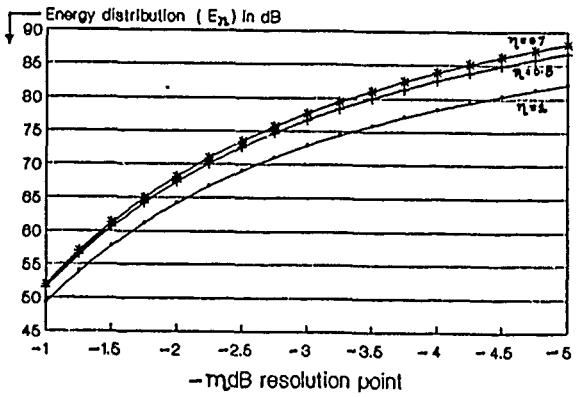


Fig.4 Variation of energy distribution E_η (m dB) with η

Fig.4 clearly shows that E_η increases with weighting. Infact, the energy distribution corresponding to -3dB resolution for η = 1.0 can be obtained by say ~ -2.4 dB resolution width for η = 0.8. This indicates that energy distribution E_1(-3dB) can be more or less maintained by weighting for the same resolution, even if the mainlobe broadens. As explained earlier, R_r represents the -3 dB

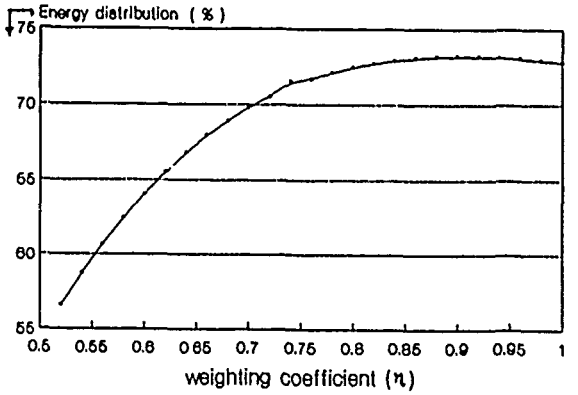


Fig.5 Variation of energy distribution E_a with η

resolution width on the unweighted image impulse response. If the -m dB reolution of the mainlobe width of the weighted impulse response is kept R_r (-3 dB resolution width for η = 1.0), the energy distribution E_a shows deterioration with increased weighting. The variation of E_a with η is plotted in fig.5. It clearly indicates that for η < 0.7, E_a falls significantly. It is found that for η > 0.7, E_a is reduced by less than 5 %. Hence, η > 0.7 alters the energy distribution in R_r very marginally. The resolution in case of Hanning weighting (η > .7) can equivalently be defined by that -m dB point which corresponds to a main lobe width of R_r in the weighted image impulse response.

ISLR as defined by E_b is one of the other significant image parameter which now needs to be analysed for weighted condition. The variation of ISLR with η is shown in fig. 6. When weighting is nominal (broadening is less), more and more energy is pushed inside the main lobe resulting in ISLR improvement which is maintained upto η = ~0.7. When weighting is heavier, broadening becomes significant and more and more energy starts occupying the area beyond $2R_r$ in the main lobe which results in degradation of ISLR. Hence, maximum improvement in ISLR is obtained for 0.68 < η < 0.78. This range of η gives an ISLR better than -14 dB.

TABLE - I

S.No.	Image Quality Parameter	Condition	Range of n
1.	Pulse broadening	broadening < 50 %	0.54 < η < 1.0
2.	Peak side-lobe level	side-lobe level < -20 dB	η < 0.77
3.	Mismatch loss	< 0.5 dB in each compression	0.67 < η < 1.0
4.	Energy distribution E_a	degradation < 5 %	0.70 < η < 1.0
5.	ISLR (E_b)	< -14 dB	0.68 < η < 0 78

TABLE-I summarises the discussions in this section. It lists the range of η obtained for various specified conditions associated with the image quality parameters. Thus , the optimum range of Hanning weighting

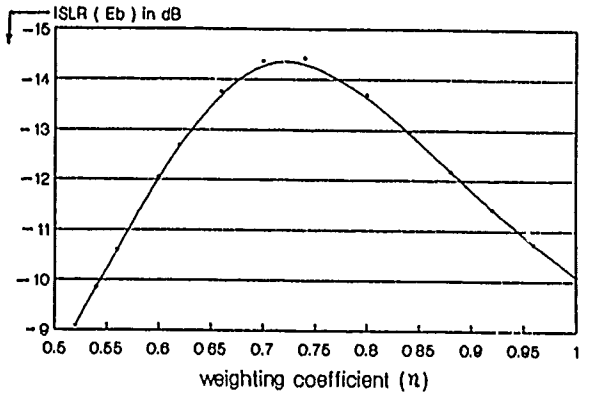


Fig.6 Variation of ISLR (E_b) with η

coefficients for the range and azimuth compression for SAR, which satisfies the various image quality requirement as listed in TABLE-I, is obtained as follows :

$$0.70 < \eta < 0.77$$

4.0 Conclusion

Impact of Hanning weighting on various SAR image quality parameters is presented. The range of Hanning weighting coefficients corresponding to a specified condition for the respective image quality parameter, is derived. An optimum range of Hanning weighting coefficient , $0.70 < \eta < 0.77$, is found appropriate for SAR imaging applications. It is found that for this range of η the resolution in case of weighted condition can equivalently be defined by that -m dB point ($m < 3$) which corresponds to a main lobe width same as -3 dB in unweighted condition.

References

1. Cook C.E. and Bernfield M.,'Radar Signals - An introduction to theory and application', Academic press,New York, London.

2. Mitchel R.H. and Marder S.,'Synthetic Aperture ,Radar (SAR) image quality considerations', Optical Engineering, Jan. /Feb. 1982/Vol. 21 No. 1.

3. Cook C.E., Paolillo J.,Bernfield M. and Palmieri C.A., ' Matched filtering, pulse compression and waveform design',Microwave Journal, Oct. 1964 - Jan. 1965.

4. Klauder J.R.,Price A.C., Darlington S. and Albersheim W.J., 'The theory and design of chirp radars ', Bell System Technical Journal 39, No. 4,July 1960.

DOPPLER PARAMETERS ESTIMATION CRITERIA FOR THE SPACEBORNE SAR AND CORRESPONDING ATTITUDE DETERMINATION ACCURACY LIMITS

TAPAN MISRA, A M JHA AND N S PILLAI

MICROWAVE SENSORS SYSTEM DIVISION
MICROWAVE SENSORS DEVELOPMENT GROUP
SPACE APPLICATIONS CENTRE (ISRO)
AHMEDABAD INDIA 380053

ABSTRACT

Two Doppler spectrum parameters namely Doppler centroid and Doppler rate fully characterise the Doppler phase history for a spaceborne SAR. Their estimate is important in generating azimuth matched filter and for Range Cell Migration (RCM) correction. The criteria of arriving at attitude determination accuracy and tolerable attitude drift specification for a spaceborne SAR has been developed.

Keywords: SAR, RCM, Doppler Centroid, Doppler Rate, LFM, Azimuth Compression, Matched Filter, PRF, Attitude Error, Attitude Drift, Localisation Accuracy.

Introduction

Synthetic Aperture Radar (SAR) is an imaging sensor where fine resolution is obtained by appropriate signal processing schemes. Fine range resolution is obtained by transmitting a very narrow pulse or using pulse compression technique. The real azimuth beamwidth of the SAR antenna is manifold larger than that required for achieving fine azimuth resolution. Fine azimuth resolution is achieved by storing the Doppler phase history of the backscattered signal (which incidentally is the same as that for an LFM signal) and by correlating with the expected phase history. This sort of signal processing is known as azimuth compression by matched filtering. In order to generate proper matched filter coefficients for azimuth compression and also to correct Range Cell Migration (RCM), the Doppler frequency has to be known precisely. Otherwise, the SAR image will be defocussed leading to image quality degradation. Two Doppler spectrum parameters namely Doppler centroid (f_{DC}) and Doppler rate (f_{DR}) fully characterise the Doppler phase history of a point target. Unfortunately, they are not constant for a particular SAR system but are sensitive functions of orbital position of the spaceborne SAR, attitude errors and misalignment between electrical and mechanical axes of the SAR antenna along azimuth. So both f_{DC} and f_{DR} are to be estimated for every image scene and their estimation accuracy will greatly dictate the image quality. Since orbital parameters are known fairly accurately, the estimation error in attitude parameters and beam misalignment along azimuth will dictate the accuracy of Doppler parameters estimated. In this paper, the criteria for arriving at measurement accuracy of attitude errors and absolute bound of attitude drift rate have been developed and the required specifications have been arrived at for a conceptual C-band spaceborne SAR system, operating in two modes SAR-1 and SAR-2 (look angles 20° and 35° respectively). The major system parameters of the same are listed in Table 1. As a corollary, a methodology of arriving at SAR image pixel location from the estimated Doppler centroid and orbital parameters has been derived.

TABLE - 1
System Parameters of the Conceptual C-band Spaceborne SAR

Altitude	800 km.
Frequency	5.3 GHz (C-band)
Off Nadir Look Angle	20° (SAR-1) 35° (SAR-2)
Swath	70 km. (SAR-1 & SAR-2)
Near Swath Ground Range	259 km. (SAR-1) 545 km. (SAR-2)
Far Swath Ground Range	329 km. (SAR-1) 615 km. (SAR-2)
Antenna	12 m. x 0.8 m.
PRF	1500 Hz - 1700 Hz
Nominal Satellite velocity	7.45 km./sec
Orbit Inclination	98° (with equator plane)
Resolution	36 m. x 36 m.
No. of Looks	6

Definition of Doppler Spectrum Parameters

Since the Doppler history of a point target follows closely as that of an LFM

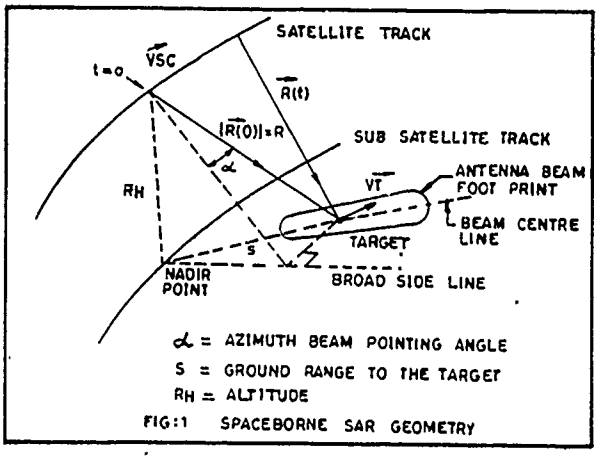


FIG:1 SPACEBORNE SAR GEOMETRY

signal, the Doppler frequency $f_d(t)$ at any azimuth time t can be expressed as,

$$f_d(t) = f_{DC} + f_{DR}t \; ; \quad -T/2 \le t \le T/2 \qquad ..(1)$$

where T is azimuth aperture time and t = 0 corresponting to the passing of beam centre line over the target [Fig. 1]).

The Doppler centroid (f_{DC}) is the Doppler frequency of the target when the beam centre line passes over the target. Since the received Doppler spectrum is weighted by the two way azimuth antenna pattern, f_{DC} is the frequency corresponding to the peak of the spectrum or the central frequency of the spectrum, if the azimuth antenna pattern is symmetric. Doppler rate (f_{DR}) is the chirp rate corresponding to Doppler LFM signal. As derived in refs. [1] & [2] f_{DC} and f_{DR} can be expressed as,

$$f_{DC} = -(2/\lambda R)[(V_T(0) - V_{SC}(0)).R(0)] \qquad ..(2)$$

$$f_{DR} = -(2/\lambda R)[(V_T(0)-V_{SC}(0))^2 - A_{SC}(0).R(0)] \qquad ..(3)$$

where,

$V_T(t)$ = velocity vector of the target
$V_{SC}(t)$ = velocity vector of the satellite
$R(t)$ = slant range to the target
$A_{SC}(t)$ = centripetal acceleration of the satellite
λ = radar wavelength

The above expressions of f_{DC} and f_{DR} are sensitive functions of orbital position of SAR, the attitude parameters and beam misalignment between electrical and mechanical axes in the azimuth direction. However, of all the three attitude parameters, yaw and pitch significantly influence f_{DC} and f_{DR}. The roll has negligible effect on them.

Need for Estimation of Doppler Centroid and Doppler Rate

Doppler centroid and Doppler rate estimation is needed to generate the azimuth matched filter coefficients and also to correct RCM. Let the received Doppler history of a target be,

$$h(t) = \exp(j2\pi f_{DC}t + j\pi f_{DR}t^2) \qquad ..(4)$$

$$-T/2 \le t \le T/2$$

So the corresponding matched filter function is given by,

$$h^*(-t) = \exp(j2\pi f_{DC}t - j\pi f_{DR}t^2) \; ; \qquad ..(5)$$

$$-T/2 \le t \le T/2$$

Hence, both f_{DC} and f_{DR} are to be estimated to generate the matched filter function. Since, azimuth phase history is sampled at PRF rate, only aliased f_{DC} i.e., f_{DC} modulo PRF only is needed to generate the azimuth matched filter. This aliased f_{DC} can be estimated by spectral analysis of azimuth data.

Further, the slant range to the target R(t) at any azimuth time t [Fig. 1] can be expressed as,

$$R(t) = R(0) - \lambda f_{DC}t/2 - \lambda f_{DR}t^2/4 \; ; \qquad ..(6)$$

$$-T/2 \le t \le T/2$$

Eqn. (6) gives the estimate of RCM. In order to correct both range walk and range curvature (i.e., linear and quadratic components of RCM) the estimate of unambiguous f_{DC} and f_{DR} is required.

Effects of Error in Estimation of Doppler Spectrum Parameters

An error in the estimate of f_{DC} will cause a degradation in SNR and AAR (Azimuth Ambiguity Ratio) of the SAR image. Figs. 2 & 3 show the degradation in SNR and AAR for various estimation errors in f_{DC} for the conceptual SAR system under consideration [Table 1]. AAR degrades more compared to SNR for same error in f_{DC} estimation.

Further, for multilook processing of SAR data, different looks will be misregistered along range due to inaccurate range walk correction. The misregistration will be maximum between extreme looks. The net misregistration x_r between extreme looks along slant range for f_{DC} estimation error of Δf_{DC} is given by,

$$x_r = \lambda \Delta f_{DC} T/2 \qquad ..(7)$$

where, T is the azimuth aperture time.

With erroneous estimate of f_{DR}, the looks will be misregistered along azimuth [1]. The total misregistration will be maximum between extreme looks. The misregistration x_{az} between extreme looks along azimuth can be expressed as,

$$x_{az} = \Delta f_{DR} \, T \, V_G/f_{DR} \qquad ..(8)$$

where,

Δf_{DR} = accuracy in f_{DR} estimate and
V_G = ground trace velocity

For the conceptual SAR system under consideration, for $1/6^{th}$ pixel misregistration between extreme looks, the accuracy requirement in f_{DR} estimate is

$|\Delta f_{DR}| \le 3.5$ Hz/sec (20° look angle)
and $|\Delta f_{DR}| \le 2.5$ Hz/sec (35° look angle)

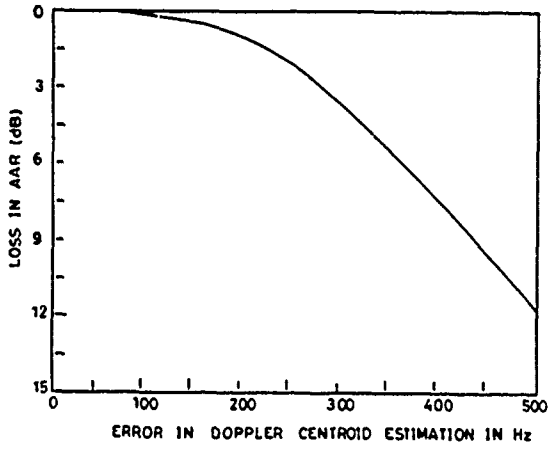


FIG. 2: LOSS IN AAR vs f_{DC} ESTIMATION ERROR

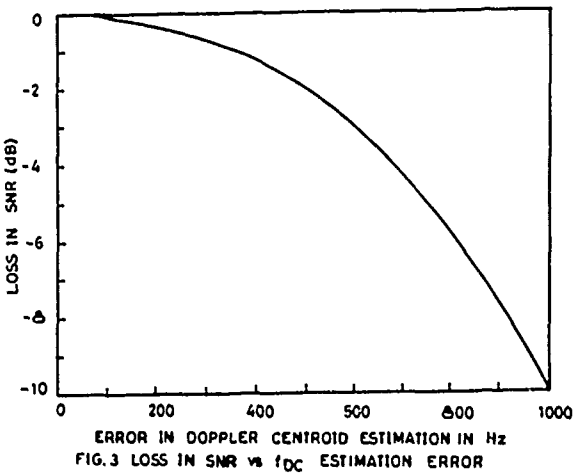


FIG.3 LOSS IN SNR vs f_{DC} ESTIMATION ERROR

The accuracy requirement in f_{DR} estimate for tolerable misregistration between extreme looks is much more stringent than that required for meeting acceptable quadratic phase error over the look bandwidth or acceptable range curvature correction.

Attitude Estimation Accuracy Requirement

For the C-band SAR under consideration, the above mentioned f_{DR} estimation accuracy can be met by relaxed accuracy in attitude estimate of the order of 1°. However, to achieve unambiguous Doppler centroid estimate from both the spectral analysis of azimuth data and ancillary information (which includes both attitude and orbit data), a much more stringent attitude estimation accuracy is called for. One point is worth noting : the requirement of attitude estimation accuracy for unambiguous f_{DC} estimation with the above mentioned method is comparatively much relaxed than that for f_{DC} estimation from ancillary data alone. The unambiguous Doppler centroid can be estimated in the following way.

Let f'_{DC} be the aliased estimate of Doppler centroid i.e., f_{DC} modulo PRF. This can be obtained by correlating the two way antenna gain pattern with the received Doppler spectrum or by finding the centroid of the Doppler spectrum, if the antenna pattern is symmetric. For generation of the matched filter, f'_{DC} is sufficient. In order to carry out the range walk correction, unambiguous Doppler centroid is to be estimated. So unambiguous Doppler centroid is

$$f_{DC} = f'_{DC} + n\ PRF \quad (n \text{ is an integer}) \quad ..(9)$$

Let f''_{DC} be the Doppler centroid estimated from attitude and orbit data only. The value of n (Eqn. (9)) which minimises the expression $|f''_{DC} - f'_{DC} - n\ PRF|$ is the unambiguous n provided the following condition is met :

$$|\Delta f_{DC}| = |f''_{DC} - f_{DC}| \leq PRF/2 \quad ..(10)$$

From Fig. 1 and eqn. (2),

$$|\Delta f_{DC}| = (2/\lambda R)|\Delta(V_{SC}(0).R(0))|$$

[Orbital position data gives fairly accurate estimate of the term $-(2/\lambda R)V_T(0).R(0)$ in eqn. (2)]

$$\text{Or, } |\Delta f_{DC}| = (2/\lambda)\, V_{SC}\, |\Delta\alpha|$$

$$\text{By eqn.(10) } |\Delta\alpha| \leq (\lambda PRF)/(4\ V_{SC}) \quad ..(11)$$

where, $\Delta\alpha$ is the estimation accuracy of azimuth beam pointing angle [Fig.1].

But, $PRF \geq (2/\lambda)V_{SC}\theta_{az}$ = Doppler bandwidth

where, θ_{az} is 3dB azimuth beamwidth

$$\text{Hence, } |\Delta\alpha| \leq \theta_{az}/2 \quad ..(12)$$

If $\Delta\theta$ be the attitude estimation accuracy and $\Delta\phi$ be the accuracy of beam misalignment along azimuth,

$$|\Delta\alpha| = (S+R_H)|\Delta\theta|/R + |\Delta\phi| \leq (\theta_{az}/2) \quad ..(13)$$

where, S = off nadir ground range, and
R_H = altitude of the satellite

For the C-band spaceborne SAR under consideration, $|\Delta\alpha| \leq 0.12°$

From eqn. (13) it can be concluded that the attitude estimation accuracy requirement is also dependent upon the accuracy with which beam misalignment along azimuth can be measured.

For a C-band SAR, the attitude estimation accuracy requirement is driven by the requirement of unambiguous Doppler centroid estimation. However, for a system in L-band (1.275 GHz), with rest of the system parameters identical to those in Table 1, due to large aperture time the attitude estimation accuracy will be governed by Doppler rate estimation accuracy requirement. The attitude estimation accuracy for the said L-band system is of the order of 0.25° for $1/6^{th}$ pixel misregistration. Whereas, for the same system $|\Delta\alpha| \leq 0.5°$ from unambiguous f_{DC} estimation requirement.

For the conceptual C-band SAR under

consideration, if the condition laid down in eqn. (13) is not met, due to poor estimation accuracy of attitude and beam misalignment, the clutterlock technique has to be employed to estimate f_{DC} unambiguously. In this method the residual ambiguity in the f_{DC} estimate is inferred by measuring the misregistration of extreme looks along range using eqn. (7). However, clutterlock technique may fail for imaging over uniform terrain or ocean due to the failure of detection of misregistration.

Attitude Drift Rate Requirement

The attitude drift rate does not affect f_{DR} since it is dependent upon the relative location of the target with respect to SAR. But unlike the optical sensors, the attitude drift does not cause relative location error between the targets within the same scene. Assume two targets which are separated along azimuth but lying in the same range gate. The azimuth time difference corresponding to beam centre line crossing over the two targets will be different for the attitude drift case than that in the non-drift case. Also, the Doppler centroid for the two targets will be different due to different attitude values when the beam centre line passes over them. However, when they are processed as part of the same scene, identical f_{DC} figure is used for azimuth compression. So the two targets will drift relatively in azimuth time due to the mismatch in their corresponding f_{DC} values from the f_{DC} value used for compression and this will exactly compensate the change in timing difference in azimuth corresponding to beam centre line passing over them. So after processing, the two targets will appear relative two each other as if the satellite was not drifting in attitude.

The drift in the attitude of the satellite will change the ground trace speed of the antenna beam over a target from the non-drifting case and consequently change the aperture time. Since f_{DR} remains constant, the Doppler bandwidth will change in the same proportion of the change in aperture time. The required attitude drift rate specification for 1% change in the Doppler bandwidth for the C-band SAR under consideration is

$$\text{attitude drift} \leq 3.0 \times 10^{-3}\ ^{o}/\text{sec} \quad (\ 20^{o}\ \text{Look Angle}\ )$$
$$\leq 2.3 \times 10^{-3}\ ^{o}/\text{sec} \quad (\ 35^{o}\ \text{Look Angle}\ )$$

The above result is interestingly independent of SAR frequency but dependent upon the system geometry.

Pixel Location from Doppler Centroid Estimate

The Doppler centroid can be expressed as a function of the projection of the slant range vector along azimuth. The pixel location in the SAR imagery along azimuth from the broad side line passing through the nadir point when the beam centre line passes over the target, can be obtained precisely from the unambiguous Doppler centroid estimate. In fact this accuracy will be limited mainly by the accuracy of aliased f_{DC} estimate by spectral analysis and the orbit accuracy which is the same for nadir point location accuracy. The localisation accuracy along azimuth is given by

$$|\Delta S_{az}| = \lambda R|\Delta f_{DC}|/(2\ V_{SC}) + |\Delta S_{nadir}| \quad ..(14)$$

where,

$|\Delta f_{DC}|$ = Doppler centroid estimation accuracy

$|\Delta S_{nadir}|$ = Nadir point location accuracy or orbit accuracy

Assuming $\Delta f_{DC} = \pm 50$ Hz and $\Delta S_{nadir} = \pm 0.5$ km., for the C-band SAR under consideration,

$$|\Delta S_{az}| \leq 0.7 \text{ km.}$$

Since the received signal is time gated along range precisely, the location accuracy in the broadside direction will be mainly limited by nadir point location accuracy. In this particular case, the localisation accuracy along broadside will be 0.5 km. same as nadir point location accuracy.

Conclusion

In this paper, two main Doppler spectrum parameters namely Doppler centroid and Doppler rate have been defined and their importance in SAR azimuth compression has been brought forth. Further, the criteria of arriving at required attitude estimation accuracy and tolerable attitude drift rate specification have been arrived at. The requisite attitude specifications have been derived for a conceptual C-band SAR. It has also been explained that unlike the spaceborne optical imaging system, the attitude drift does not pose problem of relative location error along azimuth direction for a spaceborne SAR. Further, a methodology of determining SAR image pixel location using Doppler centroid estimate and range gating information has been suggested.

References

1. **Doppler Parameter Estimation for Spaceborne Synthetic Aperture Radar**
By Fuk-Kwok Li, Daniel N Held, John C Curlander and Chialin Wu ; IEEE Trans. on Geoscience and Remotesensing, Vol. GE-23 No. 1, January 1985, pp. 47-55

2. **Study of the Effects of Yaw and Pitch Errors on Doppler Spectrum Parameters for a Spaceborne SAR**
By Tapan Misra, A M Jha and N S Pillai; SAC/MRSP/TR/01 January 1987

MILLIMETER-WAVE SENSOR IMAGE ENHANCEMENT

William J. Wilson and Helmut Suess*

Jet Propulsion Laboratory
California Institute of Technology
Mail Stop 168-327
4800 Oak Grove Drive
Pasadena, California 91109
Tel (818) 354-5699
FAX (818) 354-3437

ABSTRACT

Images of an airborne, scanning radiometer operating at a frequency of 98 GHz, have been analyzed. The mm-wave images were obtained in 1985/1986 using the JPL mm-wave imaging sensor. The goal of this study was to enhance the information content of these images and make their interpretation easier for human analysis. In this study, a visual interpretative approach was used for information extraction from the images. This included application of nonlinear transform techniques for noise reduction and for color, contrast and edge enhancement. Results of the techniques on selected mm-wave images will be presented.

INTRODUCTION

Images of an airborne, scanning radiometer operating at a frequency of 98 GHz, have been analyzed. The mm-wave images were obtained in 1985/1986 using the JPL mm-wave imaging sensor. This sensor has a 0.5° beamwidth antenna which is mechanically scanned ±22° across the flight track. The bean is stabilized to correct for aircraft movement due to turbulence and uses a butterfly scanning pattern to produce a linear raster scan pattern on the ground. This feature allows the mm-wave image to be displayed in real-time. The operational frequency of 98 GHz was chosen to give the smallest spatial resolution with the best poor weather performance. At an altitude of 750 m, a ground resolution of 7 m was obtained with a scan width of 620 m.

Passive millimeter-wave sensors are very useful for remote sensing because they provide information unique to this wavelength. Since the mm-wave emission is directly proportional to the emissivity of the object, the images can show very large contrasts. They also provide information during conditions of foggy and cloudy weather when visual and IR sensors are unusable. One of the main advantages of the passive mm-wave imager, is that the interpretation of these images by human operators is similar to the interpretation of visual and IR images because of the high correlation between the intensities of adjacent points. This is in contrast to radar images which are complicated by speckle, polarization and edge backscatter effects. However, one of the limitations of the passive mm-wave sensor is that it does not have the fine spatial resolution of visual and IR imagers. The goal of this study was to enhance the information content of these mm-wave images to help in their interpretation. Image processing techniques which were found to be most useful, especially for real-time analysis, were the choice of the false color map, smoothing for noise reduction, nonlinear modification of the brightness temperature histogram for contrast enhancement, and edge enhancement for detection of small features. The display of the original mm-wave data along with the contrast enhanced and the edge detection images appears to be a good combination to assist an observer in the interpretation of the data. Allowing the observer to quickly adjust the color scale and gain factors, using their intuition to bring out the details, is also very helpful.

FALSE COLOR ENHANCEMENT

The motivation for using color in image processing is the fact that the human eye can discern thousands of color shades and intensities. This is in sharp contrast with the eye's relatively poor performance with gray levels. Only about 16 shades of gray are detectable at any one point in an image by the average observer. Therefore, a color display has the capacity to convey much more information than an achromatic display and has the potential for improving the interpretation of images.

A careful selection of the false color scale for the brightness temperature range is an important item in the interpretation of the mm-wave image. Letting the observer adjust the color scale and temperature range in real-time, gives excellent results in this analysis. One way to think of this is that each observer has a matched filter in his brain to process information, and being able to

*Dr. Suess was a 1987-1988 National Research Council Resident Research Associate at the Jet Propulsion Laboratory from the Institute for High Frequency Technology (DFVLR), 8031 Oberpfaffenhofen, W. GERMANY.

change the color scale to match this filter, will maximize the information transfer between the observer and the computer. Choosing the color scale to appear similar to visual images seems to be one of the best choices to give the observer the best information transfer.

The key to selecting the best color scale is the ability to do this quickly and easily. In our current data reduction system, the color scale offset (the position of the color map on the brightness temperature scale) and the gain (the spread of colors vs brightness temperature) are quickly adjusted with a track ball to give the best image. Thus in a few minutes the best colors can be selected by each observer, and only minor changes are required for different images.

SMOOTHING

Smoothing of the image data reduces the random noise and other spurious effects that may be present in an image as a result of sampling, quantization or disturbances during image acquisition. These algorithms provide a better signal-to-noise ratio especially for distributed targets. However, the spatial resolution will be degraded and the edges are blurred. Three different smoothing algorithms were used on our mm-wave data: a moving average, a median filter in the spatial domain, and a low pass filter in the spatial frequency domain.

A comparison of these algorithms shows only small differences. Thus it was concluded that the median filter, which is easier to implement for real time processing, is the best smoothing filter to use in the mm-wave image processing, because of the better edge preserving properties.

BRIGHTNESS TEMPERATURE HISTOGRAM MODIFICATION

Techniques of image enhancement have the goal to improve the visual appearance of an image and to convert the image to a form more amenable to human and machine analysis. Unfortunately, at the present time there exists no unifying theory of image enhancement since it is difficult to define a criterion characterizing a "good quality" image. One possible reason lies in the multiplicity of application fields. In our case, a technique which was found very useful, was to equalize the brightness temperature histogram equalization or contrast enhancement (Pratt, 1978). This method is described in Fig. 1. The grey value histogram of the original image, which may have many of its pixels in a narrow temperature range, is mapped into a histogram with a nearly uniform distribution across the entire temperature range. After this modification, the cumulative distribution would be nearly a straight line. This means that an initial grey level interval with a high population will cause a larger stretch of the output grey levels than an interval with a lower population.

The processing enhances larger areas while suppressing smaller objects. This resultant image is especially useful for enhancing distributed targets and identifying large scale geographic features such as types of vegetation, roads, and hills. However, this technique suppresses small targets with only a few pixels.

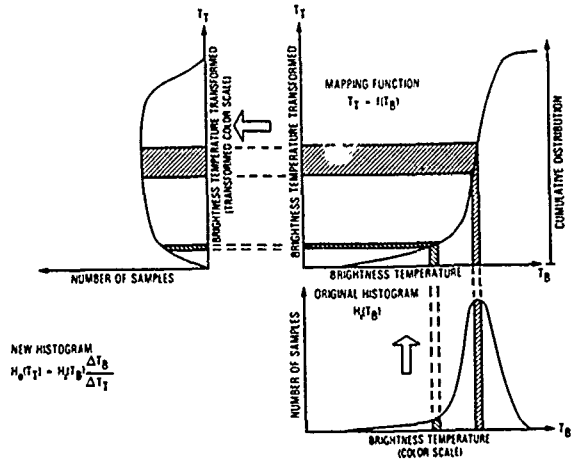


Fig. 1. Diagram explaining the brightness temperature histogram, H(T), modification for the contrast enhancement algorithm.

EDGE DETECTION

In mm-wave images, boundaries of objects appear as brightness temperature or intensity discontinuities. For the human observer, an object can often be recognized from only a crude outline. To enhance the small objects and edges with high brightness temperature contrasts in the mm-wave image, special edge enhancement algorithms are used. Two algorithms from the main class of gradient and template matching operators tried on this mm-wave data, were the Laplace and the Sobel operators. These operators reduce the temperature of all the pixels, except those on the border of a temperature discontinuity. The temperature increase at the object is proportional to the degree of the transition. Thus cold spots turn into hot doughnuts. The advantage of this algorithm is that it enhances small details that may not have been obvious on a uniform background.

ACKNOWLEDGEMENTS

This work was carried out by the Jet Propulsion Laboratory, California Institute of Technology and was supported by the National Aeronautics and Space Administration and the U.S. Army's Harry Diamond Laboratory.

BIBLIOGRAPHY

Pratt, W. K., Digital Image Processing, New York: Wiley-Interscience, 1978.

APPLICATION OF REAL TIME DIGITAL IMAGE ANALYSIS AND ENHANCEMENT TO AIRBORNE INFRARED FIRE DETECTION AND MAPPING

W.D. McColl, L.E. Neily
Innotech Avaition Enterprises Limited
Remote Sensing Division
6 Gurdwara Road
Suite 102
Nepean, Ontario
K2E 8A3

R.A. Neville
Canada Centre for Remote Sensing
Data Acquisition Division
2464 Sheffield Road
Ottawa, Ontario
K1A 0Y7

Airborne infrared imaging has been employed for forest fire detection and mapping for some time. Infrared remote sensing, sensitive to thermal emitted radiation potentially provides detection and quantifiction of fire targets that may not be visually apparent or that are obscurred by smoke.

Cost effective fire mapping requires the timely detection, location and quantification of fire targets ranging from smoldering hot spots to open flame fire fronts. Additionally these data must be recorded and transmitted to the fire manager efficiently to provide effective fire treatment.

Currently many fire mapping systems using primarily analog techniques are limited both in detection and in the medium to record and transmit data to the fire manager. Typically single or dual wavelength infrared sensors image the background terrain and the fire targets to a real time monochrome hard copy. The hard copy which is used for both navigation, by terrain feature matching, and fire mapping is limited in dynamic range with respect to the detectors for quantification of the fire targets. Where dual wavelength sensors are used generally the short wavelength detector is set such that a signal above a threshold level identifies a potential fire anomally. Detection is limited by the background thermal variation to targets with emission signals above this level. Spectral processing to reduce this background noise is not attempted.

Innotech Aviation, Remote Sensing Division has applied airborne digital imaging and real time image analysis and enhancements to commercial fire detection and mapping. The application of this technology, developed by the Canada Centre for Remote Sensing, has been demonstrated by Innotech under contract to the U.S. Forestry Service during September of 1988. Innotech used its Falcon 20 aircraft fitted with the standard sensor compliment to provide seventeen missions in Idaho, Washington and Oregon.

The primary sensor for fire mapping applications is a modified Daedalus 1260 MSS using IR detectors at 4 and 10 micrometres wavelength. Auxillary systems include the Alice II real time image processor and display, an inertial navigation system, MAID data processor and a high density digital tape recorder (HDDT). The 8 bit digital dual band infrared images with navigation data overlay are recorded on high density tape, output to the Alice II display and recorded on video cassette. This format provides for access to the full spatial resolution of the data and the digital sensor dynamic range by using standard image enhancement techniques of contrast stretch, histogram equalization, frame freeze and zoom enlargement. Colour image data are displayed on the aircraft high resolution colour monitor and recorded on the video cassette both during acquisition and during reproduction from the aircraft HDDT.

Colour encoding of the dual band imagery utilizes correlation of vegetation emissivities at 4 and 10 micometres to reduce detection limitations due to the background thermal noise and to provide a high contrast grey tone background image. The fire targets are presented in shades of red/pink through white depending on the thermal intensity. Colour separation is dependent on the differential increase in spectral black body emission at 4 vs. 10 micrometres for the higher temperature sources. The video navigation data overlay provides image location and scaling for accurate geographic plotting of the fire targets.

The technology provides enhanced commercial product easily interpretable in terms of hot spot/fire intensities on a high quality background terrain image.

UNSUPERVISED COASTLINE DETECTION AND TRACING IN SAR IMAGES

J. S. LEE AND I. JURKEVICH

Naval Research Laboratory, Washington, D.C. 20375-5000

ABSTRACT

Images of coastlines generated by Synthetic Aperture Radars (SAR) suffer from a number of deficiencies which arise from the presence of the speckle effect and the strong signal return from a wind roughened, wave modulated sea. The frequent lack of contrast caused by these effects makes coastline detection difficult by most conventional procedures such as, gray level thresholding or segmentation by edge detection. This paper describes an algorithm for the global detection of coastlines based on a sequence of basic image processing procedures and a new edge tracing algorithm. The application of the proposed procedure to SEASAT SAR and SIR-B (Shuttle Imaging Radar B) images demonstrates that it achieves good performance with only a modest computational burden. The computational burden can potentially be reduced to a level permitting a near real time performance in the sense of completing computations within a few frame times of typical image displays.

Keywords: Coastline detection, SAR, Edge tracing, Image segmentation.

I. INTRODUCTION

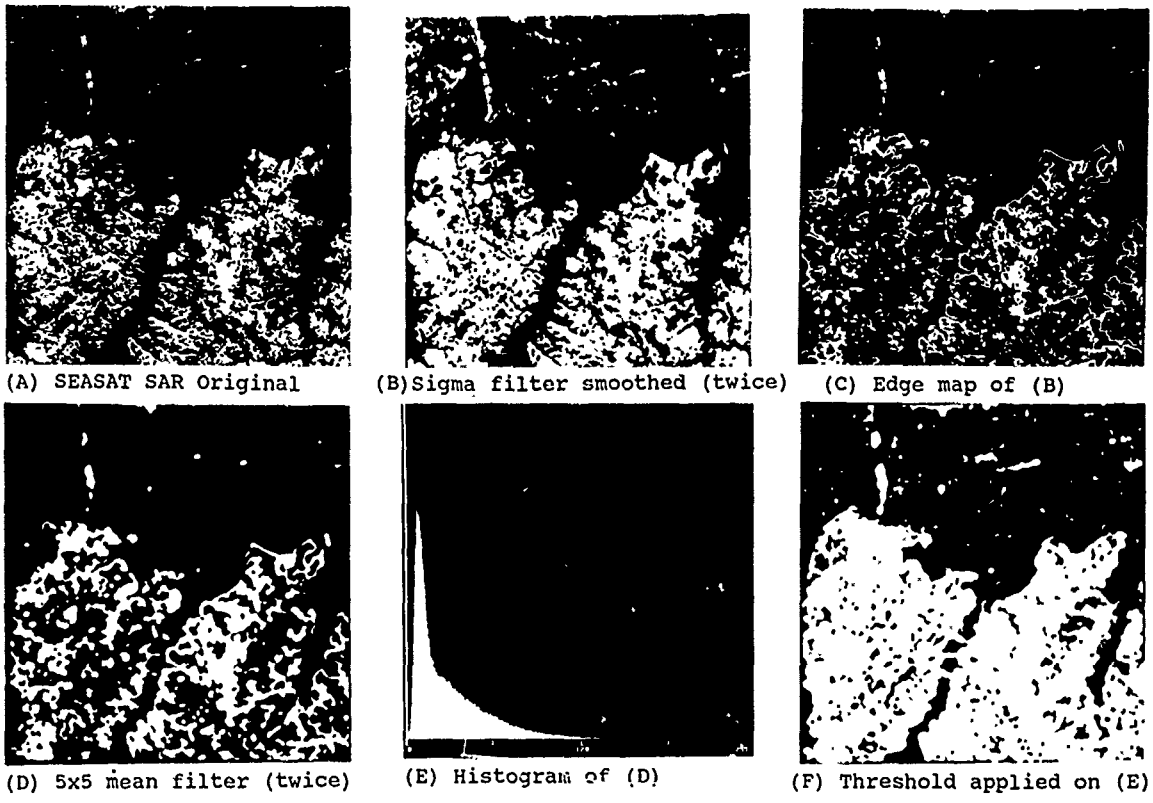
Unsupervised coastline detection in SAR images is an important step toward computer scene description. Knowledge of coastline is valuable in activities such as autonomous navigation, verification of radar platform's attitude and position, geolocation of targets (e.g. ships), etc. Coastline detection in SAR images is not, however, as simple a procedure as it is, for instance, in photographic or Landsat Thematic Mapper (TM) images. In the latter case, especially in Band 4 images, a simple act of thresholding followed by edge detection will effectively extract the coastline from the image. The difficulties in SAR images are associated with the nature of the signal return from the ocean and land areas. The return from the wind roughened and wave modulated water can frequently equal or exceed the return from a nearby land area, resulting in inadequate contrast for unambiguous separation. Furthermore, the presence of the speckle effect, generated by the coherent signal scattering within SAR resolution elements, complicates the detection problem. As a consequence, the coastline in many SAR images is so indistinct that even experienced observers have difficulties discerning it without the aid of a topographic map or other geographic knowledge.

Problems associated with coastline detection in SAR images have been considered in studies such as [1]. In the latter digital contour maps are used to simulate a SAR image, and then the coastline is defined by matching the simulated SAR image with the one under investigation. The purpose of this paper is to describe a simpler method for detecting coastlines with reasonable accuracy. Coastline detection is in the class of boundary detection problems. Similar problems have been encountered in areas such as, the determination of chest and heart boundaries in radiographs and cineangiograms [2] as well as object recognition in various remotely sensed images. The algorithms developed in these studies are problem dependent. Specific knowledge concerning the boundary is used to form rules which guide the grouping of pixels into boundaries. We have examined many SAR images containing coastlines and observe that the ocean areas, in general, are more homogeneous in gray level (fewer edge elements) than the land areas. The difficulty in using edge maps in defining a boundary is that the strong edges are not continuous. The boundary tracing by edges is an untractable programming problem which, even in a crude approximation, is very computationally intensive [3]. The local edge tracing in coastline detection must be guided by the global information about the coastline, that is, information extending over the whole image. Hence, the first step is to obtain a rough separation between the land and water. Refinement is then made in the neighborhood of the roughly defined boundaries.

The development of the coastline detection algorithm in this paper is covered in the following sections. A SEASAT SAR image of Chesapeake Bay area (Fig. 1 (A)) processed at four looks is used to illustrate the algorithm development. In section II a simple speckle smoothing algorithm [4] is applied to reduce speckle while retaining major edge features. A global ocean and land detection algorithm is described in Section III. Section IV is devoted to the rough coastline detection and the description of a contour tracing algorithm. The refinement of the detected coastline is carried out in Section V. Section VI is devoted to remarks and discussions.

II. PRE-PROCESSING BY SPECKLE REDUCTION

It is well known that speckle appearing in SAR images impedes the interpretation of the image either by a human operator or a computer. Several algorithms [4-6] based on the multiplicative noise model of the



(A) SEASAT SAR Original (B) Sigma filter smoothed (twice) (C) Edge map of (B)

(D) 5x5 mean filter (twice) (E) Histogram of (D) (F) Threshold applied on (E)

Fig. 1 SEASAT SAR image of Chesapeake Bay is used to illustrate the procedure of coastline detection

speckle effect have been developed to smooth the speckle noise without degrading the sharpness of the major edges in the image. The sigma filter [4] is selected for this purpose. The basic relation of this model is given by

$$Z_{i,j} = X_{i,j}\, v_{i,j} \quad \text{and} \quad v_{i,j} \sim (1, \sigma^2)$$

where $Z_{i,j}$ is the gray level of the observed SAR pixel, $x_{i,j}$ is its ideal or noise free counterpart and $v_{i,j}$ is the noise characterized by a distribution with mean=1 and variance, σ^2. It has been shown [5], [7] that this model fits the speckle distribution quite accurately. For 1-look amplitude SAR images. σ=0.52 and for 4-look amplitude images, σ=0.25.

The sigma filter used in this study employed a 5x5 or 7x7 moving window. A filtered pixel value is represented by an average of those pixels within the range of two standard deviations (two-sigma) of the center pixel. For SAR images (multiplicative noise), the two sigma range is given by

$$(Z_{i,j} - 2\sigma Z_{i,j},\ Z_{i,j} + 2\sigma Z_{i,j}).$$

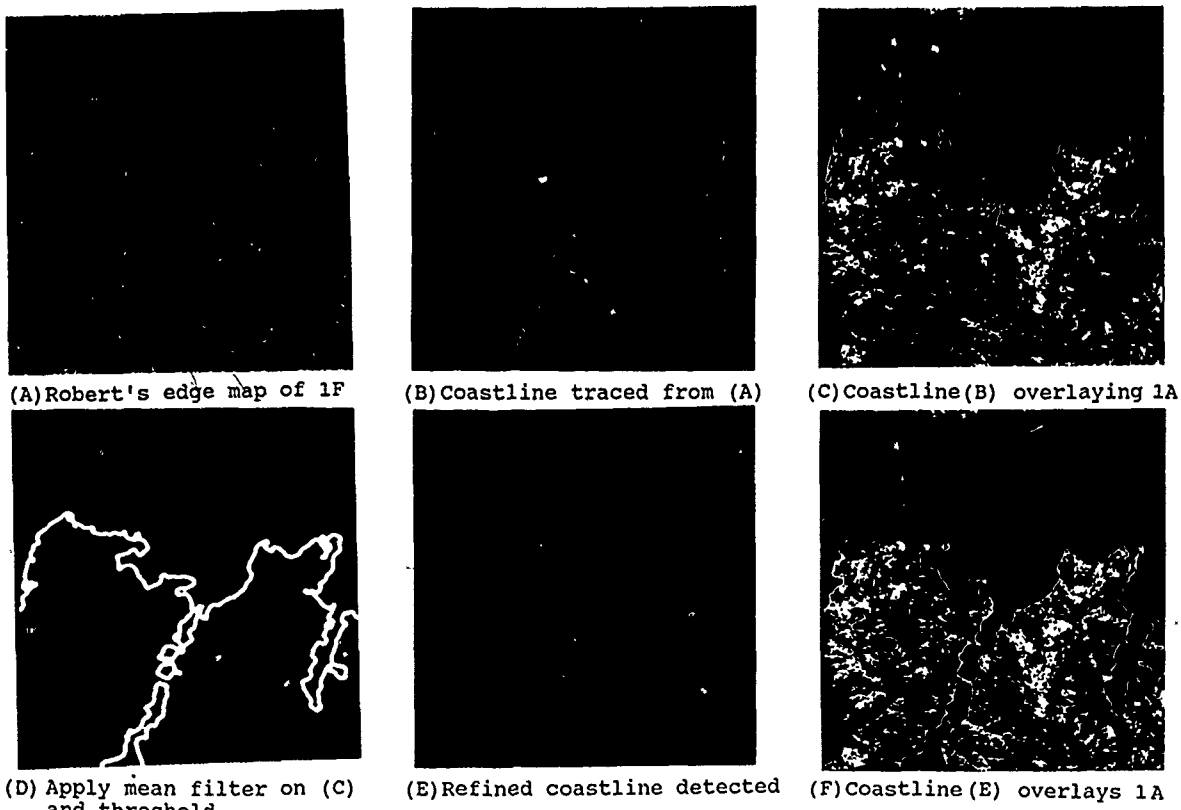
Details of this procedure are described in [4]. The result of smoothing the image twice with a 5x5 window is shown in Fig. 1(B). The speckle has been reduced considerably while coastline edges remain unaffected. It is impossible to segment either Fig. 1(A) or Fig. 1(B) by gray level thresholding, since many land areas have lower gray levels than some of the ocean areas. However, the edge map generated with the Sobel edge operator [9] (Fig. 1(C)) reveals great promise in separating the land from the sea.

III. SEA AND LAND DETECTION

To link the coastline edges displayed in Fig. 1(C) is not a simple task. As can be seen, gaps exist in coastline and some strong edges would mislead the tracing into the inland area. For this reason a simple procedure is implemented which generates an approximation to the land boundary. The edge map of Fig 1(C) is dilated by applying a 5x5 mean filter twice (Fig. 1(D)). The histogram of the latter image is then computed and displayed in Fig. 1(E). Assume at this point that at least 10% of the total number of pixels belong to the land or the sea. Since the sea area is more homogeneous than the land area, the corresponding pixels will peak around the lower gray levels. Furthermore, the averaging operation performed on Fig. 1(C) makes the sea pixels in Fig. 1(D) normally distributed. The mean and the standard deviation of the truncated normal distribution are estimated according to the procedure in [8]. A threshold is then set at (mean + 2x (standard deviation)). The result of this operation is shown in Fig. 1(F). The land and the sea can be readily distinguished by further eliminating the small blobs in both areas. In the next section, the coastline will be defined by a newly devised tracing algorithm.

IV PRELIMINARY COASTLINE TRACING

There exists a rich literature on various edge following techniques. For instance, the end-point method can be applied to the problem with low curvature edges. The Hough transform is suitable for the detection of straight edges or edges of a



(A) Robert's edge map of 1F (B) Coastline traced from (A) (C) Coastline (B) overlaying 1A

(D) Apply mean filter on (C) and threshold (E) Refined coastline detected (F) Coastline (E) overlays 1A

Fig. 2 Coastline tracing and refinement

particular shape. Under appropriate conditions even dynamic programming techniques can be used [8]. In the present application, since the land and sea are well separated as shown in Fig. 1(F), the following approach suggest itself. Data in Fig. 1(F) are operated upon by the Robert's edge operator and the result is shown in Fig. 2(A). The reason for applying Robert's operator [9] is that the edges generated are one pixel wide which makes edge tracing more precise. Since the edge map is obtained from the thresholded image, the coastline in the edge map will be continuous. The following procedure is then used to trace the coastline.

STEP 1.
Since the areas of interest in Fig. 1(F) are reasonably well defined one can proceed to scan the image starting in a sea area until a potential land pixel is located.

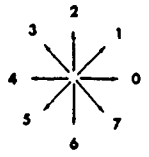


Fig. 3 Directional index

STEP 2.
From this pixel, find the 8-connected neighboring edge points and use one of the 8-connected pixels to define the starting tracing direction. Eight tracing directions are allowed as indicated in Fig. 3. The starting pixel and the edge pixel (current pixel), which define the tracing direction, are written to the output coastline image and the previous starting pixel is removed from further consideration.

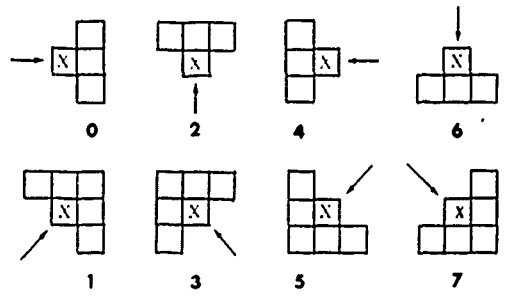


Fig. 4 Directional masks

STEP 3.
The next coastline pixel is chosen from the current pixel in the tracing direction. Only pixels associated with each direction index are considered as candidates. They are shown as square boxes in Fig. 4 where the pixel marked with x is the current pixel and (0,1,2,3,4,5,6,7,) are direction indices.
Ifthere is more than one candidate pixel available, this procedure will carry one step further to determine if a new edge pixel exists. If it does, the new coastline pixel is found and the tracing continues. Otherwise, the next candidate is tried.

It should be noted that due to contrast and resolution limitations of SAR images one can encounter various coastline features which cannot be handled with the assumed directional mask. Typically, loops smaller than the directional mask are in this category. As the algorithm proceeds along such a

feature a point is reached where no new candidate pixel is available, and the algorithm stalls. This problem is overcome by backtracing the coastline until it rejoins a major trend (Fig. 2 (A)).

STEP 4.
The tracing ends when the procedure encounters an image boundary. In such a case, the procedure returns to Step 2 to start the tracing in the other direction. If the tracing terminated at the original starting pixel, an island or a lake is detected.

STEP 5.
Large lakes or waterway inlets can also be traced if necessary.

The result of such computations is demonstrated in Fig. 2(B) where one inlet is traced. To gain an appreciation of the precision of the algorithm, as formulated to this point, we overlay the coastline with the original image, and show the result in Fig. 2(C). The detected coastline pixels are, on average, six-to-eight pixels away from the actual coastline pixels. Further refinement will be presented in the next section. However, for the purpose of geolocation, the procedure is sufficiently accurate even without this added step.

V. REFINEMENT

This section deals with the discrepancy between the detected coastline and the actual one. A simple algorithm is introduced which effectively decreases this error. The approximately defined coastline of Fig. 2(C) is dilated by applying a (5x5) mean filter twice and then thresholding the result to produce Fig. 2(D). The coastline is retraced with the algorithm of Section IV. However, only the inside edges are traced. For closed features, e.g. the lakes, interior to the coastline, the outside edge is traced. The result of this operation in shown in Fig. 2(E) and its overlay with the original image in Fig. 2(F). The refined coastline matches the actual one quite closely. The inlet in the right lower corner is not detected because of the bridge and the fact that this body of water is too narrow. Further refinements are possible due to the fact that backscattering from inlets and lakes are lower and more homogeneous than their surrounding land areas.

Other examples from SIR-B and SEASAT SAR images have been processed. The refined coastline traces show reasonably good fit based on visual inspection. This substantiates the development of this algorithm.

VI. REMARKS AND DISCUSSION

1. In our judgement, the refined coastline is accurate enough for the applications stated in Section I such as, autonomous navigation and geolocation. For geographic mapping, however, further processing is required. This involves tracing narrow inlets, small lakes, bridges, etc. The complexity of the algorithm would be increased greatly. Further refinement can be achieved by searching the pixel neighborhood in the direction normal to the detected coastline. The edge, the gray level variation and even the texture should be considered as attributes in determining the true water-land boundaries. If narrow inlets are to be detected, algorithms should be developed to trace them. In essence, the refinement should be guided by the results shown in Fig. 2(E).

2. The refinement of Section V not only improves the definition of the detected coastline, but also eliminates small fluctuations which yield false detection. A comparison of Fig. 2 (B) and Fig. 2 (E) will clearly reveal the effectiveness of this important step.

3. The algorithm developed in this paper uses many basic image processing operations, such as mean filter, Sobel edge operator, thresholding and histogram computation. With the currently available image processing hardware such as, VICOM, I^2S, etc., these operations can be performed in realtime. The sigma filter and the edge tracing scheme can likewise be efficiently implemented. Hence, this algorithm has the potential for the near real-time applications.

VII. CONCLUSION

An unsupervised algorithm has been developed for coastline detection in SAR images. To validate its performance it was applied to typical SEASAT SAR and SIR-B images. The algorithm performs reasonably well for geolocation and autonomous navigation. Further improvement in accuracy for the purpose of geographic mapping would require additional processing guided by the global information on the detected coastline. The algorithm is conceptually simple and computationaly efficient and has the potential for achieving real-time digital processing performance.

REFERENCES

[1] K. Eldhuset,, "Automatic Ship and Ship Wake Detection in Spaceborne SAR Images from Coastal Regions", Proceedings of IGARSS '88 p. 1529-1533, August 1988.

[2] C. Chow and T. Kaneko, "Automatic Boundary Detection of the Left Ventricle from Cineangiogram", Comp., and Bimod. Res, Vol. 5, pp. 388-410, 1972.

[3] D. Ballard and C. Brown, "Computer Vision", Printice Hall, 1982.

[4] J. S. Lee, "A Simple Speckle Smoothing Algorithm for Synthetic Aperture Radar Images", IEEE Trans. Syst. Man & Cybernetics, SMC-13(1), 85-89, 1983.

[5] J. S. Lee, "Speckle Analysis and Smoothing of Synthetic Aperture Radar Images," Computer Graphics and Image Processing 17, 24-32, (1981).

[6] V.S. Frost, et. al., "A Model of Radar Images and its Application to Adaptive Digital Filtering of Multiplicative Noise", IEEE Tran, PAMI, March 1982.

[7] J. S. Lee, "Speckle Suppression and Analysis for Synthetic Aperture Radar Images", Optical Engineering, p. 636-643. May 1986.

[8] J. S. Lee, and M. Yang, "Threshold Selection Using Estimates From a Truncated Normal Distribution," IEEE Tran. SMC, 1989 (in press).

[9] W. K. Pratt, "Digital Image Processing", Wiley, New York, 1978.

A LINEAR TRANSFORMATION TECHNIQUE FOR SPATIAL ENHANCEMENT OF MULTISPECTRAL IMAGES USING A HIGHER RESOLUTION DATA SET

K.V.Shettigara

Centre for Remote Sensing
Univ. of New South Wales
P.O.Box 1, Kensington, NSW 2033
Sydney Australia

ABSTRACT A new multivariate statistical process called Canonical Variable Substitution (CVS) technique is presented for enhancing the spatial resolution of multispectral bands using a second set of higher resolution data. The process involves a simple one-step linear transformation of the combined data space. The weights used in the linear transformation are scene dependent and are determined from canonical correlation analysis. The process is very general. An example of the enhancement of a SPOT multispectral image of Sydney using the panchromatic data is presented. The results are compared with that of the Intensity Hue Saturation (IHS) technique.

KEYWORDS resolution enhancement, SPOT images, canonical correlation, multiple correlation, intensity hue saturation

1 INTRODUCTION

The acquisition of higher resolution panchromatic (PS) data in the form of aerial photographs and satellite images has given rise to a host of procedures to enhance the spatial resolution of multispectral (XS) data. These procedures will become more popular with the launching of Landsat-6 in which the Enchanced Thematic Mapper (ETM) sensor will have a 15 m. PS band coregistered with XS bands.

The merging of higher resolution images is used to enhance the spatial resolution of the XS bands in a variety of combinations. Chavez (1984), and Cliche and Bonn (1985) have enhanced the SPOT XS bands of pixel resolution 20m using the PS band of 10m resolution. Welch (1984) has used the higher resolution SIR-A image to enhance Landsat images. Welch and Ehlers (1987) have merged SPOT PS data with Landsat-TM images. The high resolution aerial photographs were used for enhancing Landsat-TM data by Chavez (1986). The applications of merging procedures are increasing rapidly.

The known techniques for merging the higher resolution PS data with the lower resolution XS data utilize linear and/or non-linear weighted function of both the PS and XS data. The weights used in the published techniques are empirical. Clearly, there is a need to formalize the mathematical procedure involved in the merging exercise.

In this study a multivariate statistical process will be presented for enhancing the spatial resolution of XS bands. The process involves a simple one-step linear transformation of the combined data space.

2 GENERAL PRINCIPLES OF SPATIAL ENHANCEMENT TECHNIQUES

Welch and Ehlers (1987), and Cliche and Bonn (1985) have used both linear and non-linear functions to estimate the modified pixel values of the XS bands. Separate relationship is used for each band to compute the modified pixel values. Each relationship is a weighted function of the pixel value of the particular XS band and of the PS band. The function as well as the weights used are empirical in nature. However, some help in choosing the weights may be availed by studying the correlation matrix of the combined data set (Chavez et al.,1983)

In another approach Chavez et al. (1984) have used principal component analysis to merge the two data sets. The choice of the data bands is crucial for the success of this technique. The inclusion of the XS bands with very low correlation with the PS band can adversely affect the process. Again, the choice of the bands is based on the correlation matrix and personal judgement.

A popular linear process used for the merging purposes is the Intensity Hue Saturation (IHS) technique which will be described in more detail below. Welch and Ehlers (1987) have compared the results of the empirical linear/non-linear processes described above with the results of the IHS technique. They are of the opinion that the IHS technique performs better.

3 INTENSITY HUE SATURATION (IHS) TECHNIQUE

The IHS technique involves the rotation of the red, green and blue (RGB) color space, defined by the three XS bands of interest, to intensity, hue and saturation (IHS) space. The intensity component (variable) is then replaced by the PS band. The modified variable in the IHS space is then inverted back to the original RGB space (Welch and Ehlers, 1987). By this process the higher resolution spatial information of the PS band is incorporated in the XS bands.

In the above process it is implicit that the intensity axis in the IHS space has (statistical) properties similar to that of the PS band. Correlation coefficient is a good measure of similarity between two variables. It is shown in table II that the intensity axis is not necessarily the best linear combination of XS bands to achieve maximum similarity (correlation) with the PS band.

Apart from the above difficulty, the IHS technique has an additional constraint. The technique can deal with only three XS bands. If more than three bands are involved in the first set, like in Landsat-TM or MSS images, three bands have to be selected depending on the user's preference and judgement.

4 CANONICAL VARIABLE SUBSTITUTION (CVS) TECHNIQUE

The procedure presented here is discussed from the point of view of merging two data sets involving a XS data set and a PS data set of higher resolution. The procedure can be easily modified to allow multispectral images in the second set. The principle involved in the new technique is similar to that of the IHS technique described above. The main difference is in that the axis (variable) to be replaced by the PS band is computed by multiple correlation analysis. Considering that multiple correlation is a particular case of canonical correlation we will use canonical correlation in our discussion to keep the treatment as general as possible.

The procedure has three conceptual steps as shown in figure 1. In the first step (figure 1a) a linear combination of XS bands of the first set that would maximally correlate with the PS band is determined. Canonical correlation analysis is used to determine the required linear combination or the first canonical correlation vector. In the figure x_1,x_2 and x_3 are the original multispectral variables of set 1 and y_1,y_2 and y_3 are the transformed variables. y_1 is the first canonical variable. Axes y_1,y_2 and y_3 are mutually orthogonal. In practice y_2 and y_3 are never determined.

In the second step (figure 1b) the first canonical variable (y_1) is replaced by the PS band x_p, provided that y_1 accounts for a significant fraction of the information in x_p. Before replacing y_1 it is necessary to stretch and shift x_p so that it will have the variance and mean of y_1. This is to maintain the statistical properties of the original variables.

In the third step (figure 1c) the pixel values are inverse transformed to the original multispectral space. It will be shown in the next section that the above three conceptual steps can be accomplished in a single linear transformation.

5 MATHEMATICAL PROCEDURE

Consider a case of k XS bands and one PS band forming two subsets of a data set. In a general case if the second subset is also a multivariate data set then the first canonical variable of the second subset will replace the PS band in the following discussion. In the first step determine a canonical variable for subset 1 that would maximally correlate with the PS band.

Let x be the vector of dimension (k+1) consisting of k multispectral and 1 panchromatic values for a pixel.

$$\text{Let} \quad x' = (x'_s, x_p) \tag{1}$$

$$x'_s = (x_1, x_2, x_3 \ldots x_k) \tag{2}$$

Here x_s is the multispectral data vector. Let W be the matrix containing the orthonormal vectors for forward transformation along the columns. Vector w_1 in (3) is the first canonical correlation vector for determining the first canonical variable y_1 for subset 1 which would maximally correlate with x_p. Vectors $w_2, w_3 .. w_k$ in (3) are orthonormal to w_1, and are arbitrary.

$$W = (w_1, w_2, w_3, .. w_k) \tag{3}$$

If y_s is the transformed vector, we have

$$y_s = W' x_s \tag{4}$$

$$\text{where } y'_s = (y_1, y_2, y_3 .. y_k) \tag{5}$$

Assuming that the correlation between y_1 and x_p is satisfactory, for which some guidelines are suggested later, replace y_1 by x_p. Now the vector in (5) will be modified to:

$$y'_m = (x_p, y_2, y_3 .. y_k) \tag{6}$$

$$= (y_1 + \delta y_1, y_2, .. y_k) \tag{7}$$

$$= (y'_s + \Delta y'_s) \tag{8}$$

$$\text{where } \Delta y'_s = (\delta y, 0, 0, \ldots 0) \tag{9}$$

$$\text{and} \quad \delta y = (x_p - y_1)$$

$$= x_p - w'_1 x_s \tag{10}$$

Inverting back to the original space x_s:

$$x_m = W y_m$$

$$= W y_s + W \Delta y_s$$

$$= x_s + \delta y \, w_1$$

$$= x_s + (x_p - w'_1 x_s) \, w_1 \tag{11}$$

The vector x_m is the required vector which will have most of the higher resolution information of the PS band. As shown in (11), the conceptual forward and inverse transformations are achieved in one step. The only vector required for the whole procedure is the first canonical correlation vector w_1. As such, there is no need to determine the other orthonormal vectors $w_2, w_3 \ldots w_k$ to achieve the result.

If we choose $w_1' = (0.3333, 0.3333, 0.3333)$ or while normalised $(0.5774, 0.5774, 0.5774)$, y_1 in y_s would be the intensity value in the IHS space. In that case x_m in (11) would be the result from the IHS technique.

6 SIGNIFICANCE OF CANONICAL CORRELATION ANALYSIS

Canonical correlation analysis is extensively used in studying the relationship between two sets of multivariate data. The first canonical correlation coefficient provides a measure of maximum possible association between the first and the second set.

In the present study the second set is a univariate. In this case the square of canonical correlation also provides the fraction of information (variance) in the second set (PS band) that is explained by the first canonical variable of the first set. i.e. the XS bands. This follows from Stewart and Love's redundancy measure for the first canonical variable

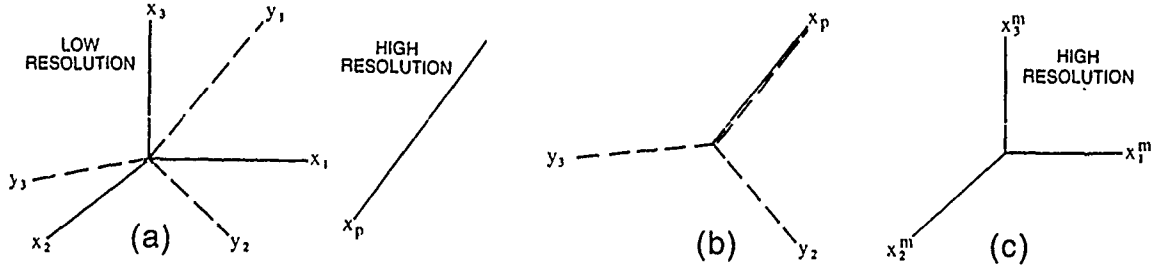


Figure 1 Geometric view of the CVS technique. (a) forward canonical correlation transformation. x_1,x_2 and x_3 are the original bands. x_p is the higher resolution panchromatic band. y_1 is the first canonical variable (b) replacement of y_1 by x_p. (c) inverse transformation of the variables in (b) to the original space.

of the second set (PS band itself in this case) given the first canonical vector of the first set. For more details see Green (1978; p 275).

The redundancy measure is given by

$$Rdu1|y1 = (v'1v1)\lambda 1/l \quad (12)$$

where $u1$ = first canonical variable of second set = x_p
$y1$ = first canonical variable of first set
$Rdu1|y1$ = redundancy measure for u1 given y1
$v1$ = first canonical vector for second set = (1.)
l = number of bands in second set
$\lambda 1$ = squared canonical correlation

In our case $\lambda 1$ directly gives the redundancy measure in (12). From the above relationship it is possible to decide if the merging process should be undertaken. At present there are no reliable rules to decide on the cutoff level on the redundancy measure. However, it is reasonable to assume that if $\lambda 1$ is less than $^{\wedge}$ 5 or the canonical correlation is less than 0.707 the merging process may not produce spectrally reliable results. This is irrespective of the merging process used.

7 ENHANCEMENT OF MULTISPECTRAL SPOT IMAGE OF SYDNEY AREA

In figures 2 b,c and d the three XS bands of a SPOT image of Sydney city area are shown. In these images the 20m pixels are replicated to 10m pixel size (resoln.tion 20m) and then coregistered with the PS band shown in figure 2a. Nearest neighbour algorithm is used for resampling. The main land covers in the image are parks covered with green grass and trees, water ponds and urban residential areas.

At the top left hand corner of the image is Centennial Park. The park has green grass and tree cover which appear very bright to moderate bright in XS3 band. In bands XS1 and XS2 they have low reflectance. The park has water ponds of different shapes which appear very dark in all the bands. The top and bottom right hand segments of the image are residential areas with mostly orthogonal roads. Concrete top buildings appear very bright in visible bands XS1 and XS2 while tiled roofs appear moderately bright. Roads have low reflectance in all the bands.

TABLE I IMAGE STATISTICS

band	XS1	XS2	XS3	PS
mean	50.7	38.9	66.4	52.9
	covariance matrix			
XS1	138.8			
XS2	120.6	133.4		
XS3	1.5	-27.7	416.0	
PS	145.7	152.0	-4.2	229.6
	correlation matrix			
XS1	1.000			
XS2	0.886	1.000		
XS3	0.004	-0.117	1.000	
PS	0.816	0.869	-0.013	1.000

In Table I the statistics of the combined set of XS and PS bands are presented. Notice that the visible bands and the PS band show a high degree of correlation. The infrared band is uncorrelated to other bands for all practical purposes.

8 RESULT AND DISCUSSION

Figure 3 shows the XS bands enhanced by merging the PS band using the IHS transformation technique. In table II the statistical characteristics of the intensity axis in the IHS transformation are shown. The intensity axis accounts for 42% of the total variance in the XS bands. However, the correlation between the intensity axis and the panchromatic axis is only 0.65. The intensity axis accounts for only 43% (0.65^2x100) of the information available in the PS band. A comparison of the enhanced XS bands with the original bands in figure 1 reveals significant enhancement achieved in all the bands, but at the expense of spectral properties.

TABLE II STATISTICS OF COMBINATION AXES

INTENSITY AXIS (IHS technique)	CAN.CORRELATION AXIS (CVS technique)
normalised vector coefficients	
(0.577,0.577,0.577)	(0.238,0.970,0.054)
variance accounted in XS space	
42%	27%
variance accounted in PS space	
43%	77%

In figure 4 the enhanced images using the canonical variable substitution (CVS) technique are presented. In table II the statistical characteristics of the canonical coefficients (vector) used for the computation of the first canonical variable are shown. Although the canonical correlation axis explains 77% of the information available in the PS band it accounts for only 27% of the multispectral variance. The low percentage of the variance accounted by this axis is easily understandable. More than 60% of the information in the XS space is concentrated in the infrared band (XS3) which is virtually uncorrelated to the PS band. The canonical coefficients reveal that the second band (XS2) mostly contributes to the canonical variable. This also means that the second band receives largest proportion of the high resolution information from the PS band followed by the first band XS1. The infrared band gets negligible contribution from the PS band. The images in figure 4 clearly show the above aspects.

The real quality of the merging process is dependent on both the resolution enhancement and the preservation of the spectral characteristics. The color composition of false color composites are a good measure to check if the spectral characteristics are preserved after the enhancement. The color aspects are found to be better preserved in the CVS transformation technique compared to the IHS technique. The sharpness of various features in the color composites from the IHS and the CVS techniques appear equally good.

One of the important aspects to be noted in table II is the relationship between the canonical weights and the correlation matrix. Although the visible bands XS1 and XS2 are equally correlated with the PS band (table I) their contribution to the canonical variable are totally different (table II). This is because the correlation coefficients reveal the relationship between the two variables, like between XS1 and PS. However, when one of the variables (XS1) in the relationship is a part of a multivariate data set the correlation coefficient does not reveal the variable's relative importance in prediciting the variable (PS) in the other set. For more information on this important property see Thorndike (1978). This points to the dangers of using the correlation matrix as a guide to empirically arrive at the weights for merging two data sets.

The CVS technique has many advantages ove r techniques. 1. The process involves a simple one ear transformation for which the weights are statistic determined. No empirical factors are involved. 2. T algorithm provides a statistical estimate to assess the suitability of merging two data sets. 3. The procedure is applicable even if the low resolution image set consists of more than 3 spectral bands. In contrast, the IHS technique requires that only 3 spectral bands are used. 4. The CVS process is general enough to utilize a higher resolution XS data set in the place of a PS band. In such a case the first canonical variable of the second set of bands replaces the PS band.

9 CONCLUSION

The canonical variable substituition technique is a simple mathematical process to enhance the spatial resolution of a multispectral image set using a higher resolution univariate or multivariate data set. The process enhances the resolution while preserving the original spectral characteristics. In this regard the CVS technique performs better than the commonly used IHS technique.

10 ACKNOWLEDGEMENT

The author is sponsored by the Dept. of Industries Technology and Commerce, Commonwealth of Australia. Thanks are due to Dr B.C.Forster for useful comments.

11 BIBLIOGRAPHY

1.Chavez,P.S.Jr.,"Digital merging of Landsat-TM and Digitized NHAP data for 1:24,000 scale image mapping", Photogrammetric Engineering and Remote Sensing,Vol.52,No.10,pp140-146,1986.

2.Chavez,P.S.Jr.,S.C.Guptill, and J.Bowell,"Image processing techniques for thematic mapper data",Proceedings:50th Annual ASP-ACSM Symposium,American Society of Photogrammetry,Washington D.C.,pp728-743,1984.

3.Chavez,P.S.Jr.,G.L.Berlin,and M.A.Tarabzouni, "Discriminating lithologies and surfacial deposits in the Al Hisma Plateau Region of Saudi Arabia with digitally combined Landsat MSS and SIR-A images", Proceedings: National Conference on Resource Management Applications: Energy and Environment, Vol.4,San Francisco,pp22-34,1983.

4.Cliche,G. and F.Bonn,"Integration of the SPOT panchromatic channel into its multispectral mode for image sharpness enhancement",Photogrammetric Engineering and Remote Sensing,Vol.51,No.3,pp311-316,1985.

5.Green,P.E.,"Analyzing multivariate data",The Dryden Press,Illinois,519p,1978.

6.Thorndike,R.M.,"Correlational procedures for research",Gardner Press,Inc.,N.Y., 340p,1978

7.Welch,R.,"Merging Landsat and SIR-A image data in digital form",Imaging Technology in Research and Development,pp.11-12,July 1984.

8.Welch,R. and M.Ehlers,"Merging multiresolution SPOT HRV and Landsat TM data",Photgrammmetric Engineering and Remote Sensing,Vol.53,No.3,pp.301-303,1987.

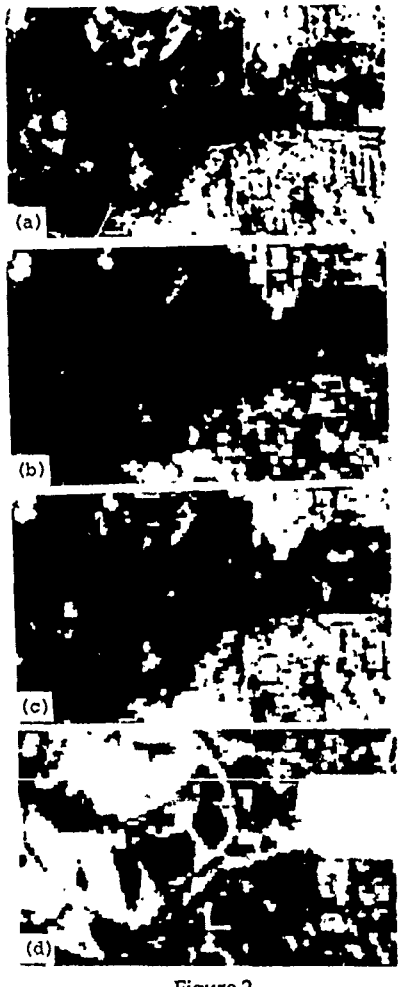


Figure 2
Original SPOT bands

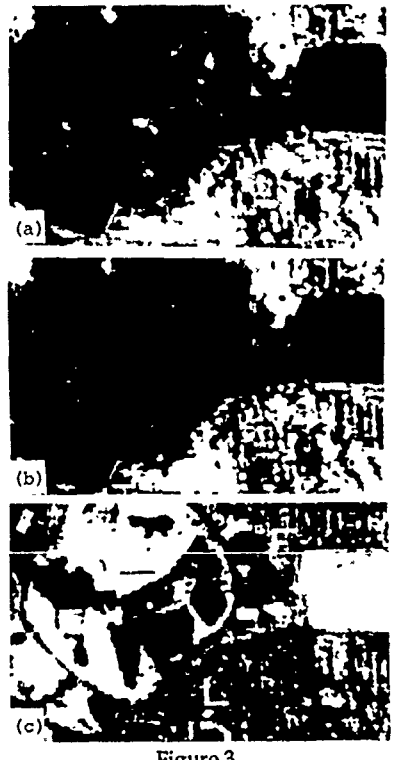


Figure 3
IHS enhanced bands

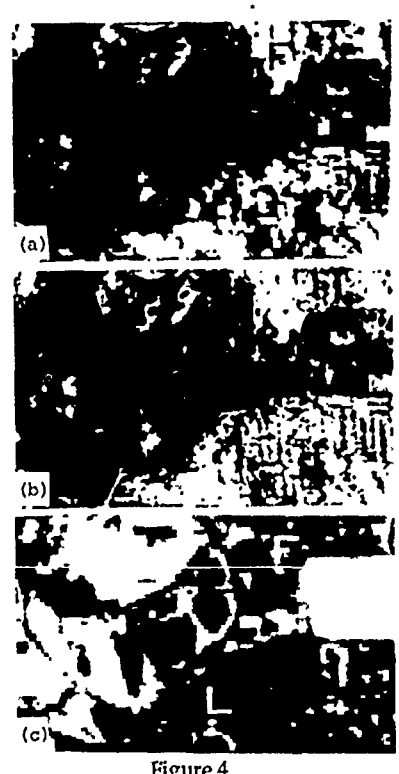


Figure 4
CVS enhanced bands

TM/LANDSAT DATA PROCESSING FOR INLAND WATER MONITORING

M. GODOY JR [1]; E.M.L.M.NOVO [1]

[1] INPE (INSTITUTO DE PESQUISAS ESPACIAIS) - São José dos Campos, São Paulo, Brasil - 12201

ABSTRACT

This paper presents results from an experiment performed to assess the performance of the median recurrent filter for stripping removal from TM/LANDSAT data. The experiment consisted of comparing the resulting images in terms of their potential for water and land use classification. This comparison was performed taking into account two types of remote sensor data user: the user interested in manual interpretation and the user interested in digital interpretation. As far as the potential for manual interpretation is concerned, the different images were compared in terms of the digital filtering effects on the elements of importance for the information extraction: color, texture, shape, contrast, etc. The image performance in terms of digital classification was assessed by comparing the classification matrix resulting from the application of a maximum likelihood algorithm to each data set: Reflectance Image; Reflectance Filtered Image; Composite Image. The performance of each data set was assessed through the percent of correct classification produced by the samples used to train the system. The results indicate that the filtered image allowed greater separability between optically distinct water classes, presenting however a poor performance for non-aquatic targets as far as digital analysis was concerned. For manual analysis, however, the composite image was the best.

1. INTRODUCTION

There are experimental evidences (Tassan and Sturm, 1986; Tassan, 1987, Tassan, 1988) that information on water quality such as color, turbidity, chlorophyll concentration, etc can be drawn from TM/LANDSAT data. For that, TM data must be subjected to various corrections for radiometric and atmospheric effects. Among radiometric corrections, special attention is given to the stripping removal, since this sort of noise is conspicuous in areas of complex boundaries between water and land surfaces (Shimoda et al., 1986).

Linear filtering procedures have been applied to remove this type of noise (Mehl et al., 1980). The main constraint to those procedures is that aberrant pixel values have strong effects on the output imagery. In this sense, non-linear filtering procedures, based on median-filter-type produce output images in which the aberrant digital values are not taken into account (Arce et al., 1986).

In this paper the median filter performance is assessed considering both: manual and digital interpretation procedures for land and aquatic ecosystems.

2. METHODS

To evaluate the suitability of applying the *median filter* for stripping removal a subscene of 512 by 512 pixel was selected over Barra Bonita Reservoir (Figure 1). This subcene was digitally processed using the SITIM-150 (Engespaço, 1988) according to the procedures presented on Figure 2)..

Three output images were produced: 1) Reflectance Image (RI), 2) Filtered Reflectance Image (FRI) and 3) Composite Image (CI). Those images were assessed in terms of both: suitability for manual and digital interpretation.

The suitability for manual interpretation was assessed by computing the number of features such as contrast, texture, etc., preserved in the output images. The suitability of each output image digital classification was assessed by implementing a maximum likelihood classification and by analyzing the resulting classification matrices.

3. RESULTS

The Reflectance Image proved to be unsuitable for water classes differentiation when manual interpretation is applied to the data. Stripping effects are quite conspicuous in the Reflectance Image preventing a clear definition of water class boundaries. As far as the surrounding terrestrial ecosystem is concerned, the Reflectance Image presented quite good performance because of the negligible stripping effects present in it. Several land use classes were easely mapped: natural pastureland, improved pastureland, permanent crops (sugar cane), temporary crops; reforestation stands; forest land, etc.

When the median recurrent filter (Figure 3) is applied to the Reflectance Image, the stripping is removed from the water producing sharp boundaries between water classes in the filtered Reflectance

Image. The filter, however, blurs the remaining scene making the output image unsuitable for manual interpretation of land use classes. The filtered Reflectance Image does not keep interpretation elements such as texture, boundaries between classes, etc. being only use ful for water classes differentiation.

The Composite Image is produced by digitally combining the Reflectance Image and the Filtered Reflectance Image in such a way that only water classes are subjected to filtering for stripping removal. Therefore the CI represents on excellent product being useful for both water classes and land use classes differentiation by manual interpretation of LANDSAT/TM data.

The performance of each output image when digital classification techniques were applied can be assessed through their classification matrices.

Table 1 presents the percent of correct classification for each water class and the average correct classification for land use classes as a whole. Table 1 clearly shows that the best image for land use and water digital classification is the Filtered Reflectance Image. The Reflectance Image presented the poorest performance for both land use and water classification. Since the same sample set was applied to train the system for the classification, the improvement in classification performance can only be related to the "stripping" removal.

TABLE 1

PERCENT CORRECT CLASSIFICATION

THRESHOLD 7-8

CLASSES	TYPE OF OUTPUT IMAGES		
	REFLECT. IMAGE	FILTERED REFLECT. IMAGE	COMPOSITE IMAGE
Water 1	39.9	99.2	99.2
Water 2	67.9	94.6	88.5
Water 3	47.9	96.7	96.8
Water 4	65.0	98.3	98.0
Water 5	83.2	99.2	92.3
Land use classes As a Whole	77.2	90.7	82.8

Table 2 presents the average value of pixels which were not classified for each data set and also the average confusion between classes. Those global results agree with those shown in Table 1.

TABLE 2

AVERAGE ABSTENTION AND AVERAGE CONFUSION FROM THE CLASSIFICIATION OF THE THREE IMAGE TYPES

IMAGE TYPE	RI	FRI	CI
Average abstention	2.74	5.20	3.9
Average confusion	20.38	0.79	8.7

Taking into account that the transformation of digital TM/LANDSAT image into a Reflectance Image aims to ease the acquisition of quantitative information on water reflectance, the filtering effect on the reflectance values was also assessed. Table 3 presents the reflectance values in the visible bands for all the output images. The data are already converted to apparent reflectance.

TABLE 3

PERCENT OF APPARENT REFLECTANCE OF DIFFERENT WATER CLASSES IN THE OUTPUT IMAGES

OUTPUT IMAGES	RI			FRI			CI		
CLASSES	TM1	TM2	TM3	TM1	TM2	TM3	TM1	TM2	TM3
Water 1	9.7	8.2	6.5	9.7	8.2	6.6	9.7	8.1	6.6
Water 2	9.9	8.7	6.6	10.0	8.9	6.6	10.0	8.9	6.6
Water 3	9.8	8.0	6.0	9.7	8.1	6.2	9.7	8.2	6.2
Water 4	9.6	7.9	5.9	9.6	7.8	5.8	8.6	7.8	5.8
Water 5	8.8	6.7	5.2	8.6	6.4	4.7	8.7	6.5	4.8

As shown in Table 3, the filtering process does not change the average reflectance of each water class. The highest difference between the reflectance percent derived from the three different images is found in the class "water 5". This differente can be explained by the highest percentage of commission error found in this class when the Reflectance Image was used to implement the digital classification. In fact, 9% of the pixels in this class are inclusions from land use classes. Owing the low water reflectance the red portion of the electromagnetic spectrum, the difference between the Reflectance Image and the remaining products in highlighted in that wave band.

The inclusion of pixels belonging to land use classes with higher reflectance explains the increase in the red reflectance in 05%. The blue and green wave bands having higher reflectance are less affected by this inclusion.

4. CONCLUSION

The results indicate that the filtered image allows a higher separability between optically distinct water classes, presenting however a poorer performance for non-aquatic as far as digital analysis is concerned. For manual analysis however, the composite image was the best. Results also proved that the filtering process does not change the apparent reflectance estimates of the different water classes.

BIBLIOGRAPHY

1. Arce, G.R.; Gallaghe, N.C.; Nodes, T.A. Median Filters: Theory for one-and-two dimensional filters. IN: (Huang, T.S., Editor): ADVANCES IN COMPUTER VISION AND IMAGE PROCESSING V2, p. 89-166, London, JAI Press, 1986.
2. Engespaço. SENSORIAMENTO REMOTO - SITIM-150, Manual do Usuário. São José dos Campos, 1988.
3. Mehl, W.; Sturm, B.; Melchior, N. Analysis of coastal zone colour scanner imagery over mediterranean coastal waters. INTERNATIONAL SYMPOSIUM ON REMOTE SENSING OF ENVIRONMENT, Proceedings. April 1980. São José, Costa Rica. ERIM. Ann Arbor, MI, V2, p.653-662, 1980.
4. Shimoda, H.; Elaya, M.; Sakata, T.; Gada, L.; Stelczer, K. Water quality monitoring of lake Balaton using Landsat Mss. SYMPOSIUM ON REMOTE SENSING FOR RESOURCE DEVELOPMENT. Proceedings. ENSCHEDE. ISPRS. p.745-758, 1986.
5. Tassan, S.; Sturm, B. An algorithm for the retrieval of sediment content in turbid coastal waters form CZCS data. INTERNATIONAL JOURNAL OF REMOTE SENSING, 7:643-656, 1986.
6. Tassan, S. Evaluation of the Thematic Mapper for Marine Application. INTERNATIONAL JOURNAL OF REMOTE SENSING, 8:1455-1478, 1987.
7. Tassan, S. The use of Thematic Mapper for coastal water analysis. INTERNATIONAL SYMPOSIUM OF PHOTOGRAMMETRY AND REMOTE SENSING, Proceedings, Kyoto, July, 1988, V7, p. 564-575, 1988.

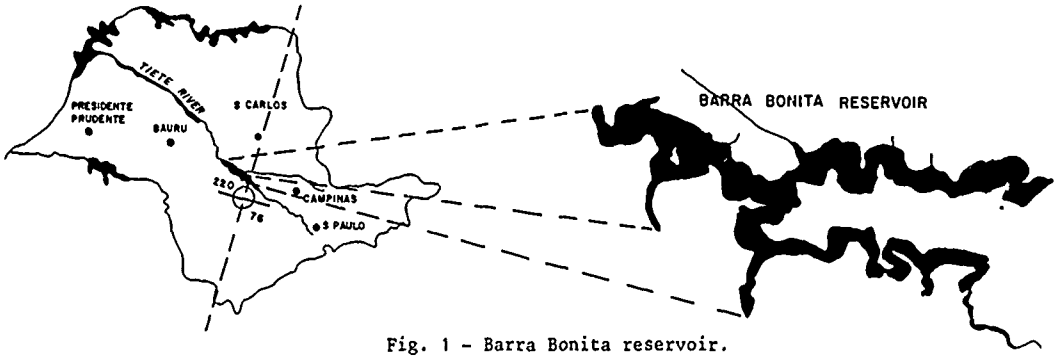


Fig. 1 - Barra Bonita reservoir.

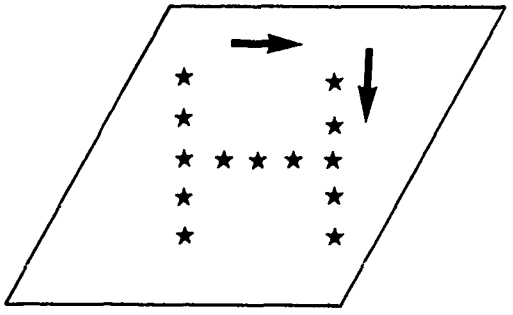


Fig. 3 - Median recurrent filter mask.

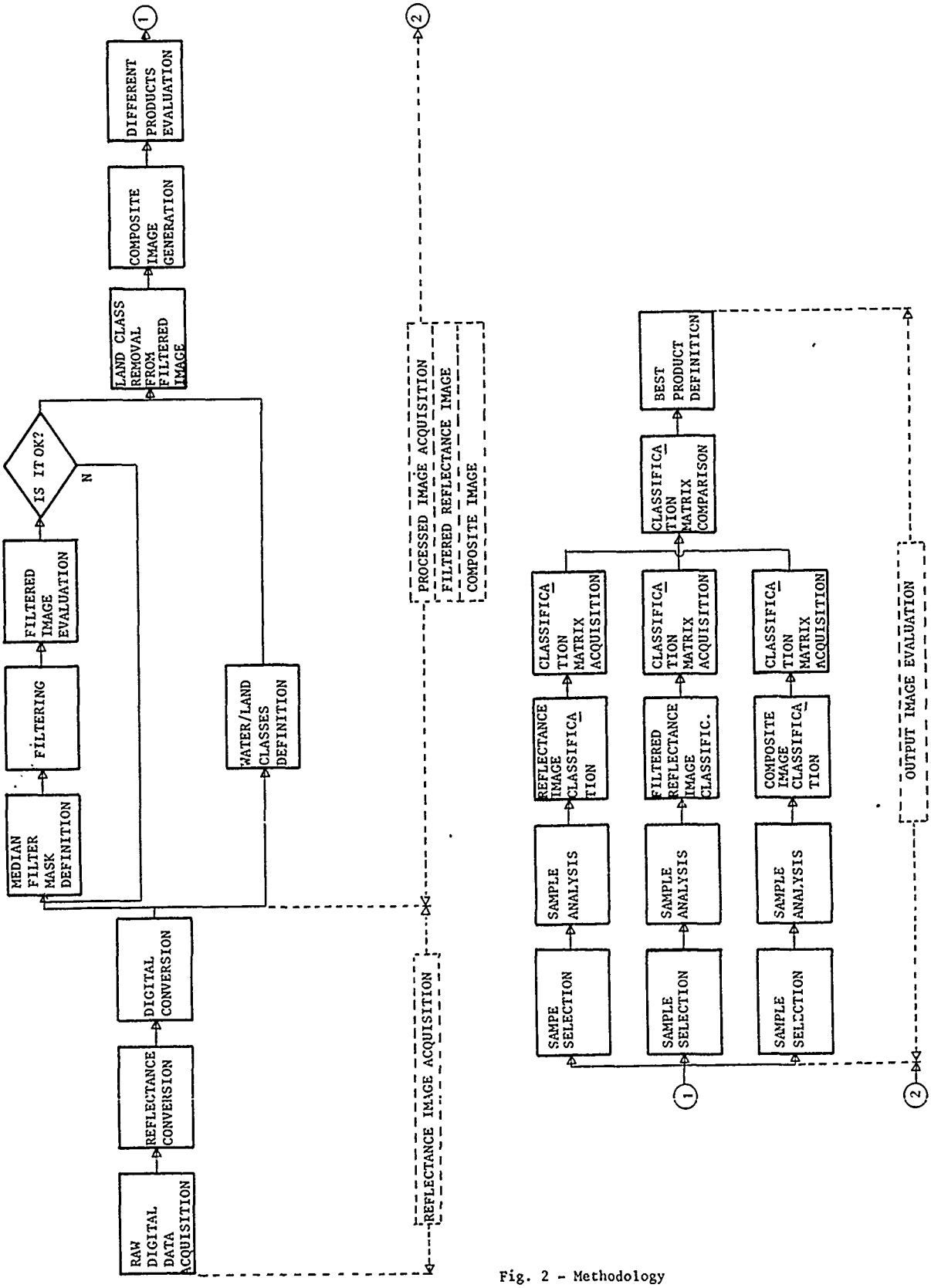


Fig. 2 - Methodology

REMOTE SENSING THE SUSPENDED SEDIMENT DISTRIBUTION ON THE MINGJIANG RIVER MOUTH IN DIFFERENT TIDE PHASES AS SHOWN BY MSS AND TM IMAGERY

Pan Delu & Lin Shouren

Second Institute of Oceanography
State Oceanic Administration,P.R.China
P.O.Box 507, Hangzhou

Abstract:

For studying distribution pattern of sediments on the Mingjing river mouth, Fujian Province China, series of Mss and Tm imageries were collected in different tide phases from 1975 to 1987. An algorithm for estimation of sediment concentration has been developed by in-situ spectral measurement with Mss and Tm bands. The values of concentration derived from satellite data are in agreement with the ship data collected near mouth whithen ±16% over the range 51 to 186 mg/m3. The diagrams of sediment fronts were drawn. The enhancement satellite imageries also shows four sediment tongues patteren on river mouth in different tide phases in great detail. Finally the machenism of sediment transportation on the mouth was discussed.

1. Introduction

Mingjing River discharges 57.5 billion m3 fresh water to the East-China Sea annually. Thousands and thousands boats nevigate in the inner river,as well as mang big ships pass through the river mouth and nevigate in the sea to Hongkong, Macao and other countries. So called 'TRANSPORTATION ARTERY' Mingjiang have been taking important play in economical development of Fujian Province, China. In the other hand,the river is carried about 10 Million tonne Sediments into the East-China Sea every year. Some of them was silted near mouth of the river to cause some problems such as obstruct of channel and ecologic change ect.The local government of Fujian is planning to regulate the river mouth for nevigation of 20,000 Tonnage steamer. Before beginning this plan, the first needs to know the sediment distribution patterns,transfer path and their quantitative concentration carried to the Sea. The second needs to understand the machenism, such as the relationship with tidal phase and others. But very few historical data about sediment tranbsportation are archived during different tide phases. The remote sensing was employed to be one kind of good tool for above task because of its timely and economic.

2.Imagery Series

In general, the chanel 1 of AVHRR data shows the sediment distribution pattern well. Unfortunately, Its spacial resolution (about 1.1*1.1 km/Piexl) is not good enough for Mingjiang River mouth. The width of the mouth inner is only several handred to two thousand meter. So the series of 10 landsat imageries (MSS 56*79m/pixel and TM 30*30m/pixel) were collected during September. 1975-july.1987, which presented in different tide phases,height, currents and wind, Listed in Table 1 and shown in Figure 1, eventhough the Landsat imageries archived in China for this area very limited.

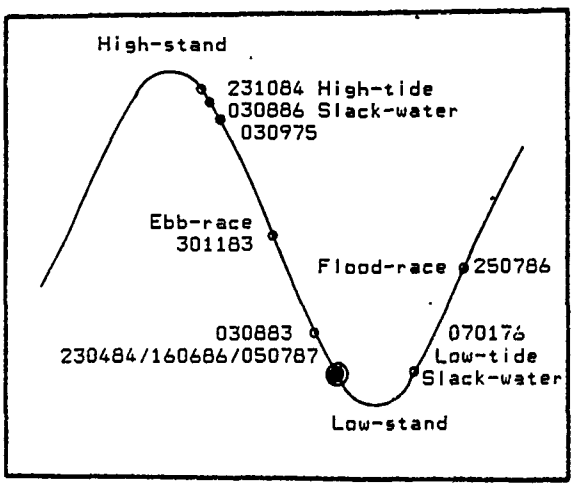


Fig1 Series of imagery relative to tidal phases

3. Method

3.1 Modelling

During june 14-16,1986, the investigation ship team collected in-situ data sycronously with Landsat 5 passing over the river mouth, such as sediment concentration as well as the spectral data with Mss and TM bands. The

model were deriven from in-situ data, responsed to the sediment concentration and spectral reflectance as following:

$$R(MSS5)=0.05058+0.0421LOG(S) \quad (1)$$

$$R(TM3)=0.506081+0.0812LOG(S) \quad (2)$$

In the Fomula (1) and (2), R(MSS5) and R(TM3) are the reflectance of MSS Band 5(0.6-0.7u) and of TM Band 3 (0.63-0.69u) measured at ship respectively; S is the sediment concentration. The correlation coefficient is 0.878 and 0.853 for Formula (1) and (2) respectively.

DATA DDMMYY	PATH/ROW	DATA	FLOOD-DRY SEASON	WIND (m/s)
030975	127/42	MSS	Dry	2.2 SW
070176	127/42	MSS	Dry	6.4 NW
030883	118/42	MSS	Dry	4.3 N
301183	119/42	MSS	Dry	5.3 N
230484	118/42	MSS	Wet	9.5 N
230184	119/42	MSS	Dry	6.4 NWN
160686	118/42	MSS	Wet	4.3 SW
250786	119/42	TM	Wet	0.0 C
030866	118/42	TM	Dry	6.4 N
050787	118/42	TM	Wet	1.0 SSE

Table 1. Series of imagery and hydrogical season and wind.

3.2 Atmospheric and geometric correction

Very simple and experiental method was employed to correct the atmospheric effective of imageries. Finding the black water supposed the sediment concentration Less than 0.5 mg/m3, calculated by Formula (1)or(2), then the following formula are used to atmospheric correction:

$$Aw(MSS5)=Ao(MSS5)-Abw(MSS6) \quad (3)$$

$$Aw(TM3)=Ao(TM3)-0.43Abw(TM4) \quad (4)$$

In Formula (3) and (4), Aw(MSS5) and Aw (TM3) are the corrected radiation of MSS Band 5 and TM Band 3 respectively. Ao(MSS5) and Ao(TM3) are the original radiation of MSS Band 5 and TM Band 3. Abw (MSS6) and Abw (TM4) are the radiance of black water with MSS Band 6 (0.7-0.8u) and TM BAND 6 (0.76-0.90u) respectively.

Seven geometric control points were applied in the polynomial enpression for geometric correction. The precision is about 1-2 pixels in the imager.

3.3 Quantitative mapping

Computing reflectance from corrected radiance, then mapping Sediment concentration by Formula (1) or (2), the example shown in Figure 2. The in-situ sediment concentration data collected from ship in june 14-16,1986 was used to text the mapping data. The value of concentration derived from the satellite data are in agreement with the ship data collected near mouth withen ± 16% over the range 51 to 186 mg/m3.

Fig2. The sediment concentration Mapping from TM AUG.3,1986 and enhancement.

3.4 Imager enhancement

The wallis transformation was applied to performs a space variant constrast stretch to normalize the MSS and TM image. the computational algorithms has the following form:

$$P'(x,y)=[S'/(S+S'/MAX)]*[P(x,y)-M+A*M'+(1-A)*M] \quad (5)$$

Where:P(x,y)=original pixel intensity; P'(x,y)=normalized pixel intensity; S=localized standard deviation (local neighborhood defined by 41*41 pixel window); S'=desired standard deviation (76.8); M=local mean; M'=desired mean (128);Max=maximum gain(6); A=factor to govern mean value shifting (0.8).Then the transfered imagers of tree bands such as MSS(5,6,7) or TM(2,3,4), were pseudo-combined. the enhanced imagers show very great datail concentration pattern and sediment path line, example see Figure 3.

4. Discussion and results

4.1 Distribution patteren

From enhanced imageries, It is obvious that four turbid tongues torward sea existed from north to south on the mouth all seasons of year,called Wuzhu, Cuanshi, Waisha and Meihua tongue shown in Figure 4. The pattern of tongue Wuzhu and Meihua are stable, but tongue Cuanshi and Waisha are varied in

differen seasons.Tho spread direction of all tongues are as same as the directiun of abb current. Their shape and spread area varied

Fig.3 The MSS (Band 5,6,7) imagers enhanced by Wallis Transformation and Pseudo-combination

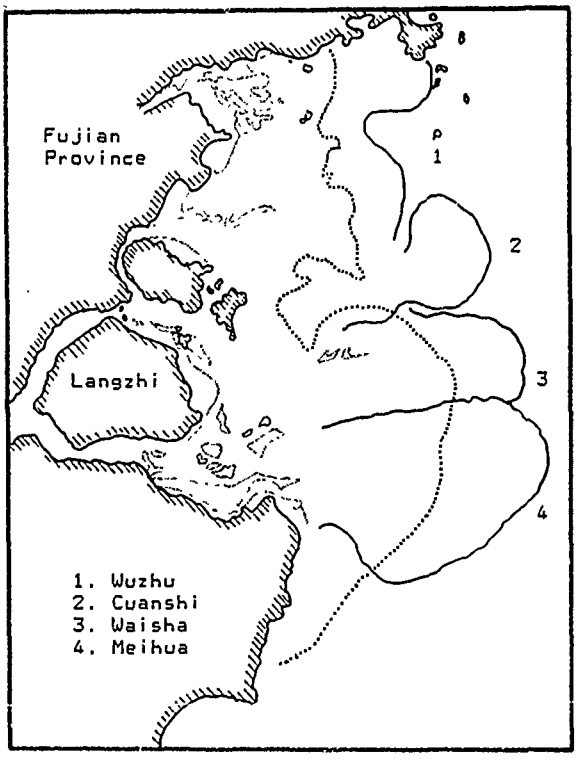


Fig4 Sediment distributiong pattern at high and low stands from imageries

with tidal level.(a) Durins inetiation of abb on the imageries (such as 030886 TM), The turbid water shows multi-layer radiant fan. The contour creek is not obvious,but the structure is well-developed like breard shown the spread direction fine. (b) During strong abb-slack water,(such as 050787 TM), The contour creek become obvious. The jet of different sedment concentration crossed each other.The obvious path line is focussing and scattering and trajectory shows the axle of tongues and different channel current very well. The imageries also show that the turbid waters run off from the constrain by the topography of creeks and scatter radiantly forward the sea to formalize obvious turbid water tongues.

4.2 Sediment transportation and concentration

Based on the path of sediment transportation from imagery, there are four different areas shown in Figure 5. (a) Sediment active area located on area I, near the mouth run-off and tidal current are as main motivation. the suspended sediment moves back and forward repeatly. The direction of movement is parallel to coast approximately. (b)Sediment diffusion eros-silt area located on the area 2. The sediment was carried by tidal current. The main motivation is tidal current, also dependent on topography construction. because of erosion and siltation, the turbid water shows different shape such as plume, cyclone, and jet, Usually the sediment concentration is much higher then others. (c) Sediment diffusion and diposition area located area 3, motivation is tidal current, also affected by run-off,wave and near water system. In this area as the turbid waters run off from the beach off mouth and scatter forward the sea.the sediment diffused and deposited. (d) Sediment drift area Located area 4. The low concentration water departed from turbid water and the sediment drifted and deposited with tidal and coastal currents and branch of Krushio.

Based on the sediment concentration, the mouth can be divided 3 areas showing in Figure 5. (a) High concentration area H: In this area,the effect of tidal current and shallows is strong, So, the turbid water well-mixed. The maximum value evaluated from imageries is 201.9 g/m3 in the dry year ,183,5 g/m3 in the wet-year. The minimum value is 46.5g/m3 at High level. (b) Midium concentration area M: The concentration is strongly affected by topography and tidal height, the maximum 337.1 g/m3 at Lower level in dry season and the minimum 6.1 g/m3 at higher level in wet season. (c) Low concentration arearl located diffusion-deposition area, the concentration is stable. In general, the value is about 56.0-44.1 g/m3. the maximum 80.2 g/m3 in occasional case of wind above grade 5 at lower level.

4.3 Mechanism

Because the Sediment is yielded of drainage basin and controled by tidal current.The mechanism of sediment transportation was found by comparing Mss and TM imagery with tidal data, The turbid water pattern, spread area, shape, direction, and concentration

distribution firstly dependent on tidal current. Secondly dependent on run-off, usually, the concentration is higher in dry season then at in the wet season. Thirdly dependent on wind wave and near water system, when the wind was above grade 5 the obvious pattern begain to be obscure. During june

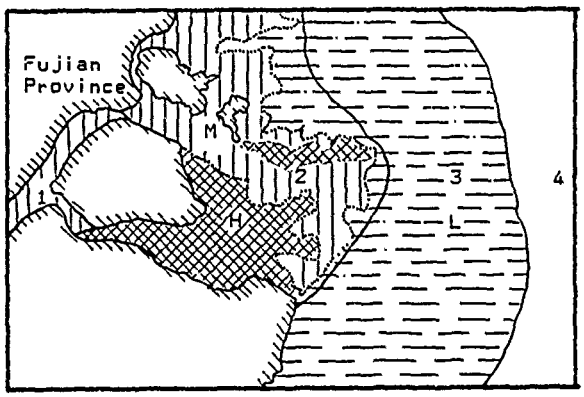


Fig 5. Sediment transportation areas and concentration distribution

to September, the branch of krushio up to north is main near water system, the krushio water usually is very Low concentration water,while from October to next April,main near water system is controled by MingZhe coastal current down to south, whose concentration is little higher then krushio. Both of them did not affect the sediment distribution much, but some time, they would make the turbid tongues to deviate from normal.

5. Suggestion of regulation

A suggestion for regution of nevigation channel has been proposed by Navigation Port Administration of Fujian Province, based on above analysis and other hydrological in-situ data. Its principal ideal is strengthening main current-Cuansn to rush the beach off mouth which obstruct the main nevigation channel with building several dams shown in Figure 6. The Dam 1 ajusted the water transportion radio of Meihua and Changem current. Dam 2 obstructs high sediment concentration water rushing into main water way. Dam 3,4 (two locking dam) cut-off branch and strengthen main channel. Dam 5 seperates the main channel and Meihua channel whose water is highest concentration.

Acknowledgement

The authors would like to acknowlege the direct and indirect contribution made by numerous participants, special thanks go to Mr. Dou Yawei and Mr. Lin Minji from the Third Institute of Oceanography, S.O.A, China for providing the final report on this project, Also go to Mr. Zhao Taichu and Zhou Jianru ect. from Second Institute of Oceanography for collecting in situ data and some analysis. Finally thanks go to investigation ship team and Image Processing Lab. of Zhejiang university for providing ships and Image processing system respectively.

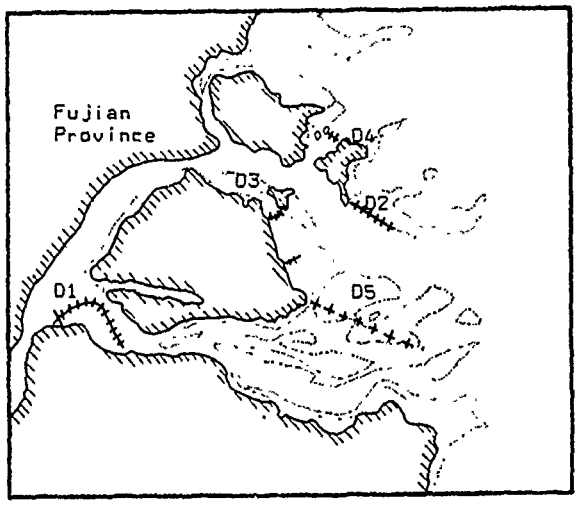


Fig.6 Diagram for suggestion of regulation on the Mingjiang river mouth

Reference

(1) Gordon, H.R., and D.K.. Clark, "Clear water radiances for atmospheric correction of Coastal Zone Color Scanner imagery",Applied optics,20 PP4175-4180, 1981.

(2) Jonnso,R.W."Quatitative sediment mapping from remotely sensed multispectral data," Remote sensing of resources,IV,PP 556-576, 1975

(3) Mccauley, J.R.,ect."Kansas water quality using ERTS-1," Remote sensing of earth resources, IV, 522-541. 1975.

(4) Pti.F,"Optinianl distribution of control points to minimize landsat image registtration errors,"Photo grammetric Engineering and Remote Sensing,PP101-110.

(5) Siamak Knorram, " Remote sensing of salinity in the San Francisco Bay Delta," Remote sensing of environment, 12,1,PP15-22,1982.

(6) Viollier,M.," Radiometric calibration of the Coastal Zone Color Scanner on Nimbus-7: a proposed adjustment,"Applied Optics,21,PP1142-1145, 1982.

Geometric and Radiometric Correction of Airborne MSS Imageries over Rugged Terrain Area

N. Y. Chen and H. T. Wang
Center for Space and Remote Sensing Research
National Central University
Chung-Li, Taiwan, R. O. C.
Tel:886-3-4257232, Fax:886-3-4254908

There are three major factors for the distortion of Airborne MSS imagery, namely, the irregularity of scanner's gyroscope, six orientation parameters being time dependent, and the terrain effect over the rugged areas due to the large scanning angle. We use the cross correlation between adjacent lines to find the relative shift of each line. Care has been taken into account the fact that there are three components of the shift, i.e., slow shift caused by aircraft motion, rapid shift caused by the gyroscope and the small random shift caused by ground feature variations. We assume that six orientation parameters are second or third order polynomial with respect to time and use control points to determine the coefficients. We finally use DEM and ray tracing technique to produce the "orthographic" Airborne image. We apply the above procedure to some images over highly rugged area in Taiwan. The results indicate that for 2000 scan lines or less, the r.m.s. error of check points are about 2 pixels(10m), while the traditional method will produce a r.m.s. error of about 10 pixels. Considering the large grid size (40m) of the DEM we used, the above results are satisfactory.

Concerning radiometric correction, we assume that the radiance measured at the position $\vec{x}_v$ in the direction Ω_v can be approximated by

$$M(\vec{x}_v,\Omega_v) = [I_d(\vec{x}_s)+I_s(\vec{x}_s)+I_g(\vec{x}_s)]R(\vec{x}_s)T(z_v,z_s,\Omega)/\pi + P(z_v,z_s,\Omega_v)$$

Where $\vec{x}_s$ is the target position, $R(\vec{x}_s)$ the target reflectance, T the transmittance, P the path radiance, I_d the the direct solar irradiance, I_s the diffuse sky irradiance and I_g the direct solar irradiance reflected from the nearby ground pixels. Note those quantities are functions of position (terrain effect) and atmospheric optical properties. For the atmospheric effects we assume that the atmosphere can be described as a standard rural tropical model with parameter d, the meteorological range (detailed descriptions see Lowtran 6 Package). The meteorological range can be estimated by assuming that the energy of the darkest pixels only contain the path radiance, then by fitting the path radiance for different height and bands. Due to the lack of absolute radiometric calibration of the Daedulus scanner, we only perform the terrain parameter calculation and two simulations under the following condition: d=10 km, $R(\vec{x}_s)$ = constant =0.2, λ =0.85μm , flying height = 8000m, azimuth angle = 45° , constant speed, solar zenith angle = 45°, solar azimuth angle = 135°. The terrain parameters, i.e. the cross section ratio (the cosine of the incident light and surface normal), shadow and horizon etc., are calculated for each pixel using DEM. For test site 1 which is rolling terrain with average height about 500 m, the average (over pixels) ratio of path radiance contributions to the total measured radiance is 16% with the maximum up to 33%. The average I_d, I_s, and I_g contribution to the irradiance are about 30%, 67%, and 3% respectively. For the test site 2 which is rugged terrain with average height 1500m, the average path radiance to the measured radiance is 18% with maximum 96%. The average I_d, I_s, and I_g contributions to the irradiance are 65%, 22% and 13% respectively. This means (1) for the low altitude, the diffuse sky light can be very important (more important than direct solar irradiance even in near IR region). (2) for the valley pixels, the contribution I_g can not be ignored.

EVALUATION DU SATELLITE SPOT POUR L'EXTRACTION DES ELEMENTS PLANIMETRIQUES DES CARTES TOPOGRAPHIQUES AU 1: 50 000: RÉSULTATS PRÉLIMINAIRES

Charbonneau, L., Coulombe, A., Perras, S., Royer, A. et Bonn, F.

Centre d'applications et de recherches en télédétection (CARTEL), Université de Sherbrooke, Sherbrooke, Québec, J1K 2R1

RÉSUMÉ

Cette recherche vise l'évaluation du potentiel de l' imagerie SPOT pour l'extraction des données planimétriques des cartes topographiques au 1: 50 000. Il s'agit en particulier de mesurer l'apport des données SPOT dans le Programme canadien de révision cartographique. Le secteur d'étude de cette recherche est la région de Sherbrooke au Québec (carte topographique 21 E/5). Le fichier numérique des données planimétriques correspondant à l'édition 6 de la carte topographique servira de document de vérité de terrain. Une conversion du format vectoriel à un format matriciel sera réalisée afin de rendre les données planimétriques compatibles aux données satellitaires. L'analyse est effectuée à partir d'une image acquise en été en mode multibande (02-08-87), une image acquise en automne en mode multibande et en mode panchromatique (25-10-86) et une image acquise en hiver en mode panchromatique (21-02-87).

On présente les méthodes de rehaussement d'images utilisées pour améliorer la qualité visuelle des données satellitaires et leur intégration à la carte topographique. On discute d'autre part des techniques de détection automatique des changements pour les éléments planimétriques.

ABSTRACT

The main object of this research work is to evaluate the potential of SPOT imagery for extracting planimetric data from 1:50 000 scale topographical maps. More precisely, the aim is to determine the input of SPOT data in the framework of the Canadian map revision program. The area under consideration is located in the Sherbrooke region in Quebec (topographical map 21 E/5). The digital planimetric data files corresponding to edition 6 of the topographical map is used for ground truth purposes. A conversion of the planimetric data from the vector format to a raster format will be applied to make the data compatible with satellite data. The analysis is focused on the use of an image acquired during the summer in multiband mode (08-02-87), one acquired during the fall in multiband and panchromatic mode (10-25-86) and another image acquired during the winter in panchromatic mode (02-21-87).

We present the image enhancement methods used for increasing the visual quality of the satellite data and their integration into the topographical map. We also discuss the techniques used for the automatic determination of changes in the planimetric elements.

1. INTRODUCTION

Depuis le lancement du satellite SPOT, le développement des applications cartographiques à l'aide de données satellitaires connaît un essor considérable. Le Ministère de l'Énergie, des Mines et des Ressources du Canada utilise déjà les images satellitaires pour la révision de cartes topographiques. Entre autre, le ministère planifie l'acquisition des photographies aériennes dans les secteurs où les changements observés sont considérables et fréquents. Un programme de mise à jour des éléments cartographiques pour les cartes au 1: 50 000 et au 1: 250 000 est réalisé à partir des images LANDSAT-MSS et il est opérationel depuis plusieurs années. Cependant, cette révision est réalisée jusqu'à maintenant de façon manuelle en superposant l'image satellitaire à la carte topographique et en identifiant les changements survenus entre l'image et le document topographique (Gregory and Moore, 1986).

2. BUT ET OBJECTIFS

Le projet présenté vise l'évaluation du potentiel du satellite SPOT pour l'extraction des éléments planimétriques des cartes topographiques au 1: 50 000. Dans un premier temps, il s'agit de développer des méthodes de rehaussements d'images en vue d'améliorer la qualité visuelle des images satellitaires. D'autre part, le projet met l'accent sur la mise au point de techniques de détection des éléments planimétriques. L'objectif visé est de trouver la technique optimale pour la détection de chaque élément. Suite à cette analyse, une comparaison sera faite entre les fichiers graphiques (contenu thématique de la carte 21 E/5) et les meilleures méthodes.

3. DONNÉES SATELLITAIRES ET THÉMATIQUES

Le secteur d'étude correspond au feuillet Sherbrooke 21 E/5 de la carte topographique au 1: 50 000 situé en Estrie dans le sud du Québec. Il occupe une région comprise entre 45° 15' et 45° 30' de latitude nord et entre 71° 30' et 72° de longitude ouest. Ce secteur couvre une superficie de 1 086 km^2.

Nous avons utilisé la série suivante d'images du satellite SPOT:

- une image SPOT acquise en été dans le mode multibande(MLA) le 02/08/87;
- une image SPOT acquise en automne dans le mode multibande et une autre dans le mode panchromatique (PLA) le 25/10/86,
- une image SPOT acquise en hiver dans le mode panchromatique le 21/02/87.

4. MÉTHODOLOGIE

4.1 Prétraitement des données satellitaires

La majorité des traitements ont été réalisée au CARTEL à l'aide de notre système de traitement numérique d'images ARIES III de DIPIX (DIPIX, 1986). Toutes les images sont du niveau de base de prétraitement 1B (Bégni et *al.*, 1987) et elles ont été traitées à partir du système MOSAICS du Centre canadien de télédétection (EMR) Ce niveau comprend les corrections radiométriques et les corrections géométriques liées aux déformations systématiques introduites par les conditions de prise de vue (rotation de la terre, effet panoramique, effet de filé). Une correction radiométrique transforme les valeurs de niveaux de gris relatifs (DN) en luminance

apparente. La réflectance apparente (ρ*) au capteur a ensuite été calculée pour normaliser les valeurs de luminance par rapport aux conditions de prises de vue, différentes pour chaque image, et par rapport aux différences entre capteurs (HRV1 et HRV2) (Bégni et *al.*, 1986).

Une **correction géométrique** permet de rectifier les images satellitaires selon la grille de projection cartographique UTM propre aux cartes topographiques (1: 50 000) au Canada. Nous avons utilisé une série de points géodésiques fournis par le Centre canadien de géomatique de Sherbrooke pour les points de contrôle. Une équation polynominale du second degré détermine la transformation requise pour rendre les données satellitaires conformes à la grille de projection. En moyenne, l'erreur résiduelle obtenue en ligne et en colonne est de 3 mètres.

Nous avons créé trois **sous-images** représentatives du secteur d'étude et caractéristiques de l'ensemble des éléments planimétriques de la carte topographique. Une première sous-image se situe en milieu agricole près de la station de recherche d'Agriculture Canada à Lennoxville, une seconde couvre la ville de Sherbrooke et une troisième se localise en milieu forestier près de l'aéroport de Sherbrooke.

4.2 Traitements numériques

a) Rehaussement de contraste

Les différentes techniques de rehaussement visent à accentuer dans l'image les valeurs correspondant aux éléments d'intérêt tout en atténuant ou en éliminant celles correspondant aux objets inopportuns. Le rehaussement des contrastes qui en résultent permet une meilleure identification de la nature et de la localisation des éléments analysés.

Plusieurs techniques de rehaussement de contraste par étalement de l'histogramme ont été appliquées sur les sous-images tests. On peut modifier le contraste d'une image en effectuant un étirement de cet histogramme, c'est-à-dire une redistribution de l'étendue des valeurs numériques. Nous avons évalué 6 types d'étirements:

- . étirement linéaire
- . étirement linéaire par intervalle
- . étirement par égalisation
- . étirement par une loi puissance
- . étirement logarithmique
- . étirement par une loi exponentielle

b) Intégration du mode panchromatique au mode multibande

La méthode mise au point par Cliche et *al.*, (1985) permet l'intégration des données satellitaires provenant des deux modes d'acquisition. Ce rehaussement associe la qualité spatiale des données panchromatiques avec la qualité spectrale des données multibandes (Charbonneau et *al.*, 1988). Cette combinaison devrait améliorer l'identification des éléments linéaires et augmenter le rendu visuel, atténuant ainsi l'apparence plus floue des données multibandes. Quatre images satellitaires à 10 m de résolution spatiale ont été créées à partir de l'intégration des données de diverses dates:

1) intégration PLA (25-10-86) - MLA (02-08-87);
2) intégration PLA (25-10-86) - MLA (25-10-86);
3) intégration PLA (21-02-87) - MLA (25-10-86);
4) intégration PLA (21-02-87) - MLA (02-08-87).

c) Analyse en composantes principales

Cette technique vise à diminuer la corrélation entre les différents canaux en effectuant une rotation des axes de référence dans l'espace à N dimensions, N correspondant au nombre de canaux utilisés. La génération de signatures spectrales se fait en considérant l'ensemble de l'image test comme site d'entraînement. Les nouveaux axes, ou composantes principales sont définis par les valeurs propres des canaux d'entrée et sont calculées à partir des signatures spectrales définies précédemment.

d) Filtrage des images

Les filtres spatiaux sont des méthodes de rehaussement d'images. Il s'agit d'une combinaison arithmétique appliquée sur des pixels voisins pour augmenter ou réduire le contraste entre les différents éléments. Les filtres utilisés dans ce projet de recherche ne concernent que l'analyse des variations spatiales dans l'image (fréquence spatiale). Le filtrage repose sur l'utilisation d'une fenêtre mobile d'une taille spécifiée par l'opérateur et une nouvelle valeur du pixel central est calculée à partir de ses voisins dans cette fenêtre. De nombreux tests ont été expérimentés avec des fenêtres de tailles différentes (3 x 3, 5 x 5, 7 x 7, 9 x 9) mais les filtres utilisant une fenêtre 3 x 3 ont donné de meilleurs résultats. Divers filtres ont été évalués:

- . filtre passe-bas
- . filtre passe-haut
- . filtre contour
- . filtre directionnel

e) Combinaison de canaux

La création de néo-canaux résulte de la combinaison de différents canaux par des opérations mathématiques de base (+, /, -, *).

Nous avons expérimenté quelques combinaisons classiques:

a) IR / R ou R / IR
b) IR * R
c) (IR - R) / (IR + R)
d) IR + R
e) (IR - R) / [(IR + R + L) * (1.0 + L)]
f) IR - R ou R - IR

IR = canal infrarouge
R = canal rouge
L = constante variant entre 0,25 et 1,0

5. ANALYSE PRÉLIMINAIRE DES RÉSULTATS

Nous avons évalué le potentiel des divers traitements effectués pour la détection des éléments cartographiques. Les résultats présentés dans cet article correspondent à l'analyse des images test et non à l'ensemble du secteur d'étude. Parmi la liste du contenu thématique (tableau 1), on compte cinq thèmes qui ne sont pas présents sur la carte 21 E/5. Ces thèmes sont: route plus de deux voies non pavées, piste de karting, tunnel/passage souterrain, quai et voie ferrée multiple.

Dans une première étape d'analyse, nous avons fait des regroupements de éléments cartographiques car les caractéristiques spectrales de ces variables sont proches et ils sont donc difficiles à dissocier. On retrouve les associations suivantes:

1) bâtiment et installation industrielle
2) friche et pâturage
3) clairière, coupe à blanc, feu de forêt et percée
4) sol nu et labour

Par ailleurs, le thème cours d'eau a été divisé en deux catégories (rivière et ruisseau) et on a formé deux types de ligne de transmission d'énergie (petite et grande en fonction de l'importance du déboisement). On compte une douzaine d'éléments dans la liste qui ne sont pas identifiés comme données planimétriques traditionnelles. Il s'agit de: surface boisée-conifère, surface boisée-caduque, surface boisée-mixte, labour, sol nu, type de culture et les deux thèmes regroupés. friche/pâturage et clairière/coupe à blanc/feu de forêt/percée. Normalement, la carte topographique regroupe tous ces éléments en deux grandes familles. surface boisée et surface agricole.

ÉLÉMENTS CARTOGRAPHIQUES	
Autoroute	ENSEMBLE LINÉAIRE
Route plus de 2 voies pavées	
Route 2 voies ou moins pavées	
Route plus de 2 voies non pavées	
Route moins de 2 voies non pavées	
Chemin de charroi/sentier/chemin d'hiver	
Voie ferrée unique	
Piste d'atterissage - pavée	
Piste d'atterissage - non pavée	
Piste de karting	
Pont/passerelle	
Ligne de transmission d'énergie - petite	
Ligne de transmission d'énergie - grande	
Remonte pente	
Cours d'eau - rivière	
Cours d'eau - ruisseau	
Esker	
Chute	ENSEMBLE PONCTUEL
Rapide	
Roche	
Barrage	
Quai	
Bâtiment/installation industrielle	
Tour	
Cheminée	
Lac/étang	ENSEMBLE ZONAL
Abreuvoir/réservoir/piscine	
Bassin de filtration	
Île	
Terre inondable	
Marais/marécage	
Agglomération urbaine	
Terrain de golf	
Terrain de sport	
Carrière (gravière/sablière/mine/etc.)	
Cour de ferraille	
Dépotoir	
Surface boisée - conifère	AUTRES ÉLÉMENTS
Surface boisée - caduque	
Surface boisée - mixte	
Friche/pâturage	
Clairère/coupe à blanc/feu de forêt/percée	
Sol nu/labour	
Culture	
Type de cultures	

Tableau 1. Éléments du contenu thématique

Une cote allant de 1 à 4, soit de nulle à excellente, a été donnée pour chaque traitement en fonction de leur potentiel pour la détection des éléments cartographiques. Nous avons ajouté la lettre "c" pour confusion lorsque la reconnaissance d'un élément peut se confondre avec une ou plusieurs autres variables du contenu thématique. Pour faciliter la compréhension de l'analyse, les éléments sont groupés en quatre grands ensembles soit linéaire, ponctuel, zonal et autres éléments. Le dernier regroupement réfère aux éléments de type utilisation du sol et non inclus dans le contenu thématique habituel des cartes topographiques. Le choix des meilleurs traitements est basé sur les cotes allouées à chacun en fonction de leur potentiel de détection visuel sur écran vidéo à haute résolution. L'analyse visuelle a été réalisée par trois personnes et la cote allouée est la synthèse de l'équipe. Plus de cinquante types de traitements différents ont été analysés.

5.1 Ensemble linéaire

L'évaluation des données satellitaires brutes montre le fort potentiel de l'image PLA du 21-02-87 pour la distinction des éléments linéaires. Celle-ci s'est avérée la meilleure parmi les quatre images analysées. On retrouve un excellent contraste entre le réseau routier non-enneigé à cette date et l'ensemble de l'image. Le parallèle entre surface enneigée et non-enneigée fait ressortir tous les éléments linéaires.

Les résultats de l'interprétation des quatre images créées par l'intégration MLA + PLA s'avèrent presque similaires pour trois d'entre elles. Les images intégrées (PLA-aut. + MLA-été), (PLA-aut. + MLA-aut.) et (PLA-hiver + MLA-été) permettent une bonne distinction de ces éléments. La combinaison de l'aspect spatial des données panchromatiques associée à l'aspect spectral des données multibandes SPOT élimine l'effet "d'escalier" que l'on retrouve particulièrement pour les variables linéaires. Cependant on doit demeurer prudent pour l'interprétation des intégrations réalisées avec des images de différentes dates.

Dans l'ensemble, les images créées par une modification de l'histogramme ont procuré des résultats intéressants pour la détection des thèmes linéaires. Parmi les traitements effectués, l'image obtenue par un étirement de type loi puissance sur la première composante principale des données MLA du 02-08-87 à (10 m) permet la meilleure distinction du contenu linéaire. Au niveau des techniques de filtrage, l'application d'un filtre contour (Laplacien croisé rehaussé) sur les données PLA du 21-02-87 s'avère supérieure aux autres méthodes testées pour les éléments linéaires. Les composés fausses couleurs du 25-10-86 et 02-08-87 se sont avérés excellents pour la distinction des éléments linéaires. Cependant, il est à noter qu'un filtre contour (différence moyenne rehaussée avec une pondération du pixel central à 9) est appliqué sur la bande verte et sur la bande rouge. Les données de l'infrarouge demeurent intactes. Les filtres rehaussent le contenu linéaire des images.

En résumé, on constate que les thèmes autoroute et route plus de deux voies pavées ont été détectées par la majorité des traitements. Ces éléments ne demandent pas de traitement particulier pour leur distinction. On les détecte sur les images brutes mais la méthode d'intégration des données PLA aux données MLA facilite la détection de ces éléments. L'apport de la fine résolution spatiale des données PLA est considérable pour la discrimination des variables linéaires. La distinction des routes non pavées est seulement possible avec les données à 10 m de résolution spatiale, plus particulièrement sur l'image PLA d'hiver d'où l'importance de l'aspect spatial. La détection des thèmes chemin d'hiver/sentier/chemin de charroi, voie ferrée et ligne de transmission d'énergie est étroitement reliée au contexte environnant des éléments ainsi qu'à leur dimension (superficie). Par exemple, on discrimine facilement un sentier à l'intérieur d'une zone boisée. Par contre, on perd complètement le tracé d'une ligne de transmission dans les secteurs agricoles.

L'élément piste d'atterissage non pavée a été impossible à détecter. Au niveau spectral, ce thème se confond avec les zones de friches et/ou de pâturages. Un problème différent se pose pour l'élément remonte-pente. On peut déduire qu'il existe un remonte-pente lorsque l'on distingue des pistes de ski bien que l'on ne voit pas l'infrastructure proprement dite. Le réseau hydrograpique et plus particulièrement la distinction des rivières n'a pas montré de difficulté majeure.

5.2 Ensemble ponctuel

La détection des éléments ponctuels est dans l'ensemble peu convaincante. Ces thèmes sont généralement de petite taille et leur distinction est limitée à la résolution spatiale des données satellitaires. Pour l'ensemble des thèmes ponctuels, l'élément bâtiment/installation industrielle a été le plus facile à identifier parmi toutes les images retenues en ne prenant en considération que les constructions ayant une grande superficie. Une confusion demeure entre le bâtiment en tant que tel et l'espace réservé pour le stationnement. Aucun traitement n'a permis la détection des éléments tour et cheminée.

Il est intéressant de noter que l'on distingue des zones d'eau libre (non gelée) sur les deux principales rivières sur l'image PLA d'hiver. Il est donc possible de localiser les rapides, les barrages et

des zones de pollution bien que cette distinction puisse être liée au débit des cours d'eau.

5.3 Ensemble zonal

Les données MLA brutes permettent une bonne distinction des éléments zonaux. Cependant l'image d'automne offre plus de potentiel de détection. On arrive facilement à différencier les zones de feuillus des zones de conifères, les terrains de golf des pâturages et des secteurs en friche. Les images issues des intégrations n'augmentent pas la détection des éléments zonaux car ces thèmes possèdent des dimensions assez grandes. Une résolution spatiale plus fine n'améliore pas la différenciation de ces variables.

Les images rehaussées par modification de l'histogramme, de techniques de filtrage et de combinaisons de canaux ont été moins efficaces pour la détection des éléments zonaux que des thèmes linéaires. Les composés fausses couleurs avec les bandes verte et rouge filtrées ont été aussi efficaces que les données MLA habituelles. La détection des éléments zonaux de petite taille n'est pas possible (lac/étang, abreuvoir/réservoir/piscine et petite île). La discrimination de ces thèmes est limitée à la résolution spatiale des données satellitaires. Le contraste entre la piste de course à pied et les zones couvertes de gazon facilite la détection de l'élément terrain de sport.

Les variables carrière, cour de ferraille, dépotoir, sol nu et labour se confondent et l'on doit analyser le contexte environnant pour les différencier. L'élément agglomération urbaine ne demande pas de traitement particulier, on le détecte sur toutes les images.De façon générale, les éléments zonaux se sont avérés être les variables nécessitant les traitements les moins sophistiqués.

6. CONCLUSION ET SYNTHÈSE DES RÉSULTATS

La synthèse des résultats est résumée dans le tableau 2. On présente l'occurence des cotes 4 et 3 de divers traitements évalués en fonction des ensembles linéaire, ponctuel, zonal et autres éléments.

Plusieurs constatations se dégagent de ce tableau. On remarque que les composés fausses couleurs à 10 m résultant de l'intégration PLA + MLA ont été supérieurs aux autres traitements pour la détection de l'ensemble du contenu thématique (totaux entre 12 et 14 pour la cote 4). Les composés fausses couleurs avec un filtre contour sur les bandes verte et rouge se sont avérés également très performants pour la distinction des éléments cartographiques (totaux 12 et 13 pour la cote 4).

De façon globale, la détection des ensembles linéaire et zonal n'ont pas montré de difficulté majeure. Nous avons trouvé des traitements appropriés pour la distinction des divers éléments. Cependant, le tableau 3 montre le faible potentiel des traitements expérimentés pour la reconnaissance des variables ponctuelles. Ces éléments thématiques sont généralement de petites tailles et la résolution spatiale demeure toujours une contrainte pour leur détection.

Pour les variables de l'ensemble autres éléments, différents traitements testés se sont avérés efficaces pour distinguer les surfaces boisées de type conifère et caduque. Les meilleurs traitements sont les composés fausses couleurs à 10 m ou 20 m. La détection du groupe clairière/coupe à blanc/feu de forêt/perçée a été obtenue par la majorité des traitements. Les résultats sont moins probants pour l'identification des types de culture. Au niveau spectral, un champs en labour est considéré comme un sol nu d'où leur confusion.

Ce projet est réalisé grâce à une proposition spontanée financée par le MAS/ÉMR/MDN (contrat no. 23258-8-9002 / SD).

8. RÉFÉRENCES

Bégni, G., Dinguirand, M.C., Jackson, R.D. and Slater, P.N. " Absolute calibration of the SPOT-1 HRV cameras." Proceedings SPIE's Conference, Innsbruck, Austria, 1986.

Bégni, G., Henny, P., Leroy, M. et Dinguirand, M.C. " La qualité des images SPOT." Compte rendu du Colloque SPOT, Utilisation des images, Bilan, Résultats, pp. 1505-1516, 1987.

Charbonneau, L., Perras S., Bonn, F. et Brochu, R. "Analyse et interprétation d'images SPOT d'un secteur urbain et agro-forestier." Journal canadien de télédétection, Vol. 14, No. 2, pp. 92-104, 1988.

Clément, P., Bonn, F. et Dubois, J.M.M. "Télédétection et caractéristiques physiques des dépôts meubles dans un milieu anciennement englacé du sud du Québec." Géographie physique et Quaternaire, Vol. 37, pp.107-117, 1983.

Cliche, G., Bonn, F. and Teillet, P.M. "Integration of the SPOT Panchromatic Channel into its Multispectral Mode for Image Sharpness Enhancement." Photo. Eng. and Remote Sensing, Vol. 51, No. 3, pp. 311-316, 1985.

Dubois, J.M.M. " Proposition de régions physiographiques pour les Cantons de l'est: un apport à la classification de Bostock." Géoscope, Vol. V, No. 2, pp. 13-46, 1974.

Gregory, A.F. and Moore, H.D. "Economical Maintenance of a National Topographic Data Base Using LANDSAT Images." Photo. Eng. and Remote Sensing, Vol. 52, No. 4, pp. 519-524, 1986.

	ENSEMBLE LINÉAIRE		ENSEMBLE PONCTUEL		ENSEMBLE ZONAL		AUTRES ÉLÉMENTS		TOTAL	
Traitement	Cote 4	Cote 3	Cote 4	Cote 3	Cote 4	Cote 3	Cote 4	Cote 3	Cote 4	Cote 3
A	4	2	0	2	4	4	3	2	11	10
B	4	2	0	1	4	1	5	1	13	5
C	4	4	0	1	2	2	0	1	6	8
D	5	4	1	0	1	1	1	2	8	7
E	6	3	0	1	4	4	2	1	12	9
F	6	2	0	1	2	2	5	0	13	5
G	5	2	1	1	2	2	4	1	12	6
H	6	2	1	1	3	2	4	1	14	6
I	6	4	0	1	6	2	1	5	13	12
J	6	3	0	1	4	1	2	5	12	10

A: données MLA du 2 août 1987
B: données MLA du 25 octobre 1986
C: données PLA du 25 octobre 1986
D: données PLA du 21 février 1987
E: intégration PLA (25-10-86) + MLA (02-08-87)
F: intégration PLA (25-10-86) + MLA (25-10-86)
G: intégration PLA (21-02-87) + MLA (25-10-86)
H: intégration PLA (21-02-87) + MLA (02-08-87)
I : données MLA rehaussées (02-08-87) avec un filtre contour (différence moyenne avec une pondération du pixel central à 9) pour les bandes verte et rouge
J: données MLA rehaussées (25-10-86) avec un filtre contour (différence moyenne avec une pondération du pixel central à 9) pour les bandes verte et rouge

Tableau 2. Synthèse des résultats: Occurence des cotes 4 et 3 pour les divers traitements.

REMOTE SENSING DETECTION OF HYDROCARBONS AND HYDROCARBON SEEPS

Edward A. Cloutis
Department of Geology
University of Alberta
Edmonton, Alberta T6G 2E3
Tel: (403) 492 3265

The spectral reflectance properties of Athabasca tar sands in the 0.35-2.6μm wavelength range were examined because of their potential importance for remote sensing detection of surficial hydrocarbon seeps. The data are also useful for attempts to develop reflectance spectroscopy into a rapid, non-invasive analytical tool for the characterization of tar sands, heavy oils, oil shales, coals, and other hydrocarbon-bearing materials.

A common initial step in hydrocarbon exploration is the examination of available remote sensing data. The emphasis at this stage has all too often been on the surficial structural geology. Potentially useful geochemical and mineralogical information is generally ignored, perhaps due to the fact that the spectral reflectance properties of many geological materials are poorly known. The spectral properties of tar sands were examined in order to provide laboratory data of use to remote sensing data analysis. Tar sands were selected for study because they are compositionally, and perhaps spectrally, very similar to oxidized hydrocarbon outcrops.

The laboratory-derived reflectance spectra of tar sands show an abundance of diagnostic absorption features in the 0.35-2.6μm interval. Discrete absorption bands due to clay, water and bitumen/organics appear at well-defined wavelengths. The strengths of the absorption features are directly related to the various phase abundances. The primary absorption bands due to the hydrocarbons appear at ~1.7μm, and from 2.3 to 2.6μm. Clay absorption bands are found near 1.4 and 1.9μm. These bands are attributed to bound and structural water in the clay structure and the wavelength positions of the band minima are diagnostic of the particular clay(s) present. The shape and broadness of these bands is also a function of the degree of crystallinity of the clay and the number of distinct structural sites available to the water and OH molecules. Absorption bands due to free water (unattached to the clays) are also present near 1.4 and 1.9μm, but differ slightly from the wavelength positions of the clay bands. These differences are easily resolvable in moderate resolution spectra.

The laboratory data suggest that ~4 wt. % bitumen is required in a sample to be unambiguously detected. This detection limit will of course vary somewhat depending on the nature of the other phases present in the material. The 1.7μm absorption band is potentially very important for remote sensing because it lies outside the regions of atmospheric absorption and is not significantly overlapped by the most common mineral absorption bands. The spectral slope in the 2.3-2.6μm region can also be used for hydrocarbon detection. A negative slope characterizes all known clays, while a positive slope is present in the spectra of bitumen-rich (>10 wt. %) samples.

Many organic-rich materials show certain broad spectral similarities to the tar sands-low overall reflectance, an absorption near 1.7μm, and a positive slope towards longer wavelengths. While these spectral parameters are by no means unique, they can be used to severely constrain the possible compositions of a target when only very low resolution spectral data is available. As the capabilities of terrestrial remote sensing platforms improve, particularly with the upcoming EOS program in the 1990's, detailed laboratory spectral data must be available to effectively analyze the information which will be obtained.

Acknowledgements: This study and the researcher were supported by a grant and scholarship from the Alberta Oil Sands Technology and Research Authority (AOSTRA), and a Grant-in-Aid (#582-12-01) from the American Association of Petroleum Geologists.

PARAMETERS AFFECTING REFLECTANCE OF CONIFEROUS FORESTS IN THE REGION OF CHLOROPHYLL PIGMENT ABSORPTION

Brian Curtiss[1] and Susan L. Ustin[2]

[1]Center for Study of Earth from Space (CSES) / CIRES
University of Colorado Boulder, CO 80309
(303) 492-6718 (303) 492-5070 (FAX)

[2]Department of Botany
University of California Davis, CA 95616
(916) 752-0621

Abstract

The observed variation in the reflectance in the region of chlorophyll absorption (450 to 900 nanometers) of mature Pinus *ponderosa* (Ponderosa Pine) is explained in terms of pigment concentrations and stress associated with pollutant exposure Reflectance and chlorophyll concentrations were measured on each years growth on each of 20 trees at four sites along an ozone exposure gradient in the southern Sierra Nevada. Additionally, each of these sites was characterized in terms of stand density and age. Exposure to ozone, a strong oxidant, results in chloroplast injury. Granulation of the chloroplast stroma is an early metabolic response to ozone exposure. Thus, an examination of the reflectance changes observed along the ozone gradient provide a means for quantifying the spectral changes associated with chloroplast damage and reduced photosynthetic efficency. Additionally, the spectral trends associated with increasing needle age were related to the effects of cumulative ozone exposure and changes in overall chlorophyll concentration. It was found that the within site variation of the reflectance spectra was directly related to total chlorophyll concentration; an increase in chlorophyll concentration resulted in both a shift in the "red-edge" inflection to longer wavelength and an increase in the depth of the chlorophyll absorption well. Similarly, an increase in chloroplast damage, associated with higher levels of ozone exposure, also produced shift to longer wavelengths of the "red-edge" but a decrease in the depth of the absorption well. The absorption band broadening observed in the high ozone exposure sites is indicative of an increase in the disorder of absorbing medium. In the case of chlorophyll, the observed broadening is probably the result of damage to the thylacoid membranes within the chloroplasts.

Keywords. remote sensing, photosynthesis, chlorophyll

Introduction

One of the major goals within the field of botanical remote sensing is the accurate assessment of ecosystem productivity. Such existing and proposed models for the determination of ecosystem productivity assume a linear relationship between APAR (absorbed photosynthetically active radiation) and vegetation indices such as NDVI (normalized difference vegetation index), (Sellers 1985 & 1987). Additionally, it is generally assumed that APAR is a direct measure of the maximum photosynthetic capacity of an ecosystem and that the photosynthetic efficiency of an ecosystem is constant. Thus, these models calculate ecosystem productivity by integrating APAR over the growing season. These models work quite well for ecosystems where the assumptions are valid.

Several factors can result in errors in NDVI based estimates of APAR. In ecosystems with incomplete canopy closure, the spectral signature received by the sensor has components from both the vegetation and from the substrate (eg. soil, rock, plant litter). Depending on the viewing and illumination geometries, canopy architecture, and on the albedo of the substrate in the region of chlorophyll absorption, errors in NDVI estimated APAR can be as high as 50 percent. The presence of a large proportion of shadow can also produce errors in the observed reflectance of a forest canopy (Smith et al. 1983; Curtiss and Ustin 1988), that, if they are not corrected for, can result in an underestimation of NDVI. Thus, calculated NDVI can vary within a scene due to changes in viewing geometry, and can change seasonally because of illumination geometry. These changes will occur independent of any real change in ecosystems photosynthetic capacity.

Ecosystems tend to adjust their maximum photosynthetic capacity in relation to available resources through changes in leaf area index and other architectural changes, chlorophyll content, and species composition When the rate change in a stress is greater than the rate at which the ecosystem can adjust its photosynthetic rate, the assumption of constant photosynthetic efficiency is not valid. Examples of this are: short term stresses, such as diurnal water stress; longer term (3-6 months) stress for ecosystems with slow rates of photosynthetic capacity adjustment, such as coniferous forests.

In some instances, these assumptions produce errors that cancel each other, and, therefore, give the appearance that the model is working. Northern latitude boreal forests are an example of such an ecosystem. During the winter months, calculated NDVI for these ecosystems goes down as would be expected if NDVI is related to photosynthetic productivity. In reality, NDVI goes down not because of a reduction in APAR, but because of the increase importance of shadowing associated with the seasonal change in solar zenith angle. This underestimate of APAR is offset by a lowering of photosynthetic efficiency resulting from the lower winter temperatures. Because the predictions of ecosystem productivity are linked to solar zenith angle, it is likely that they will be insensitive to true changes in boreal forest productivity such as might be associated with longer, colder winters or more frequent spring freezes.

Many of the errors associated with NDVI based estimates of APAR can be overcome by using contiguous, high spectral resolution measurements made by sensors such as the High Resolution Imaging Spectrometer planed for the Eos polar platform, (Goetz and Herring 1989). These new sensors measure the reflected spectrum in many, contiguous, spectral

bands in the visible and reflected infrared wavelengths; this data can be used for the direct identification and quantification of spectral features diagnostic of leaf constituents and, will thus permit a more physiologically based understanding of vegetation spectra. Spectral patterns within the visible and shortwave infrared regions are diagnostic for the presence of vegetation (Asrar et al., 1984). They may permit the identification of fundamental physical canopy properties (Goel and Thompson, 1984a,b; Sellers, 1985, 1986), and potentially define spectral patterns related to the assessment of environmental stress (Chang and Collins, 1983; Bauer, 1985). While minerals can be identified based on the presence of diagnostic spectral features (Goetz et al. 1985), vegetation spectra are generally very similar to one another due to their small base of shared constituents, e.g. pigments, cellulose, lignin, protein, starch. Variations in shape of the reflectance spectra of vegetation will be largely controlled by differences in leaf and canopy structure and foliar optical properties. Additionally, because of the high dimensionality of these datasets, spectral un-mixing models (Adams and Adams 1984; Smith et al. 1985; Adams et al. 1986) can be used to separate the spectral signatures present in a mixed pixel (eg. vegetation, substrate, shade, see Ustin et al. 1986).

To date, there have been few experimental studies concerning the remote detection of photosynthetic properties. Most experimental studies investigating spectral characteristics associated with physiological processes have examined environmental stress(es), such as drought or mineral excess, on the reflectance characteristics of individual leaves (Chang and Collins 1983; Ripple 1986). Stress-induced shifts in the red-edge have been reported, both toward longer (i.e. red) or shorter (blue) wavelengths, and have been attributed to changes in chlorophyll content. This red-shift of the chlorophyll absorption edge was first reported by Gates et al. (1965) and later by Collins (1978).

In a pilot study of ozone treated conifers, Ustin and Curtiss (1987) found a marked increase in reflectance in the chlorophyll absorbance bands of ozone treated conifers (Figure 1). Additionally, they report a small blue shift of 3-9 nm in the derivative of the inflection point wavelength of the 'red edge' Such changes are consistent with reports of chlorophyll loss as a characteristic consequence of ozone injury. Further, Ustin and Curtiss (1987) suggest that subtle changes in the 590-690 nm spectral region indicate that chlorophyll a and b concentrations may be quantified in addition to total chlorophyll

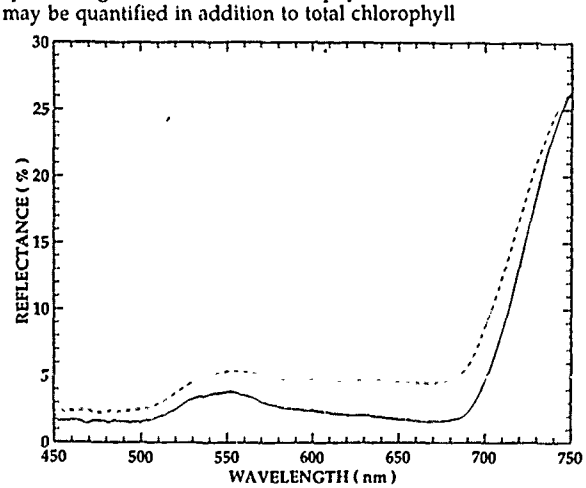


Figure 1. Mean spectra for ozone exposed (dash dot line) and control (solid line) Ponderosa Pine seedlings. The seedlings were exposed to a single seasons duration of realistic fumigation with Dose levels representative of summer air quality experienced in many parts of the western United States.

This paper presents a first step towards a better understanding of the parameters that affect that portion of the reflected light spectrum dominated by chlorophyll pigment absorption. It is hoped that this understanding will provide a basis for the development of more accurate techniques for the assessment of APAR and photosynthetic efficiency.

Methods

The reflectance spectra of ozone exposed *Pinus ponderosa* (Ponderosa Pine) seedlings and needles from mature trees were measured with the NASA Jet Propulsion Laboratories' PIDAS spectroradiometer (Portable Intelligent Display and Analysis Spectrometer), (Goetz 1987). This instrument makes reflectance measurements in 1 nm wide spectral bands in the visible and near infrared (400-900 nm) and in 5 nm wide spectral bands in the near and shortwave infrared (900-2450 nm). The foreoptics, with a 6^{o} field of view was placed approximately one meter above samples, giving about a 10 centimeter field of view. Illumination was provided by two 120 volt 1000 watt floodlamps operated from two 120 volt 10 ampere constant voltage supplies to provide a constant and stable light source The lamps were placed opposite each other, to minimize shadows, and about 0.5 meters above the canopy. Irradiance at the canopy was approximately equivalent to full sunlight.

The Ponderosa Pine seedlings were exposed to a single seasons duration of realistic fumigation profiles having stochastic daily doses and episodic peaks. Dose levels were chosen to be representative of summer air quality currently experienced in many parts of the western United States Reflectance measurements of the seedlings were made on 9 different areas on composite canopies formed from arranging about 30 seedlings in a tray. The elevation of the trays holding the seedlings were adjusted to maintain a constant distance between the canopy top and the PIDAS foreoptics FiberFrax, a uniformly reflective ceramic fiber wool, was used as the reflectance standard. This standard was placed at the same distance from the foreoptics as the Ponderosa Pine canopies and was measured before and after each set of canopy measurements.

The Southern Sierra Nevada is an area of regional ozone exposure. Reflectance spectra of three branches from each of 20 Ponderosa Pine trees at four sites experiencing a range of ozone exposure levels were measured using a setup similar the one used for the seedlings. Chlorophyll analysis and reflectance measurements where made separately on each years growth For these measurements, Halon was used as a reflectance standard. A more complete description of the site selection and sample analysis is presented in Ustin et al. (1988).

Results

Ozone tends to be regionally distributed and has been clearly implicated in forest decline (McLauglin 1986). Chronic exposure to ozone, a strong oxidant, results in direct injury of cell membranes, with chloroplast injury being an early metabolic response. General disruption of chloroplast function proceeds other cellular changes (Thompson et al. 1966). Thus, ozone exposure provides a means for examining the spectral changes associated with reductions in both photosynthetic capacity and efficiency.

The primary spectral changes observed in the seedlings is that associated with chlorophyll loss, see Figure 1. Chlorophyll loss results in an increase in reflectance in the region of chlorophyll pigment absorption (450 - 700 nm). Additionally, this chlorophyll loss results in a shift in the red edge inflection to shorter wavelengths (a blue shift). The red edge inflection is defined as the center at half maximum of the first derivative of the spectrum, see Figure 2. The use of the wavelength of the red edge inflection (REI) is useful because it is much more

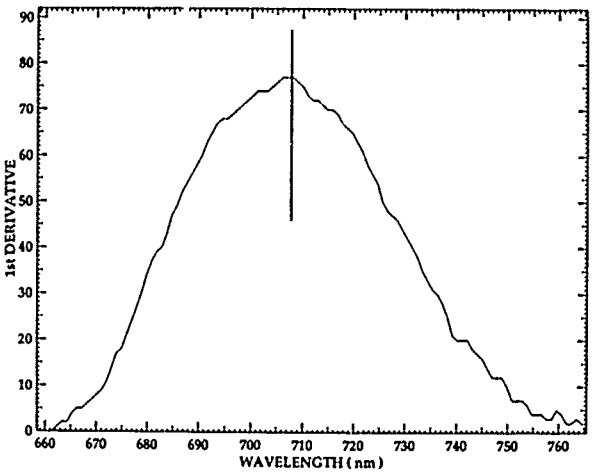


Figure 2. First derviative spectrum of a reflectance spectrum of Ponderosa Pine showing the methodology for the determination of the wavelength of the REI. The REI is defined as the center wavelength at half maximum of the first derivative spectrum. This technique is less sensitive to noise than the use of the wavelength of the first derivative maximum or the wavelength that the second derivative crosses zero.

insensitive to mixed pixel effects as compared to NDVI or other band ratio techniques for the quantification of light absorption by chlorophyll. The reflectance spectra in Figure 3 show that the reflectance spectrum of an ozone exposed seedling could be confused with the mixed signature of a control (no ozone exposure) seedling and granitic soil when using a broad band sensor such as TM or AVHRR. However, the wavelength of the REI of the mixture does not match the REI of the ozone exposed seedling.

While the relationship between chlorophyll concentration and maximum photosynthetic capacity is good within a single species, architectural variation between species and environmentally produced variation in photosynthetic efficiency result in a breakdown of this relationship. Thus, chlorophyll concentration can be used to asses the ability of parameters derived from reflectance data to predict maximum photosynthetic capacity only when the comparison is made within a species and when all samples have equivalent photosynthetic efficiency. Figure 4 show the relationship between REI wavelength and needle chlorophyll content for two site with similar ozone exposures. Each years growth (up to four) for 20 trees at 2 sites is plotted as a separate data point.

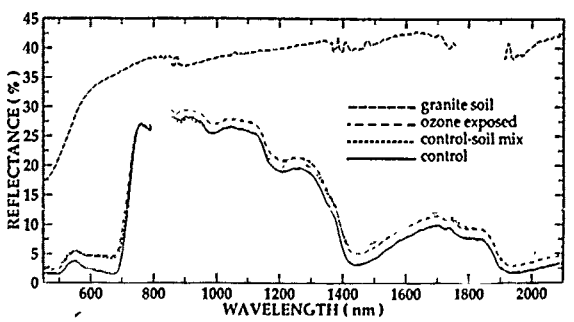


Figure 3. Simulated Ponderosa Pine and soil spectral mixtures. Tan granitic forest soil and Ponderosa Pine control spectra used in simulated mixture (92.8), and the mean ozone exposed spectra are plotted. While the mixed spectrum models the that of the ozone exposed seedling quite well, the REI wavelengths do not match.

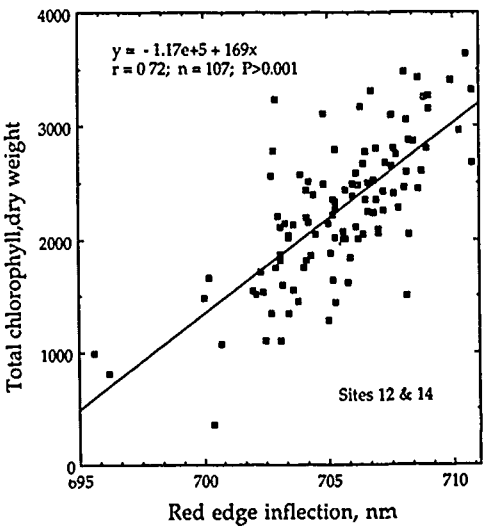


Figure 4. The relationship between needle chlorophyll concentration (micrograms/gram) and the wavelength of the red edge inflection. Data shown are the chlorophyll concentration and red edge inflection for each years growth on 20 trees at each of two sites in the Southern Sierra Nevada having moderate to low visual symtoms of ozone injury.

Some of the variance about the regression line observed in Figure 4 can be explained in terms of variance in photosynthetic efficiency. Within a single site there is considerable variation in the observed visual symptoms of ozone injury, such a chlorosis, low needle retention and fewer years of needles retained. The greater chloroplast damage in these trees produces a broadening of the chlorophyll absorption bands that results in a shift to longer wavelengths (a red shift) of the REI (see Fig. 5) that is indicative of a lower overall photosynthetic efficiency.

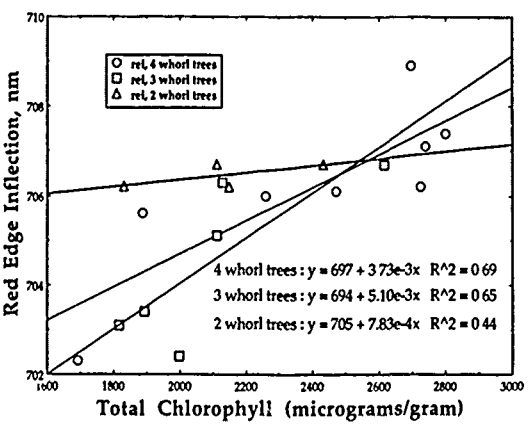


Figure 5. The relationship between needle chlorophyll concentration and the wavelength of the red edge inflection for two sites in the Southern Sierra Nevada having moderate to low visual symtoms of ozone injury. Values for individual branches have been grouped based on the number of years needle retention on the the branch. Branches having the needle classes of the same age have similar total chlorophyll concentrations and red edge inflection values. Trees experiencing higher levels of ozone damage (those trees with only two years of needles retained) have chlorophyll absorption bands that are correspondingly broader.

Summary

Photosynthetic properties are potentially measurable using parameters derived from remote sensing data such as the wavelength of the red edge inflection. The wavelength of the red edge inflection is quite insensitive to mixed pixel effects, and, therefore may be used for ecosystems with incomplete canopy closure. Shifts in the wavelength of the red edge inflection to shorter wavelength in conjunction with a reduction in the chlorophyll absorption well depth are indicative of a reduction in total chlorophyll, and, thus, a reduction in photosynthetic capacity. Actual photosynthetic rates are often less than the photosynthetic capacity due to ambient environmental conditions, such as chronic ozone exposure, drought or high temperature stress. Red edge shifts to longer wavelengths produced by ozone induced chloroplast membrane damage and are indicative of this reduction in photosynthetic efficiency.

Because of the complex interaction of chlorophyll absorption band broadening and band depth the ability to determine the components of the photosynthetic process using the remote sensing techniques discussed here is untried and problematic. However, current techniques relying on measures like NDVI have been demonstrated to be quite sensitive to canopy structure, variability of the substrate, and viewing and illumination geometry. Thus, more accurate means of predicting photosynthetic capacity and efficiency are required before remotely sensed techniques can be applied for quantitative scene-independent predictions of ecosytem productivity. If the specific chlorophyll absorption feature parameters identified in this paper can alone, or in combination, produce greater reliability of prediction, then the results of this study will have broad significance.

Acknowledgments

Although the information in this document has been funded in part by the United States Environmental Protection Agency under Grant #R-814274-01-0 and contract #7B0008NTEX, it does not necessarily reflect the view of the Agency and no official endorsement should be inferred. Additional funding for this study was provided under NASA/JPL contract #958039.

References Cited

1. Adams, J.B. and J.D. Adams. 1984. Geologic mapping using Landsat MSS and TM images: removing vegetation by modeling spectral mixtures. *Proc. Int. Symp. Rem. Sens. Environ. Third Thematic Conf: Remote Sensing for Exploration Geology.* Pg. 615-622. Environ. Rsrch. Inst. Mich.
2. Adams, J.B., M.O. Smith and P.E. Johnson. 1986. Spectral mixture modeling: a new analysis of rock and soil types at the Viking Lander I site. *J. Geophys. Res.* V80.
3. Asrar, G., E.T. Kanemasu, and M. Yoshida. 1985. *Remote Sens. Environ.* 17: 1
4. Bauer, M.E. 1985. Spectral Inputs to Crop Identification and Condition Assessment. *Proc IEEE* 73: 1071-1085.
5. Chang, S.H. and W. Collins. 1983. Confirmation of the airborne biogeophysical mineral exploration technique using laboratory methods. *Econ. Geol.* 78: 723-736.
6. Collins, W. 1978. Remote sensing of crop type and maturity. *Photogram. Eng. and Remote Sens.* 44: 43-55.
7. Curtiss B. and S.L. Ustin. 1988. Characterization of sources of illumination in Ponderosa Pine forest communities using the Portable Instantaneous Display and Analysis Spectrometer. *SPIE technical Symposium on Optic, Electro-Optics and Sensors* in Orlando, Florida. April 4-8.
8. Gates, D.M., H.J. Keegan, J.C. Schleter, and V.R. Weidner. 1965. Spectral properties of plants. *Appl. Optics* 4: 11-20.
9. Goel, N.S. and R.L. Thompson. 1984a. Inversion of vegetation canopy reflectance models for estimating agronomic variables. III. Estimation using only canopy reflectance data as illustrated by the Suits model. *Remote Sens. Environ.* 15: 223-236.
10. Goel, N.S. and R.L. Thompson. 1984b. Inversion of vegetation canopy reflectance models for estimating agronomic variables. IV. Total inversion of the SAIL model. *Remote Sens. Environ.* 15: 237-253.
11. Goetz, A.F.H., G. Vane, J.E. Solomon and B.N. Rock. 1985. Imaging spectrometry for earth remote sensing. *Science* 228:1147-1153.
12. Goetz, A.F.H. 1987. The portable instantaneous display and analysis spectrometer (PIDAS). *Proc. 3rd. Airborne Imaging Spectrometer Data Analysis Workshop,* June 2-4, JPL Publication #87-30.
13. Goetz, A.F.H. and M. Herring. 1989. The High Resolution Imaging Spectrometer (HIRIS) for Eos. *IEEE Trans. Geosci. Remote Sensing* 27: 1136-144.
14. McLaughlin, S.B. 1985. Effects of air pollution on forests: A critical review. *J. APCA* 35:512-534.
15. Ripple, W. 1986. Spectral reflectance relationships to leaf water stress. *Photogramm. Engin. and Remote Sens.* 52: 1699-1675.
16. Sellers, P.J. 1985. Canopy reflectance, photosynthesis and transpiration. *Int J. Remote Sens.* 6: 1335-1372.
17. Sellers, P.J. 1986. Canopy reflectance, photosynthesis and transpiration. II. The linearity of their response. *Rem. Sensing Environ.* 21:143-183.
18. Smith M., Cymerman C., Curtiss B. 1983. Application of reflectance models to mapping forest resources using LANDSAT images. *Proc RNRF Symp. Appl. Remote Sensing to Resource Manag.,* 137-146.
19. Smith, M.O., P.E. Johnson and J.B. Adams. 1985. Quantitative determination of mineral types and abundances from reflectance spectra using principal components analysis. *Proc. 15th Lunar and Planetary Sci. Conf. II. J. Geophy. Rsrch.* 90.C797-C804.
20. Schutt, J.B., R.R. Rowland, and W.H. Heartly. 1984. a laboratory investigation of a physical mechanism for the extended infrared absorption ('red shift') in wheat. *Int. J. Remote Sens.* 5: 92-102.
21. Thomson, W.W., W.M. Dugger and R.L. Palmer. 1966. Effects of ozone on the fine structure of the palisade prenchyma cells of bean leaves. *Can. J. Bot.* 44:1677-1682.
22. Ustin S.L., J.B. Adams, C.D. Elvidge, M. Rejmanek, B.N. Rock, M.O. Smith, R.W. Thomas and R.A. Woodward. 1986. Thematic Mapper studies of semiarid shrub communities. *BioSci.* 36: 446.
23. Ustin SL and B. Curtiss. 1987. Spectral Characteristics of Ozone Treated Conifers. Final Report, EPA contract #7B0008NTEX.
24. Ustin S.L., B. Curtiss, S. Martens, V.C. Vanderbilt . 1988. Use of high spectral resolution sensors to detect air pollution injury in conifer forests. in *Proc. EPA/ASPRS Symp. on the Remote Sensing of Air Poll.* St. Louis, MO April 2-8.

TRANSMISSION, BIDIRECTIONAL REFLECTANCE AND DEPOLARIZATION OF AN OPTICAL WAVE FOR A SINGLE PLANT LEAF

Qinglin Ma, Akira Ishimaru, Phillip Phu
Department of Electrical Engineering
University of Washington
Seattle, Washington 98195

Yasuo Kuga
Department of Electrical Engineering and Computer Science
University of Michigan
Ann Arbor, Michigan 48109

ABSTRACT

A model for optical scattering from a single leaf is developed using the vector radiative transfer theory for electromagnetic wave propagation in random media and Kirchhoff rough surface scattering theory. We model a generalized leaf as a slab of water with an irregular surface containing randomly distributed spherical scatterers. A modified vector radiative transfer equation and its boundary conditions are used to deal with the multiple scattering process within the leaf and Kirchhoff rough surface scattering theory is used to include the surface effects. The optical parameters of a single leaf in the calculation are chosen within certain ranges based on the knowledge we have about leaves The normalized scattering cross section per unit area for the transmitted and reflected vertical and horizontal intensities for a horizontally polarized plane wave normally incident on a leaf from such a model are computed for corn, potato and laurel leaves and compared with experimental results. The agreement between the theoretical results and experimental data verifies the optical scattering model for a single leaf.

Introduction

To interpret,the remote sensing data acquired from aircraft and spacecraft, it is necessary to understand the reflectance produced by features on the surface of the earth. A specific problem in agriculture is interpretation of reflectance produced by vegetation. Therefore, the problem of the interaction of light with a plant canopy has attracted much attention in recent years. There are several theories that deal with this problem, such as the two flux, or Kubelka-Munk theory (*Allen, et al.*, 1970; *Gausman and Allen*, 1973, *Allen and Richardson*, 1968) which is applicable at visible light frequencies and the strong fluctuation theory (*Fung and Ulaby*, 1978; *Tsang and Kong*, 1981; *Ulaby and Wilson*, 1985) at microwave frequencies. However, all of these studies treat the canopy as a whole object rather than looking at each individual leaf. The problem of how a single leaf affects the propagation of electromagnetic waves has not been paid much attention. In this paper, we study the propagation, scattering and depolarization of an electromagnetic wave by a single leaf.

First, based on the knowledge we have about leaves, we model a generalized single leaf in the optical wavelength range as a slab of water, with an irregular surface containing randomly distributed spherical scatterers with a Gaussian size distribution. Then, vector radiative transfer theory and Kirchhoff rough surface scattering theory are applied to this scattering model to obtain the normalized scattering cross section per unit area for the transmitted and reflected intensities for a horizontally polarized normally incident plane wave. The modified radiative transfer equation with Stokes vectors and boundary conditions is solved numerically by Fourier expansion, a discrete-ordinate technique and an eigenvalue eigenvector method using a supercomputer Cray X/MP. An experiment is made to measure the transmitted and reflected copolarized and cross-polarized intensities versus the azimuth angle for corn, potato and laurel leaves at an incident He-Ne laser wavelength of 0.6328μ. All of the control and data acquisition are carried out on an IBM-PC computer. Finally, the theoretical results are compared with the measured data.

An Optical Scattering Model for a Single Leaf

Plants depend on radiant energy to carry on photosynthesis and other physical processes. The interaction is mainly of two forms: absorption and scattering. The light absorbed by the leaves will be converted into heat or fluorescence. Scattering may be caused by the plastids inside the leaf, such as grana, ribosome, nuclei, etc.. which are of the dimension of a wavelength of light. Because of the high water content of fresh leaves, we model a generalized single leaf as a slab of water with an irregular surface containing randomly distributed spherical scatterers . For mathematical simplicity, we assume that the multiple scattering process within the leaf is independent of the surface scattering process and that the surface does not scatter the incoherent wave. Under these assumptions, the rough surface does not affect the boundary conditions of the radiative transfer equation. Because the optical distance of the leaf is large ($\tau \sim 10$), the transmitted coherent intensity is negligibly small. Therefore, we assume that the bottom surface of the leaf is not rough. This model is shown in Fig.1, where $n_1 = n_3 = 1$ is the refractive index of air, $n_2 = 1.33$ is that of water and n_0 is that of the scatterer. We assume that the size distribution of the scatterers is Gaussian with a mean diameter comparable

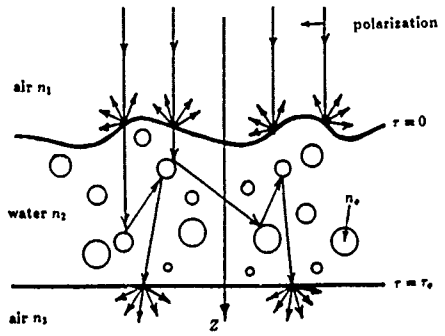


Fig.1: Geometry of the optical scattering model.

to a wavelength of light, i.e. ka~1, where k is the wavenumber in water and a is the mean radius of the scatterers, and that the standard derivation of the Gaussian distribution is 30%. Considering the high absorption of visible light by the leaf, we choose the optical distance τ about 10. The last parameter needed to be known for theoretical calculation is the refractive index n_0 of the scatterer. From the measurement of Allen and others (1968), we can choose the real part of n_0 about 1.4 for the wavelength 0.6328μ. The imaginary part of n_0 can be determined by fitting the theoretical results with experimental results.

Now we have all the parameters we need to carry out the theoretical calculation using the radiative transfer theory. Considering the short wavelength of the incident wave, we assume that the surface is a very rough normally distributed surface with a Gaussian correlation function. We can use Kirchhoff rough surface theory to include the surface effects. We choose l/σ about 10, where l is the correlation length and σ is the root-mean-square height of the surface.

A summary of the parameters we use:

$$\lambda = 0.6328\,\mu, \quad ka \sim 1, \quad \tau \sim 10, \quad l/\sigma \sim 10,$$

$$n_1 = n_3 = 1, \quad n_2 = 1.33, \quad n_0 = n_0^r + in_0^i, \quad n_0^r \sim 1.4.$$

Theoretical Calculation

Two scattering processes occure when light interacts with the leaf: multiple scattering within the leaf and rough surface scattering. The multiple scattering process can be evaluated by the radiative transfer theory. This theory assumes that the fields are uncorrelated, so the resulting energy under consideration can be obtained by simple linear summation. Then, by considering energy conservation, we obtain the integral-differential radiative transfer equation governing the propagation of the specific intensity in an isotropic medium for a plane wave normally incident on a slab of random medium with refractive index different from that of the surrounding media ($z < 0$ and $z > z_0$) (*Ma, et al.*, 1989):

$$\mu \frac{d[I(\tau,\hat{s})]}{d\tau} = -[I(\tau,\hat{s})] + \int_{4\pi} [S][I(\tau,\hat{s}')]\,d\omega' + C_p[F^+]e^{-\tau} + C_p R_{23}[F^-]\,e^{-2\tau_0+\tau}, \quad (1)$$

where

$[I(\tau,\hat{s})]$ is the 4×1 incoherent specific intensity matrix,
τ is the optical distance and is given by $\tau = \rho_n \sigma_t L$,
ρ_n is the particle number density,
σ_t is the total scattering cross section per particle,
$\hat{s}$ is the observation direction given by $\hat{s} = l\hat{x} + m\hat{y} + n\hat{z}$.
($l = \sin\theta\cos\phi$, $m = \sin\theta\sin\phi$, $n = \cos\theta$.)
$[S]$ is the 4×4 Mueller scattering matrix.

$$C_p = \frac{T_{12} I_0}{1 - R_{23} R_{21} exp(-2\tau_0)},$$

where T_{12} and R_{21} are the transmissivity and reflectivity at the boundary $\tau = 0$, and R_{23} is the reflectivity at the boundary $\tau = \tau_0$.

The boundary conditions for this problem are:
at $\tau = 0$, for $0 \le \mu \le 1$,

$$[I(0,\mu)] = [R(\mu)]_{21}\,[I(0,\mu)], \quad (2)$$

at $\tau = \tau_0$, for $-1 \le \mu \le 0$,

$$[I(\tau_0,\mu)] = [R(\mu)]_{23}\,[I(\tau_0,\mu)]. \quad (3)$$

To solve this integral-differential equation with its boundary conditions, first we expand $[I]$ and $[S]$ in a Fourier series in ϕ and making use of orthogonality, we obtain two uncorrelated integral-differential equations for $m = 0$ and $m = 2$. All other Fourier components of $[I]_m$ are zero. Then, a discrete-ordinate method and an eigenvalue-eigenvector technique are used to obtain numerical solutions (*Ishimaru*, 1978).

The normalized scattering cross section per unit area $\sigma^0_{\alpha\beta}$ is defined by

$$\sigma^0_{\alpha\beta} = \frac{4\pi I_{\alpha\beta}}{I_0}. \quad (4)$$

where α and β are respectively the scattered and incident polarizations and I is the specific intensity.

Now, we consider the surface scattering. The general and exact solution of this problem is yet unknown. But, for our problem the Kirchhoff rough surface scattering theory (*Beckman*, 1963) is applicable because of the short incident wavelength. Under Kirchhoff's assumption, Stogryn (1967) derived the incoherent reflected intensity for a horizontally polarized plane wave incident on a dielectric rough surface. To compare with experimental results we convert the formulas into the following simple forms for normal incidence:

$$\sigma^0_{hh} = \frac{l^2 |r^p_{21}|^2}{\sigma^2 (1+\cos\theta)^2} exp\left[-\frac{l^2 \sin^2\theta}{4\sigma^2(1+\cos\theta)^2}\right], \quad (5)$$

$$\sigma^0_{vh} = 0.$$

where

θ is the scattering angle and $90° < \theta < 180°$,
r^p_{21} is the field reflection coefficient for the horizontally polarized component from medium 2 to 1.

Because in our assumption $\sigma/\lambda \gg 1$, we can neglect the reflected coherent intensity.

Combining (4) and (5), we obtain the total normalized scattering cross section per unit area for the transmitted and reflected intensity for a normally incident plane wave with horizontal polarization as:

for $0 < \theta < 90°$,

$$\sigma^0 = 4\pi\langle I \rangle, \tag{6}$$

for $90° < \theta < 180°$,

$$\sigma^0 = 4\pi\langle I \rangle + \frac{l^2|r_{21}^P|^2}{\sigma^2(1+\cos\theta)^2} \exp\left[-\frac{l^2\sin^2\theta}{4\sigma^2(1+\cos\theta)^2}\right] \begin{pmatrix} 0 \\ 1 \\ 0 \\ 0 \end{pmatrix}, \tag{7}$$

Experimental Set-Up

A simplified block diagram of the experimental set-up is shown in Fig.2. A horizontally polarized He-Ne laser beam which is chopped at 400Hz with a diameter of about 15mm is normally incident on the leaf. The leaf is mounted at the center of a rotational stage. The detector consists of a 4mm receiving aperture , a set of movable polarizers and an optical fiber. The scattered light from the leaf is focused into the optical fiber and is amplified by the cooled photomultiplier. The whole detector is controlled by a stepping motor which is used to rotate the detector in the scattering plane. Thus, it allows us to measure the scattered intensities versus different scattering angles. All of the control and data acquisition are done with an IBM-PC computer.

The separation of the cross-polarization of the polarizer is at least 30 dB and the noise level of the system is 75 dB below the free space incident intensity. The field of view of the receiver is 0.15 degrees. The transmitted copolarized intensity HH and cross-polarized intensity VH are obtained for scattering angles of 0 to 50 degrees and the reflected intensities for 110 to 175 degrees with a resolution of 5 degrees. All data are normalized with respect to the free space normally incident intensity.

When we measure the scattering pattern of the leaf, the shiny side is always facing toward the incident beam. In order to obtain a statistical average, the scattered intensity at each angle is measured 1000 times at the different parts of the leaf. It is noticed that the local tilt of the leaf surface usually affects the form of the scattered intensity. Measurements are made for corn, potato and laurel leaves.

In order to compare the experimental data with theoretical results, the measured scattering intensity $\langle I \rangle_{\text{intensity}}$ has to be converted into specific intensity $\langle I \rangle_{\text{specific}}$ to obtain the normalized scattering cross section per unit area σ^0. Since the viewing angle of the receiver is very small ($\Delta\theta_0 = 0.15°$), the following formula can be used,

$$\langle I \rangle_{\text{intensity}} = \pi(\Delta\theta_0)^2 \langle I \rangle_{\text{specific}}. \tag{8}$$

Therefore, σ^0 is given by

$$\sigma_0 = \frac{4\pi\langle I \rangle_{\text{specific}}}{I_0} = \frac{4}{(\Delta\theta_0)^2}\langle I \rangle_{\text{intensity}}. \tag{9}$$

Comparison of the results

All parameters used in the theory are tabulated in Table I for corn, potato and laurel leaves. The transmitted ($0° < \theta < 90°$) and reflected ($90° < \theta < 180°$) copolarized and cross-polarized intensities versus azimuth angles for a normally incident horizontally polarized laser light have been measured and converted to the normalized scattering cross section per unit area for corn, potato and laurel leaves. The corresponding theoretical and experimental results for the normalized scattering cross section per unit area $\sigma^0_{\alpha\beta}$ are shown in Figs.3 to 5. The solid lines are for the numerical results of theory and the dashed lines are for the measured data. From these results we see that the linearly polarized incident wave has been depolarized and that there is a peak in the reflected copolarized component due to rough surface scattering. The difference between the theoretical and experimental results is caused by several factors, such as the flatness and the fiber structure of the leaf, which affect the form of the intensities. Besides, from the measurements for different leaves of the same species, we noticed that the intensities can vary about 1 dB due to a slight difference in the thickness, the fiber structure, the water content, etc..

Comparing the reflected copolarized components HH from Fig.3 to 5 for these three kinds of leaves, we see that the surface of the potato leaf ($l/\sigma = 7$) has greater roughness than that of the corn ($l/\sigma = 8$) or laurel ($l/\sigma = 11$). This result supports both theory and experiment because visual observation of these leaves gives the same result.

CONCLUSIONS

An optical scattering model for a single leaf has been developed by combining multiple scattering within the leaf and rough surface scattering. The theoretical results were compared with the experimental results and show good agreement. The results show that the scattering within the leaf is mainly caused by the scatterers whose sizes are comparable to the incident optical wavelength, i.e. ka~1. The leaf is optically thick ($\tau \sim 10$) and shows strong depolarization effect on the optical wave. The optical parameters obtained from the scattering model for corn, potato and laurel leaves are given in Table I. From Table I we know that laurel leaf is the optically thickest of the three and its surface is the smoothest.

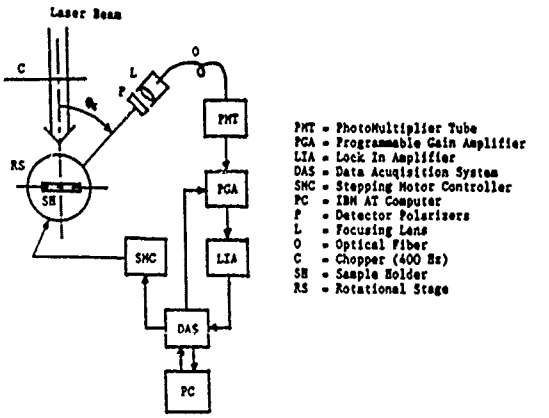


Fig.2 Experimental set-up.

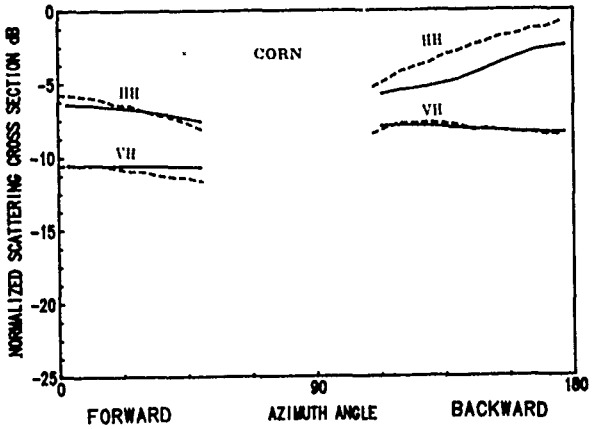


Fig.3: Comparison of the theoretical and experimental results for the corn leaf. The solid lines are for the theoretical results and the dashed lines are for the experimental results. $\phi = 0$, $n_0 = 1.42 + i0.0009$, ka=2, $\tau = 9$, $l/\sigma = 8$.

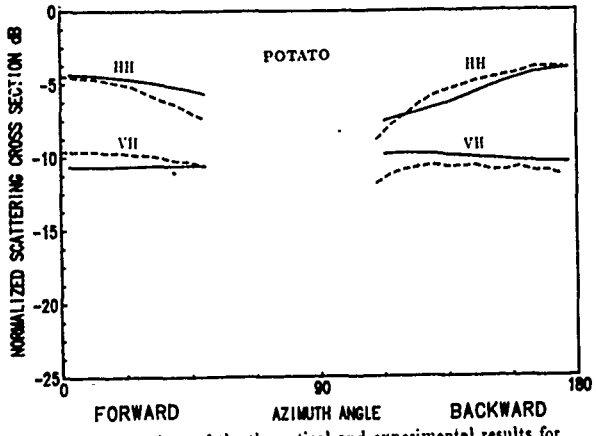


Fig.4: Comparison of the theoretical and experimental results for the potato leaf. The solid lines are for the theoretical results and the dashed lines are for the experimental results. $\phi = 0$, $n_0 = 1.41 + i0.001$, ka=2.8, $\tau = 10$, $l/\sigma = 7$.

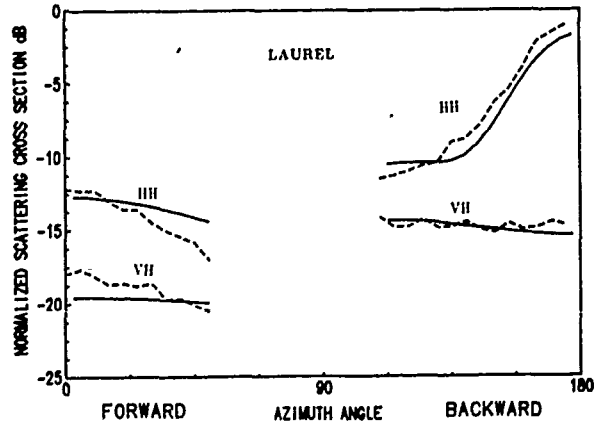


Fig.5: Comparison of the theoretical and experimental results for the laurel leaf. The solid lines are for the theoretical results and the dashed lines are for the experimental results. $\phi = 0$, $n_0 = 1.40 + i0.0016$, ka=2.5, $\tau = 11.5$, $l/\sigma = 11$.

TABLE I
OPTICAL PARAMETERS OF CORN, POTATO AND LAUREL LEAVES

Leaf Name	Corn	Potato	Laurel
ka	2	2.8	2.5
τ	9	10	11.5
n_0^r	1.42	1.41	1.40
n_0^i	0.0009	0.001	0.0016
l/σ	8	7	11

** For all the calculations $\lambda = 0.6328\mu$, $n_1 = n_3 = 1$, $n_2 = 1.33$ are chosen.

ACKNOWLEDGMENT

This work was supported by the U.S. Army Engineer Waterways Experiment Station, the National Science Foundation and the Army Research Office.

REFERENCES

Allen, W. A., H. W. Gausman, and A. J. Richardson, "Mean effective optical constants of cotton Leaves", *J. Opt. Soc. Amer.* 60, pp. 542-547, 1970.

Allen, W. A. and A. J. Richardson, "Interaction of light with a plant canopy", *J. Opt. Soc. Amer.*, 58, pp. 1023-1028, 1968.

Beckman, P., *The Scattering of Electromagnetic Waves from Rough Surfaces*, The Macmillan Company, New York, pp. 70, 1963.

Fung, A. K. and F. T. Ulaby, "A scatter model for leafy vegetation", *IEEE Trans.* GE-16, pp. 281-286, 1978.

Gausman, H. W. and W. A. Allen, "Optical parameters of leaves of 30 plant species", *Plant Physiol.* 52, pp. 57-62, 1973.

Ishimaru, A., *Wave Propagation and Scattering In Random Media*, Academic Press, New York. Vol. I, pp. 195, 1978.

Ma, Q., A. Ishimaru, and Y. Kuga, "Scattering and depolarization of waves incident upon a slab of random medium with refractive index different from that of the surrounding media", Submitted to Radio Science, 1989.

Sekera, Z., "Scattering matrices and reciprocity relationships for various representations of the state of polarization", *J. Opt. Soc. Am.* 56, pp. 1732, 1966.

Stogryn, A., "Electromagnetic scattering from rough, finitely conducting surface", *Radio Science*, Vol. 2 (New Series), No. 4, pp. 415-428, 1967.

Tsang, L. and J. A. Kong, "Application of strong fluctuation random medium theory to scattering from vegetation-like half space", *IEEE Trans.* GE-19. pp. 62-69, 1981.

Ulaby, F. T. and E. A. Wilson, "Microwave attenuation properties of vegetation canopies". *IEEE Trans.* GE-23. pp. 746. 1985.

DETECTION OF STRESS OF CONIFEROUS FOREST TREES WITH THE VIRAF SPECTROMETER

C. Buschmann, U. Rinderle and H.K. Lichtenthaler

Botanical Institute II (Plant Physiology and Plant Biochemistry)
University of Karlsruhe, Kaiserstr. 12, D-7500 Karlsruhe (FRGermany)
Tel.: 0049/721/608-4166 Telefax: 721/608 4290

ABSTRACT

The newly constructed **VIRAF**-spectrometer (**V**isible **I**nfrared **R**eflectance **A**bsorbance **F**luorescence) spectrometer has been applied in outdoor measurements to compare needles of two spruce trees differently affected by the forest decline in the Northern Black Forest. Changes of the reflectance signatures due to loss of chlorophyll or other damage effects are interpreted with the help of other techniques currently applied for the evaluation of physiological activity. The VIRAF-spectrometer is used to examine the spectra of reflectance, absorption and of chlorophyll fluorescence. The measurements are carried out with the same sample without changing its position. By means of a two-wavelength fluorometer and of a PAM-fluorometer the fluorescence induction kinetics of the needles were determined as an indicator of vitality. In addition, the rate of photosynthetic CO_2-fixation and the stomatal closure was studied using a CO_2/H_2O-porometer. The VIRAF-measurements proved to be an excellent method of physiological ground-truth and vitality test which can well be associated with remote sensing of forest decline.

Key words: reflectance signatures, chlorophyll fluorescence, photosynthetic activity, forest decline, ground truth

1. INTRODUCTION

In order to be able to interpret reflectance signatures of damaged forest trees on a physiological basis, the VIRAF-spectrometer (**V**isible **I**nfrared **R**eflectance **A**bsorbance **F**luorescence) was newly constructed (detailed description: Buschmann et al., 1988; Buschmann and Lichtenthaler, 1988). It allows the computer-controlled, non-destructive measurement of the spectra of reflectance, absorption and fluorescence emission with one sample without changing its position and without removing needles or leaves from the plant. It has been demonstrated before that this instrument can be applied to study a variety of stress symptoms affecting photosynthesis or other physiological parameters (Buschmann et al., 1988; Buschmann and Lichtenthaler, 1988a and 1988b; Buschmann et al., 1989). This study presents outdoor measurements carried out by means of the VIRAF-spectrometer and in parallel by fluorescence instruments (two-wavelength fluorometer, PAM-fluorometer) and a CO_2/H_2O-porometer.

The different methods used give individual information on the following parameters: The reflectance spectra of leaves reflect their pigment content and tissue structure (Lichtenthaler and Buschmann, 1987; Buschmann et al., 1989; Lichtenthaler, 1989). The reflectance spectra of the VIRAF-spectrometer are taken by illuminating with strong white light. Thus fluorescence emission contributes to the reflectance spectrum (Lichtenthaler and Buschmann, 1987; Buschmann et al., 1989). Fluorescence spectra and the ratio F690/F735 determined with the VIRAF-spectrometer are indicators of chlorophyll concentration and photosynthetic activity (Lichtenthaler, 1987; Lichtenthaler and Rinderle, 1988a and 1988b). From the fluorescence induction kinetics the Rfd-values (Rfd = ratio fluorescence decrease) were calculated as indicator of the potential photosynthetic capacity and as a vitality index (Lichtenthaler and Rinderle, 1988a). Rfd-values usually go parallel with the net CO_2-assimilation rates, but are independent of the stomata closure. The new PAM-fluorometer (**P**ulse **A**mplitude **M**odulation) permits determination of the photochemical and the non-photochemical quenching coefficients of fluorescence which are affected by the photosynthetic electron transport chain and photosynthetic ATP-formation (Schreiber et al., 1986; Lichtenthaler and Rinderle, 1988a). Information on the net CO_2-assimilation and transpiration as well as stomatal opening or closure is obtained by means of a CO_2/H_2O-porometer.

Our outdoor study was carried out to demonstrate the potential capability for detecting damage and stress effects of trees affected by the forest decline by means of the VIRAF-spectrometer. In order to further facilitate the interpretation of reflectance signatures on a physiological basis the VIRAF-measurements were compared with other physiological measurements, e.g. induction kinetics of fluorescence and the Rfd-values as vitality index, the fluorescence quenching processes, net photosynthesis and stomatal closure. The application of the VIRAF-spectrometer in physiological ground truth measurements is discussed in the light of the other methods.

2. MATERIAL AND METHODS

Needles of two, approx. 50-year-old spruce trees (*Picea abies*) in the Northern Black Forest (Freudenstadt, Schöllkopf, 900 m above sea level) were examined in outdoor measurements. The needles of a tree without visible damage symptoms (little needle loss and discolouring) were compared to needles of a damaged tree standing about 150 m distant. The measurements were performed on the 11th November 1988 with the last three needle years 1986, 1987 and 1988.

The **VIRAF-spectrometer** was used to determine reflectance spectra and emission spectra of the chlorophyll fluorescence. The reflectance between 400 and 800 nm was measured by illuminating the sample perpendicularly with strong white light and analyzing the light reflected perpendicularly from the sample surface (Buschmann et al., 1988; Buschmann and Lichtenthaler, 1988). Subsequently, fluorescence spectra were determined between 650 and 800 nm. The optical pathway resembled that of the reflectance, except that the fluorescence was excited with a broad-band blue light (470 $\pm$ 50 nm). Figure 1 describes the measuring set up of the VIRAF-spectrometer for reflectance, transmission (absorption) and fluorescence studies.

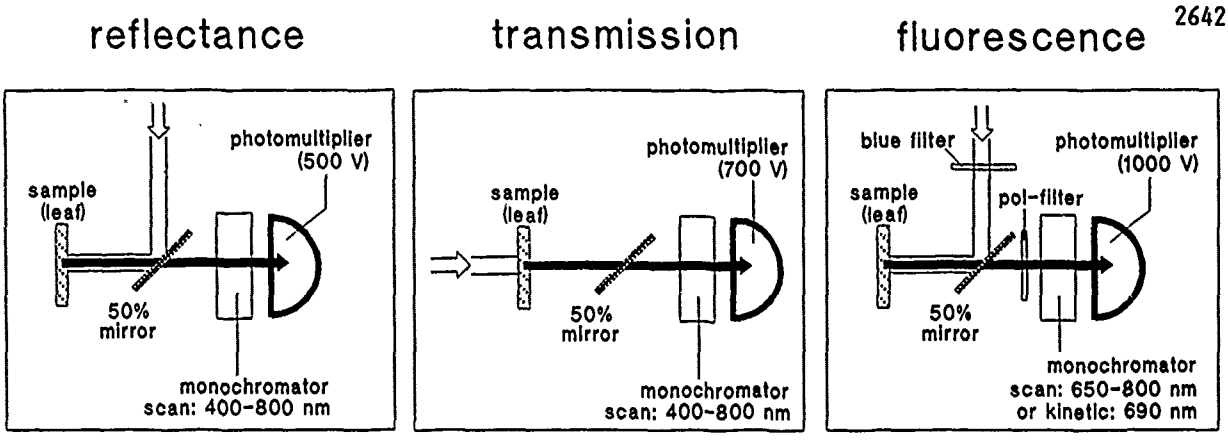


Fig. 1: Scheme of the optical pathway for measuring reflectance, absorption and chlorophyll fluorescence by means of the VIRAF-spectrometer.

These three types of measurements can be carried out without changing the position of the sample and with the leaf or needles left intact attached to the plant. From the reflectance spectra the 2nd derivative was calulated and the IR/R ratio (ratio of the reflectance signal at 800 and 680 nm) was determined. From the fluorescence spectra the ratio F690/F735 was quantified as ratio of the intensity at the two wavelengths.

Induction kinetics of the chlorophyll fluorescence were measured simultaneuously in the 690 and 730 nm regions by means of a portable **two-wavelength field fluorometer** (Lichtenthaler and Rinderle, 1988a). Excitation was performed with a He/Ne-laser (632.8 nm). The Rfd-values were calculated from the kinetics as the ratio between fluorescence decrease fd and the steady-state fluorescence fs (Rfd = fd/fs). From the Rfd-values at 690 and 730 nm the stress adaptation index Ap was calculated according to Lichtenthaler and Rinderle (1988a). The ratio F690/F730 of the fluorescence maxima was determined from the steady-state of the induction kinetics reached 5 min after the onset of illumination.

The differentiation between photochemical quenching (qQ) and non-photochemical energy quenching (qE), as well as the ratio of maximum fluorescence Fm to the ground fluorescence Fo were determined using the PAM-fluorometer of Walz (Schreiber et al., 1986).

Photosynthetic CO_2-fixation (P_N) and stomatal conductivity (gH_2O) were determined using the CO_2/H_2O-porometer of Walz (Schulze et al., 1982).

Chlorophylls (a+b) and total carotenoids (x+c) were determined from the needle extracts with 100% acetone using the new extinction coefficients of Lichtenthaler (1987b).

3. RESULTS AND DISCUSSION

Chlorophyll and carotenoid content: In contrast to the needles of the healthy spruce, the needles of the damaged spruce possessed a lower chlorophyll content, except for the youngest needle year 1988 (Table 1). There were no significant changes in the chlorophyll a/b-ratio. The (a+b)/(x+c)-ratio of the green (chlorophylls) to yellow pigments (carotenoids) was significantly decreased in the older needle years of the damaged spruce. This indicates that the chlorophylls in the damaged spruce were less stable and - in contrast to the carotenoids - were preferentially destroyed. The lower chlorophyll content may not solely be due to a degradation of chlorophyll but partially also to a lower accumulation rate.

Measurements with the VIRAF-spectrometer: The reflectance spectra of green spruce needles (Fig. 2, upper part) show the typical shape of green leaves: a low reflectance in the visible range (400 - 700 nm) with the exception of a small increase in the green-light region around 560 nm and a sharp rise of the reflectance above 680 nm towards longer wavelengths ("red edge"). Needles of damaged spruces, characterized by a lower chlorophyll content (Table 1), in general have a higher reflectance in the visible range, with a further increase in the green and yellow region of the spectrum as compared to fully green spruce needles. The rise of the "red edge" starts at a shorter wavelength. This is quantified by a "blue shift" of the inflection point (I.P. in Table 1) in the intersection of the 2nd

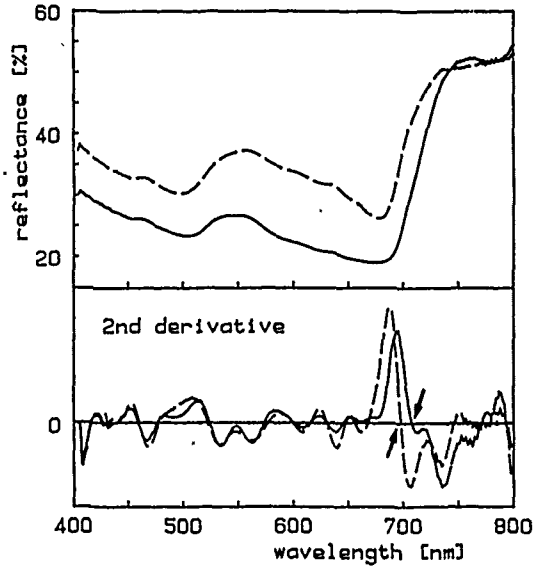


Fig. 2. Reflectance spectra (upper part) of two-year-old spruce *(Picea abies)* needles taken from a healthy (continuous line) and a damaged tree (dashed line) growing in the Northern Black Forest. The spectra are means of 3 independent experiments measured with the VIRAF-spectrometer. From the reflectance spectra the second derivative has been calculated (lower part). The inflection point of the "red edge", i.e. the intersection with the zero line, is indicated by arrows.

derivative with the zero line (arrows in Fig. 2, lower part). In the early 1980s, up to about 1985, the reflectance spectra of damaged spruces were also characterized by a much lower reflectance in the near infrared region around 800 nm (Buschmann and Lichtenthaler, 1988a and 1988b; Lichtenthaler, 1989). This decrease in the reflectance in damaged trees is much lower in recent years, which provided rather favorable conditions for needle growth. Yet the ratio IR/R of the 800 to 680 nm regions of the reflectance spectrum was lower for the 1987 and 1986 needles of the damaged spruce (Table 1).

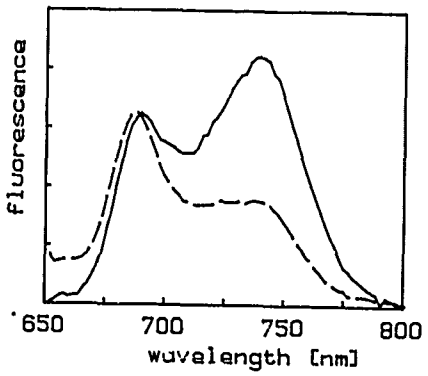


Fig. 3. Fluorescence-emission spectra of two-year-old spruce *(Picea abies)* needles taken from a healthy (continuous line) and a damaged tree (dashed line) growing in the Northern Black Forest. The spectra are means of 3 independent experiments measured with the VIRAF-spectrometer.

The fluorescence-emission spectra of spruce needles measured with the VIRAF-spectrometer showed two maxima (Fig. 3). The fluorescence intensity in general and especially the short-wavelength maximum become higher when the needles are photosynthetically inhibited and/or the chlorophyll concentration is lower (Lichtenthaler et al., 1986; Lichtenthaler and Rinderle, 1988b), as is the case in the needles taken from the damaged tree. This change was quantified by the ratio F690/F735, which is determined as the ratio between the intensities at the two fluorescence maxima. Since the absorption of the chlorophyll overlaps with its fluorescence spectrum, the fluorescence at the short-wavelength side is decreased with increasing chlorophyll concentration (Lichtenthaler et al., 1986; Lichtenthaler and Rinderle, 1988b). In addition, inhibition of photosynthesis also leads to a higher fluorescence or heat (Buschmann, 1986). Therefore the ratio F690/F735 represents a good stress indicator, which provides information on the chlorophyll content and does not require pigment extraction. The fluorescence ratio F690/F735, as induced by the blue light (VIRAF-measurements), was significantly higher in the older needles of the damaged spruce than in the healthy tree and reflects their lower chlorophyll content (Table 1).

Measurements with the portable two-wavelength fluorometer: A similar fluorescence ratio (F690/F730) can be calculated from the fluorescence induction kinetics measured with the two-wavelength field fluorometer. The values of F690/F730, as determined from the red-laser-light (632.8 nm) induced kinetics, are much lower than those calculated from the blue-light-induced fluorescence-emission spectra. That they differ in the height of their maxima in the 690 and 735/730 nm region has been shown before (Lichtenthaler and Rinderle, 1988b). In the older needles of the damaged spruce the values of F690/F730 are significantly increased to about the same degree as the values of the ratio F690/F735 determined with the VIRAF-spectrometer (Table 1). This comparison indicates that one obtains via the ratio F690/F730 or F690/F735 similar information on changes of the chlorophyll content of the needles due to damage or stress, even though the absolute values are different.

From the chlorophyll fluorescence induction kinetics measured with the two-wavelength fluorometer we obtained Rfd-values for the 690 and the 730 nm regions. The Rfd 730-values were always lower than the Rfd 690-values (Table 1). Rfd-values of

Table 1. Parameters of spruce *(Picea abies)* needles of 1986 taken from a healthy and a damaged tree (Northern Black Forest, Freudenstadt). The values were determined from measurements with the VIRAF-spectrometer, the portable two-wavelength fluorometer, the PAM-fluorometer and the CO_2/H_2O-porometer. The chlorophyll (a + b) and carotenoid (x + c) content was evaluated from a 100% acetone extract. Mean of 3 determinations per each needle year. Maximum deviation 5% or less for most of the parameters. The position of the inflection point I.P. varied by ± 2nm.

	pigments			reflectance		chlorophyll fluorescence parameters								CO_2-assimilation		stomata
						VIRAF	two-wavelength fluorometer				PAM-fluorometer					
	a+b[a]	a/b	a+b / x+c	IR/R	I.P.	F690[b] / F735	F690[c] / F730	Rfd 690	Rfd 730	Ap	Fm / Fo	qQ	qE	P_N** (per area)	P_N** (per a+b)	gH_2O***
1988 needles																
healthy tree	49	3.3	4.5	3.3	707	0.86	0.36	4.6	3.0	0.29	5.1	0.89	0.26	2.8	0.9	38
damaged tree	54	3.4	4.6	3.1	707	0.82	0.34	4.0	2.6	0.28	4.9	0.82	0.41	3.7	1.1	52
1987 needles																
healthy tree	63	3.2	4.6	3.1	704	0.86	0.39	3.8	2.6	0.25	5.9	0.85	0.23	2.5	0.7	33
damaged tree	40	3.0	4.0	2.7	700	1.15	0.49	3.3	2.5	0.19	4.6	0.81	0.46	3.4	1.4	62
1986 needles																
healthy tree	71	3.0	4.7	3.1	707	0.77	0.35	3.6	2.5	0.24	5.7	0.82	0.30	1.2	0.3	18
damaged tree	23	3.0	2.9	2.2	697	1.87	0.76	3.0	1.9	0.28	3.3	0.75	0.58	2.0	1.4	46

a) µg chlorophyll a+b per cm^2 needle area.
b) determined from the blue-light-induced fluorescence spectra measured with the VIRAF-spectrometer
c) determined from the He/Ne-laser-induced fluorescence kinetics measured in the 690 and 730 nm regions with the two-wavelength fluorometer
** net CO_2-assimilation rate P_N per area in µmol $CO_2 \cdot m^{-2} \cdot s^{-1}$ and P_N in mg $CO_2 \cdot mg\ (a+b)^{-1} h^{-1}$.
*** Stomatal conductivity gH_2O in mmol $H_2O \cdot m^{-2} \cdot s^{-1}$

the needles of the damaged spruce were only slightly lower than those of the healthy spruce and indicated good photosynthetic activity of the remaining chlorophyll in the stressed tree. The values of the stress-adaptation index Ap of 0.19 or much higher (Table 1) indicated that despite the lower chlorophyll content no irreversible damage had occurred.

Measurements with the PAM-fluorometer: The PAM-fluorometer allowed determination of the ground fluorescence Fo and maximum fluorescence Fm in the 730 nm region. The ratio Fm/Fo, which can be considered as an additional stress indicator, was reduced in the needles of the damaged spruce in a similar way to the Rfd-values (Table 1). The photochemical quenching coefficient qQ was about the same in needles of the healthy and the damaged spruce and appears not to be a suitable indicator of stress or damage to plants (see also Rinderle and Lichtenthaler, 1989). The non-photochemical quenching coefficient qE is, however, considerably higher in all needle years of the damaged spruce (Table 1).

Measurements with the CO_2/H_2O-porometer: The net CO_2-assimilation P_N (per cm^2 needle area or per chlorophyll a+b), as determined with the CO_2/H_2O porometer, showed lower values for the needles of the healthy spruce than those of the damaged tree (Table 1). The higher photosynthetic rates of the damaged spruce were paralleld by higher values of the stomatal conductivity gH_2O (Table 1). This indicates that at the measuring day (beginning November 1988) the stomata of the needles of the damaged tree were still more open than those of the healthy tree. The needles of the healthy spruce were apparently in the stage of incipient winter dormancy, though the high Rfd-values indicated that the internal photosynthesis was fully active. The damaged spruce, in turn, seemed still to be increasing their carbohydrate reserves for the winter-time. Reduced frost hardening of damaged trees is a well-known phenomenon associated with forest decline.

4. CONCLUSION

Our field studies with the newly constructed VIRAF-spectrometer and the portable two-wavelength fluorometer indicate that both instruments can be used with great success for physiological ground truth and vitality tests in connection with remote sensing of terrestrial vegetation. The advantage of the VIRAF-spectrometer is the fact that complete reflectance, absorption and fluorescence spectra can be taken. From the reflectance spectra one can determine the "blue shift" of the inflection point of the "red edge". It also allows determination of the fluorescence ratio F690/F735, as a major stress indicator. The two-wavelength fluorometer permits determination of the Rfd-values as a vitality index, the ratio F690/F730 and the stress adaptation index Ap. It is also very suitable for ground-truth measurements, but does not provide information on reflectance changes of leaves. Though both instruments give complementary information on leaves during stress appearance and regeneration, they can be used separately for outdoor screening of the physiological state of plants. The results also demonstrate that a combination of reflectance data with other physiological measurements facilitates further interpretation of damage and stress effects in plants.

Acknowledgements: This work was sponsored by grants of the Europäisches Forschungszentrum für Maßnahmen zur Luftreinhaltung (PEF), Karlsruhe which is gratefully acknowledged. We thank F. Stober, S. Rang, S. Burkart and A. Fahl for skilful technical assistance.

REFERENCES

Buschmann, C., "Fluorescence and heat emission by plants - Application in photosynthesis research", Naturwissenschaften 73, pp. 691-699, 1986.

Buschmann, C. and Lichtenthaler, H.K., "Correlation of reflectance and chlorophyll fluorescence signatures of healthy and damaged forest trees", Proceedings of IGARSS '88, Vol. 3, pp 1339-1342, ESA Publications Division, Noordwijk 1988a.

Buschmann, C. and Lichtenthaler, H.K., "Reflectance and chlorophyll fluorescence signatures of leaves", Applications of Chlorophyll Fluorescence, Lichtenthaler, H.K., ed., pp. 325-332, Kluwer Academic Publishers, Dordrecht 1988b.

Buschmann, C., Rang, S., Stober, F. and Kocsányi, L., "Reflectance spectra of leaves and needles as a basis for the physiological evaluation of forest damage. - I. Measuring set up", PEF Report 35, pp 191-196, 1988.

Buschmann, C., Rang, S., Stober, F. and Kocsányi, L., "Reflectance spectra of leaves and needles as a basis for the physiological evaluation of forest damage. - II. Basic measurements", PEF Report (in print), 1989.

Lichtenthaler, H.K., "Chlorophyll fluorescence signatures of leaves during the autumnal chlkorophyll breakdown", Journal of Plant Physiology 131, pp. 101-110, 1987a.

Lichtenthaler, H.K., "Chlorophylls and carotenoids, the pigments of the photosynthetic biomembanes", Methods in Enzymology 148, pp. 350-382, 1987b.

Lichtenthaler, H.K., "Remote sensing of chlorophyll fluorescence in oceanography and in terrestrial vegetation: an introduction", Applications of Chlorophyll Fluorescence, Lichtenthaler, H.K., ed., pp. 287-297, Kluwer Academic Publishers, Dordrecht 1988.

Lichtenthaler, H.K., "Possibilities for remote sensing of terrestrial vegetation by a combination of reflectance and laser-induced chlorophyll fluorescence", Proceedings of IGARSS '89, this volume, 1989.

Lichtenthaler, H.K. and Buschmann, C., "Reflectance and chlorophyll fluorescence signatures of leaves", Proceedings of IGARSS '87, Vol. 2, pp. 1207-1212, University of Michigan, Ann Arbor 1987.

Lichtenthaler, H.K., Buschmann, C., Rinderle, U. and Schmuck, G., "Application of chlorophyll fluorescence in ecophysiology", Radiation and Environmental Biophysics 25, pp. 297-308, 1986.

Lichtenthaler, H.K. and Rinderle, U., "The role of chlorophyll fluorescence in the detection of stress conditions in plants", CRC Critical Reviews in Analytical Chemistry 19, Suppl. I, pp. S29-S85, 1988a.

Lichtenthaler, H.K. and Rinderle, U., "Chlorophyll fluorescence spectra of leaves as induced by blue light and red laser light", 4th Internat. Colloquium on Spectral Signatures of Objects in Remote Sensing, Aussois, pp. 251-254. ESA Publications Division, Noordwijk 1988b.

Rinderle, U. and Lichtenthaler, H.K., "The various chlorophyll fluorescence signatures as a basis for physiological ground truth control in remote sensing in forest decline", Proceedings of IGARSS '89, this volume, 1989.

Schreiber, U., Schliwa, U. and Bilger, W., "Continuous recording of photochemical and non-photochemical chlorophyll fluorescence quenching with a new type of modulation fluorometer", Photosynth. Res. 10, pp. 51-62, 1986.

Schulze, E.D., Hall, A.E., Lange, O.L. and Walz, H., " A portable steady state porometer for measuring the carbon dioxide and water vapour exchange of leaves under natural conditions", Oecologia 53, pp. 141-154, 1982.

COMPARISON OF SPECTRAL DATA GATHERED FROM A LABORATORY SPECTROMETER AND TM IMAGES WITH AND WITHOUT SHADOW CORRECTION

Richard L. Thiessen and Jay R. Eliason*

Geology Department, Washington State University, Pullman, WA 99164-2812, USA; * - also at Geologic Analysis and Consulting Services, P.O. Box 315, Deary, ID 83823.

ABSTRACT

Shading of topographic features due to varying sun angle has long been a problem with computer classification of remotely collected imagery. Reflectance intensity is a function of the material type, atmospheric effects, topographic orientation, and solar illumination angle. Identification of surface materials from their spectral reflectance characteristics is one of the major goals in digital analysis of imagery. Atmospheric effects and shadows due to topography and solar illumination will create variations in spectral reflectance intensities. These effects result in a spreading out of the DN value histograms observed for the spectral bands, leading to poorly defined classification statistics. In order to analyze these effects, we examined Thematic Mapper (TM) imagery and rock hand sample spectral data from the Paiute Ridge Quadrangle at the Nevada Test Site. The TM imagery used for this study was collected in March providing a relatively low sun angle. Staff at Battelle PNL calculated a Minnaert reflectance model of the ground surface based upon the topography and the sun angle. This reflectance model was used to remove topographic reflectance variations from the original TM image for bands 1, 4, 5 and 7 creating a corrected image. A June TM image, with a relatively high sun angle, was also utilized. Histograms of each TM band were prepared for training areas of each lithologic unit. The original March image histograms showed a wider distribution of data points because of the relatively severe shadowing. Histograms from the June image showed narrower spreads for most of the training areas, indicating a lesser shadowing effect. The corrected March image produced more tightly defined histograms. For a computer classification, the tighter histograms would yield better statistics with less overlap of the histograms of the individual training areas. This would lead to an improved computer classification.

Spectral reflectance data from field samples were determined with a laboratory spectrometer (Beckman DK-2A). The spectral curves obtained with the spectrometer were correlated with the histograms determined from the images. The tightly defined histograms from the shadow corrected TM provided the best correlation with the rock data.

Several units, including the Rainier Mesa Member of the Timber Mountain Tuff, showed multiple spectral patterns on both images and rock spectra. This difference was evaluated versus geochemistry, hematitic alteration, devitrification, pumice content, and degree of welding.

Keywords: Spectral reflectance, Histograms, Minnaert reflectance model, Nevada Test Site, Rainier Mesa Member.

INTRODUCTION

The Paiute Ridge quadrangle of the Nevada Test Site (NTS) has been a site of ongoing joint Battelle PNL-Washington State University (WSU) research (Foote and others, 1985, Thiessen and others, 1987) sponsored by the Office of Basic Energy Sciences, Department of Energy. Geologic analysis of PNL prepared imagery at WSU has led to an excellent knowledge base of the structures and lithologies at the site. A field trip to NTS served three purposes. First, the image interpretations, including several modifications to the U.S.G.S. quadrangle map (Byers and Barnes, 1967) were field checked. Second, the orientations of naturally occurring joints and fractures were measured in order to compare to ones determined directly from digital elevation models using our Geologic Spatial Analysis system (Eliason, 1984; Eliason and Thiessen, 1987; Thiessen and others, 1987, 1989) which automatically determines the full three-dimensional spatial orientation of the structures. Third, a suite of rock samples was collected for spectroscopic, hand sample, thin section, and chemical analysis. Due to the logistics and time constraints of the field trip, not every lithologic unit was sampled, but for those that were, several rocks were collected from each. The purpose of this report is to present the spectral response curves of some of these rocks and compare those to each lithologic unit's characteristics on Landsat Thematic Mapper imagery. This will be a test of the applicability of the technique on the Nevada Test Site and give a reference library of spectra for the lithologies present. In addition, this will test the characteristics of three different images: a March image on which there are severe shadowing effects due to a low sun angle; the March image with those shadowing effects corrected out using a technique developed by Harlan Foote (of Battelle PNL); and a June image, which, with its higher sun angle, has a lesser shadowing effect.

METHODOLOGY

The hand samples from the study area were cataloged and an overlay was prepared for the published 1:24,000 scale geologic map (Byers and Barnes, 1967) showing the location of each sampling site. Approximately half of the collected hand samples were deemed to be representative, and were analyzed on a Beckman DK-2A Ratio Recording Spectrophotometer using a magnesium oxide reference target. The spectrometer was operated in %T mode with a scan time

of 10, scale of 1X, sensitivity of 1.30, 0-100% operation range, and a time constant of 0.2. Whole rock samples were positioned in the sample port in a variety of orientations, so that two fresh surface areas and two weathered surfaces were examined. Wavelengths were examined from 0.36 to 2.70 microns, covering all but the thermal band (band 6) of the Thematic Mapper. This wavelength range had to be split into three portions with separate charts prepared for each range, and so there were occasional minor discontinuities in the traces at about 1.20 and 1.86 micrometers. The rocks were grouped into the lithologic units from the published quadrangle map (Byers and Barnes, 1967). Some rock types appeared to have several different spectral characteristic curves, and so were divided accordingly. The spectral response curves of all of the analyzed samples of each major lithologic group determined from the fresh and weathered surfaces were compiled for comparison to histograms from the TM imagery. Several of these are presented with this report (Figures 1,2 and 3).

The Landsat image response of the various lithologic units was obtained using an IIS digital image analysis system at Battelle. The three different images, March, March corrected and June, were used. All bands of data from the June image could be used, but for the March images only bands 1, 4, 5, and 7 were available. The bands were linearly stretched to visually produce similar color responses for the major lithologic units in the three images. Training areas of each major map unit were designated and digitized interactively on the IIS screen. The histograms for each image band in each training area were plotted. These plots were scaled to produce overlays for the rock spectral curves (Figures 1C, 2C, 3C). The wavelength range of each of the TM bands is indicated on the figures, and the histograms are plotted on edge so that the dark end of the histogram (DN = 0) plots at the wavelength axis and the lighter shades plot at the upper end of the histogram. These are then direct overlays to the lithologic spectral data.

RESULTS

The lithologic response curves and image histogram data in general show good agreement. The histogram response curves are much better defined for the March corrected image than for the March original. This is particularly true for darker ridge forming units such as the basalts and several of the carbonates (Figure 1C). In these cases, the uncorrected image's histogram has a wide range in the dark portion of the spectrum. The three different images (March original, corrected, June) had to be stretched in order to get consistent visual color balances for the major lithologic units. As a result, the shadowed portions of the darker ridge formers on the March original image were pushed into saturation. On the corrected image, the histograms are tighter, and so no longer extend into the saturation point, showing that the shadow effects are removed. The June image, with its higher sun angle, tends to show this shadow effect, but to a lesser degree.

Several lithologic units produced distinctly different image responses. On the band 1/4/7 color composite images, both the Rainier Mesa (Tmr) and Ammonia Tanks Members of the Tertiary aged Timber Mountains Tuff appeared to be predominantly orange in color in the southern portion of the quadrangle, and a mottled greenish color in the north. Similarly, undivided tuffs of the Paintbrush Tuff and Indian Trail Formations went from a light to a dark blue in the north. Histograms for these units show distinct patterns, as do spectral curves obtained from rock samples. The Rainier Mesa Member (Tmr) was examined in detail in order to determine the possible causes of this division. The greenish unit was designated unit Tmr-A (Figure 2), and the orange unit Tmr-B (Figure 3). Figures 2 and 3 indicate that Tmr-A is in general darker over the analyzed spectral range, and samples of the two units do show that Tmr-A is, in general, darker than Tmr-B in visible light.

Hildreth (1981), Mills and Rose (1986), Warren and others (1986), and Broxton and others (1986) discuss geochemical variations in ash flow tuffs in general and in Tmr specifically. These studies show a compositional diversity of Tmr from relatively mafic to silicic. Hildreth (1981) hypothesizes that this is due to evolution in a zoned magma chamber, with the roof of the chamber being characterized by a higher silica content. The upper portion of the chamber will be erupted first, followed by the more mafic lower portions. The three Tmr samples from the present study are listed in Table 1 by increasing silica content. Note that the sample that exhibits a Tmr-A spectral response is intermediate in silica with respect to the Tmr-B samples. Hildreth's (1981) average values for the lower (mafic) and upper (silicic) portions of Tmr are listed on Table 1, as are Broxton and others' (1986) observed evolutionary trends of Tmr from mafic to felsic members. When the three Tmr samples are listed by increasing silica, it can be seen that MgO, P2O5, CaO and Zr all follow their trends observed elsewhere (calcium is high). K2O shows an inverted trend. For other elemental oxides, the sample with the intermediate value of silica is either at the high or low end of the trend. Since this is the one Tmr-A sample of the group, the question is whether it is more or less evolved then the two Tmr-B samples. TiO2 and iron indicate that Tmr-A is the more evolved member, whereas barium indicates that it is less evolved. Na2O is ambiguous since Hildreth (1981) and Broxton and others (1986) report different evolutionary trends for sodium.

	Tmr-B #86	Tmr-A #19	Tmr-B #28	Lower	Upper	Trend
SiO2	69.76%	73.25%	76.50%	68.9%	77.4%	increase
MgO	1.12	0.36	0.13	0.6	0.15	
P2O5	0.102	0.049	0.013	0.13	0.01	decrease
CaO	8.87	5.82	1.32	2.2	0.5	decrease
Al2O3	11.52	11.95	12.90	16.3	12.4	decrease
K2O	4.86	4.89	5.11	5.3	4.8	
TiO2	0.131	0.126	0.137	0.37	0.1	decrease
Iron	0.83	0.80	0.86	1.9	0.6	decrease
Na2O	2.68	2.62	2.89	4.1	3.8	increase
Ba	126ppm	131ppm	72ppm	2000	50	decrease
Zr	97ppm	86ppm	86ppm	250	70	

Table 1. Chemical analyses of Rainier Mesa Member (Tmr) samples from the study area compared to averages (Hildreth, 1981) for the lower and upper parts of the magma chamber for Tmr. Tmr samples are plotted with increasing silica content and the Tmr-A sample falls between the two Tmr-B samples. Evolutionary trends are ones observed by Broxton and others (1986) for Tmr and are from mafic to silicic members.

Table 1 shows that the two Tmr-B samples cover essentially the entire range of values (Hildreth, 1981) for both the lower and upper parts of the magma chamber for silica, magnesium, phosphorous, and potassium. According to aluminum, titanium, iron, barium, and zirconium, Tmr-B is probably from the upper magma chamber. Tmr-A is probably also derived from the upper portion of the magma chamber as indicated by aluminum, potassium, titanium, iron, barium, and zirconium. For silica, magnesium, and phosphorous, Tmr-A is between the upper and lower parts of the magma chamber. Calcium is too high for both Tmr-A and Tmr-B, while sodium is too low. These results indicate that both Tmr-A and Tmr-B are more likely to be derived from the upper, more silicic portion of the magma chamber which was erupted first.

There is not an obvious geochemical difference between the two which would produce the observed spectral distinctions.

Broxton and others (1986) state that the more evolved, silicic portion of Tmr has a higher sanidine to plagioclase ratio then the mafic portion. Sanidine is an abundant phenocryst in about equal portions of both Tmr-A and Tmr-B, again indicating that they are both derived from the silicic portions of the magma chamber.

Mills and Rose (1986) state that pumice from Tmr is either mafic or silicic with the base of Tmr characterized by only silica-rich pumice, and the rest by mixed pumice. Of six samples that show Tmr-A spectral signatures, two have only light pumice fragments, while the other four have light and dark pumice. Similarly, three samples of Tmr-B have mixed pumice, two have just silicic pumice fragments.

Many of the Tmr samples show hematitic alteration, with a reddish color in hand sample and obvious hematite in thin section. Similarly, some Tmr spectral curves do show good ferric iron absorption troughs below 0.5 microns. However, a range of hematitic alteration from fresh to altered is obvious in both Tmr-A and Tmr-B. Hematitic alteration does appear to be somewhat more prevalent in Tmr-B.

Strong water absorption is observed for many samples of Tmr-B, as seen in Figure 3. However, several Tmr-A samples also show water absorption. Samples exhibiting hematitic alteration have better developed water absorption bands.

Smith (1960) discussed in situ evolution of an ash flow tuff. As the tuff cools in place, it may weld, with a zone of dense welding occurring in the lower-central portion of thick flows, zones of partial welding above and below this, and zones of poor welding at the top and base of the flow. Superimposed upon this will be varying crystallization effects, including devitrification, vapor phases, granophyric crystallization, and fumarolic alteration. Devitrification of glass into spherulitic and axiolitic intergrowths was ranked in five thin sections of Tmr (Table 2). Samples of Tmr-A ranked in the middle (3 and 4), while Tmr-B thin sections showed the most (5) and least (1 and 2) devitrification. Degree of devitrification did not directly correlate to strength of water absorption bands.

	Tmr-A	Tmr-B
Devitrification (1=low, 5=high)	3,4	1,2,5
K2O/Na2O (<1.8=not hydrated)	1.81, 1.87	1.77
Welding (1=mod,3=mod well)	1,2	1.5,2,3
Density (1=light,4=heavy)	1,1,2,3,3,4	2,2,3,4,4

Table 2. Ranking of devitrification, hydration of glass, welding, and density of samples.

Vapor phase crystallization involves growth of alkali feldspar, tridymite, and cristobalite in open pore spaces (Smith, 1960). It predominantly occurs in the upper partially welded zone and part of the upper poorly welded zone. Sanidine crystals are equally common in Tmr-A and Tmr-B. The presence of tridymite and cristobalite has not been assessed. Granophyric recrystallization is relatively rare, and only occurs in the interior core of very thick ash flows. Fumaroles on the surface of the flow may lead to local alteration, characterized by distinctly variegated colors which have not been seen in either Tmr-A or Tmr-B samples.

The degree of welding was assessed for five thin sections of Tmr. Welding ranged from moderately welded (1) to moderately well welded (3). Tmr-B was in general better welded, as shown on Table 2. Relative densities of the samples showed a similar relationship, with samples of Tmr-B being more dense than Tmr-A.

The U.S. Geological Survey quadrangle map for the study area (Byers and Barnes, 1967) describes Tmr as being a "multiple-flow compound cooling unit" up to 60m thick with a "brown, densely welded vitrophyric biotite ash-flow tuff" on the top and a "pinkish-grey pumiceous partly welded ash-flow tuff" below. Comparison of the geologic map and the TM imagery shows that where the two occur together, unit Tmr-A outcrops on hill tops and Tmr-B on lower slopes. This matches the observation that samples of Tmr-A are more brownish, whereas Tmr-B are more pinkish. However, it is contrary to the observation that Tmr-B is better welded, and both range from only moderate to moderately well welded. Overlaying the sampling locations onto the image and geologic map shows that all of the Tmr-A samples were collected from portions of the image showing the greenish Tmr-A color response, while all but one of the Tmr-B samples were from orange (Tmr-B) regions.

Spectral response curves for several other rock sample groups were divisible into sub-units. This was true of the altered tuff and the Ammonia Tanks Member (Tma) of the Timber Mountain Tuff. Not all of these groups appear to match the histograms from their training areas. All of the Tma samples were collected from the northwest part of the quadrangle sandwiched between a mapped altered tuff and basalt dike. The histograms for this unit did not give a good match to any of these sub-units, but its training area was in the southern portion of the quadrangle, where the unit was more pristine. Tma and the altered tuff were also analyzed using thin sections and geochemistry, and yielded results similar to Tmr.

CONCLUSIONS

This study shows that reflectance spectral response curves are valuable for distinguishing different units and sub-units on images. It also shows the necessity of collecting rock samples from the training areas directly, rather than in areas that may be affected by different alteration and weathering processes.

ACKNOWLEDGEMENTS

We are grateful for the aid and assistance given by the operators of the spectrometer (G. Mohl, C. Brougher), staff at Battelle PNL for the loan of the spectrometer and the usage of their remote sensing systems, E. Rieken for drafting, H. P. Foote for the preparation of the shadow corrected images, and Dr. P. Larson, K. Bailey and R. Conrey for analysis of the thin sections and rock samples. We are also especially grateful to P. Larson for helpful discussions on the zonation of silicic magma chambers and nature of ash flow tuffs.

REFERENCES

Broxton, D.E., Warren, R.G., and Byers, F.M. Jr., "Petrochemical Trends in the Timber Mountain-Oasis Valley Caldera Complex, SW Nevada", Eos, Vol. 67, p 1260, 1986.

Byers, F.M. Jr. and Barnes, H., "Geologic Map of the Paiute Ridge Quadrangle, Nye and Lincoln Counties, Nevada Test Site", U.S. Geological Survey Geologic Quadrangle Map GQ-577, 1:24,000 scale, 1967.

Eliason, J.R., "A Technique for Structural Geologic Analysis of Topography", Ph.D. thesis, Washington State University, Pullman, WA, 166 p, 1984.

Eliason, J.R. and Thiessen, R.L., "Geologic Spatial Analysis, a New Multiple Data Source Exploration Tool", International Symposium on Remote Sensing of the Environment, Fifth Thematic Conference, Remote Sensing for Exploration Geology, pp 763-774, 1987.

Foote, H., Thiessen, R.L., Eliason, J., and Wukelic, G., "Advanced Techniques for Lithologic and Fracture Analysis Utilizing Thematic Mapper and Topographic Data", Presented at Final Landsat Image Data Quality Program (LIDQA) Symposium, NASA, 1985.

Hildreth, W., "Gradients in Silicic Magma Chambers: Implications for Lithospheric Magmatism", Journal of Geophysical Research, Vol. 86, pp 10153-10192, 1981.

Mills, J.G. Jr. and Rose, T.P., "Geochemistry of Glassy Pumices from the Timber Mountain Tuff, Southwestern Nevada", Eos, Vol. 67, p 1262, 1986.

Smith, R.L., "Zones and Zonal Variations in Welded Ash Flows", U.S. Geological Survey Professional Paper 354-F, pp 149-159, 1960.

Thiessen, R.L., Eliason, J.R., and Rieken, E.R., "Three-Dimensional Computer Analysis and Modelling of Remote Sensing-Structural Geologic Problems", Twelfth Canadian Symposium on Remote Sensing (this volume), 1989.

Thiessen, R.L., Johnson, L.K., Foote, H.P., and Eliason, J.R., "Surface Reflectance Correction and Stereo Enhancement of Landsat Thematic Mapper Imagery for Structural Geologic Exploration", International Symposium on Remote Sensing of the Environment, Fifth Thematic Conference, Remote Sensing for Exploration Geology, pp 763-774, 1987.

Warren, R.G., Nealey, L.D., Byers, F.M. Jr., and Freeman, S.H., "Magmatic Components of the Rainier Mesa Member of the Timber Mountain Tuff, Timber Mountain-Oasis Valley Caldera Complex", Eos, Vol. 67, p 1260, 1986.

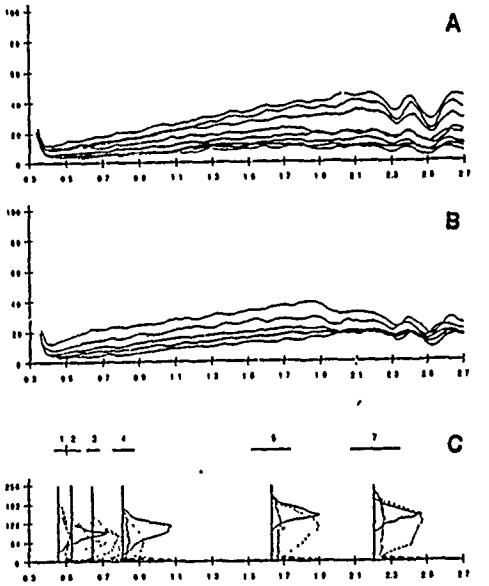


Figure 1. Spectral curves and histograms for the Smoky Member of the Nopah Formation, a carbonate. A) Spectral curves for fresh surfaces. B) Spectral curves for weathered surfaces. C) Histograms from the original March TM image (dashed line), March shadow corrected TM image (solid line) and June TM image. Notice the improvement obtained with the shadow correction technique.

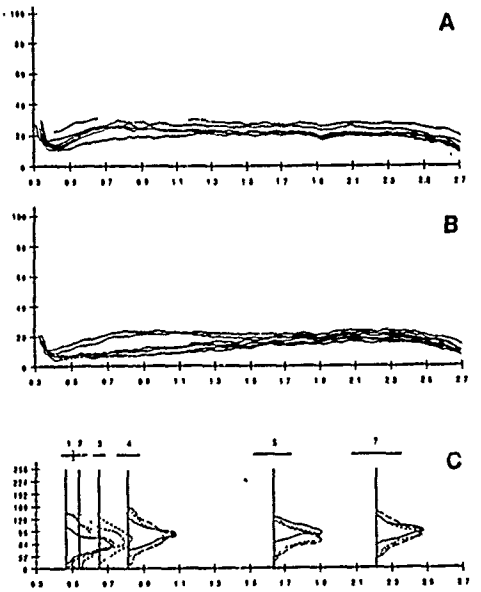


Figure 2. Spectral curves and histograms for Tmr-A, one of the observed responses from the Rainier Mesa Member of the Timber Mountain Tuff. Same conventions as Figure 1.

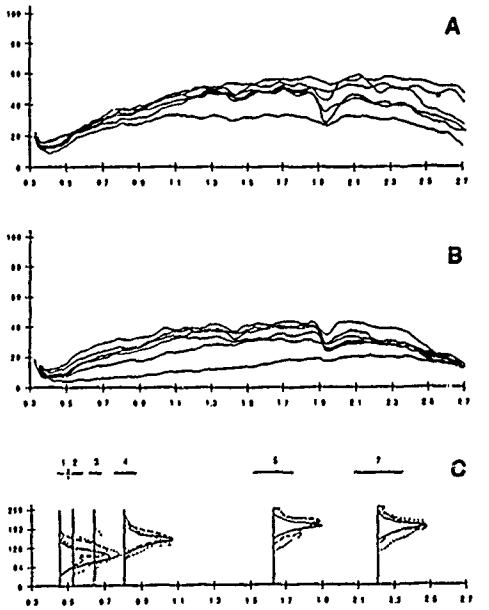


Figure 3. Spectral curves and histograms for the other response (Tmr-B) from the Rainier Mesa Member. Same conventions as Figure 1

SPECTRAL EMISSIVITY OBSERVATIONS IN HAPEX

T.J. Schmugge

USDA-ARS Hydrology Laboratory
Beltsville, MD 20705 USA

F. Becker

Universite Louis Pasteur
67000 Strasbourg, France

ABSTRACT

The Thermal Infrared multispectral Scanner (TIMS) was flown on the NASA C-130 aircraft for series of a 12 flights during the HAPEX experiment in the Southwest of France during 1986. TIMS provided coverage of the 8-12 micrometer thermal infrared band in 6 contiguous channels so it is possible to observe the spectral behavior of the surface emissivity over this wavelength interval. The surfaces observed ranged from bare soils to coniferous forest. As expected, the fully vegetated fields and coniferous forest exhibited little or no spectral variation as did a small lake. However, the bare soil surfaces had a 5 to 10°C difference in radiant temperature over the TIMS channels with the 3 shorter wavelength channels ($8.0 < \lambda < 9.6$ micrometer) being cooler than the 3 longest wavelength channels. This qualitatively agrees with laboratory measurements done by a group from the Universite Louis Pasteur in Strasbourg on soil samples from this area. This spectral variation arises from the absorption of infrared radiation due to the stretching vibrations of the silicon-oxygen bonds of silicates in the soil. These results indicate that the longer wavelength channels in the 8-12 micrometer band may be more effective for surface temperature sensing for soils rich in silicates.

1. INTRODUCTION

Understanding the spectral variation of the surface emissivity in the thermal infrared portion, 8-14 micrometers or microns, of the electromagnetic spectrum for the earth's surface is important for the determination of its surface temperature. Over the ocean the emissivity is assumed constant over this region, however for land there is considerable evidence of the spectral variation of the surface emission over the thermal infrared band. Large spectral features lie in the 8-10 micron region and are attributed to the fundamental vibrations of Si-O bonds [1]. They are most pronounced in quartz. An example of this feature is shown in Fig. 1 where laboratory measurements of the emissivity of a sand and a loess soil from France are presented [2]. For the sand the emissivity varies from less that 0.8 between 8 and 9 microns to about 0.95 at 10-12 microns. There is a similar but less pronounced variation for the soil. The substitution of other metals, e.g., Al or Mg for Si produces changes in the width, depth and center wavelength of the absorption band and thus affords possible approaches for mapping silicate rocks [3]. However, for the purpose of using thermal infrared radiation for estimating surface temperature, these variations in surface emissivity provide additional complications [4].

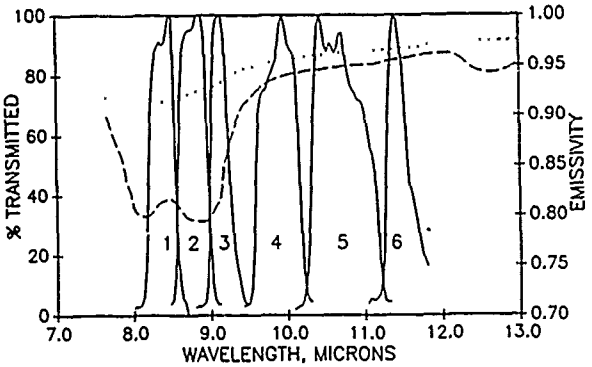


Fig. 1. Laboratory measurements of the emissivity for a sand (dashed curve) and a loess soil (dotted curve) from Nerry et al. [2]. The filter functions for the 6 TIMS channels are also plotted.

2. TIMS

The results we will present here were obtained with the TIMS sensor on board the NASA C-130 during the HAPEX aircraft campaign in 1986. HAPEX (Hydrologic Atmospheric Pilot Experiment) had as its primary objective the improved parameterization of the surface fluxes in atmospheric circulation models [5], [6]. It included an intensive program of ground and aircraft observations. The remote sensing of surface was performed by a suite of sensors on the C-130 aircraft. These included a microwave radiometer for soil moisture observations, visible and near IR scanner for vegetation and albedo data and thermal infrared scanner for surface temperature.

The Thermal Infrared Multispectral Scanner (TIMS) has six channels in the thermal infrared (8-12 microns) region of the electromagnetic spectrum. The filter functions of the of the channels are shown in Fig. 1. In the figure it is clear that the lower 3 channels will have a lower response.

The instantaneous field of view is 2.5 mrad, while the detector signals are sampled every 2.08 mrad along the scan [7]. Thus, from an altitude of 1500 m the pixel size is 3.75 m and the cross track separation is 3.2 m at nadir. For the mid altitude (~1500 m) flights to be considered here the scan rate was set at 25/sec, which at a typical aircraft speed of 80 m/sec provided complete coverage. The digitized field of view contains 638 pixels and covers 76.6° yielding a swath width of 2.4 km.

For calibration the system is equipped with cold and warm reference sources or blackbodies, approximately covering the temperature range of interest. For most of HAPEX the temperature separation between the two references was set at 30°C. The reference temperatures are known to better than 0.5 K [7].

TIMS responds to the incident radiance (W/m^2 sr cm^{-1}) and is related to the brightness temperature of the observed surface via the Planck equation for blackbody radiation. However, if the observed temperatures are close to the calibration temperatures, say within 10 or 15°C, the error arising from using a linear temperature calibration is less than 1°C and that approximation will be used here [8].

The temperatures given by the sensor represent the detected radiance at the aircraft altitude. In order to convert this result to the actual surface temperature, atmospheric effects must be taken into account. These include the absorption and emission by the atmospheric gases, primarily water vapor for this portion of the spectrum. Since the relationship between radiance and temperature is considered linear, Eq. 1 can be used to correct for these atmospheric effects.

$$T_{ir} = \tau * T_s + (1 - \tau) * T_{av} \qquad (1)$$

T_{ir} is the observed temperature (K), T_s the surface temperature, T_{av} the average air temperature for the atmospheric layer between the surface and the aircraft and τ the transmittance of this layer for a specific channel. Radiosoundings released at the central site within one hour of the pass are used to determine the transmittance and the average air temperature. The latter is calculated assuming that the principal absorption is by water vapor and so the temperatures are weighted by the vapor content of the layer. LOWTRAN-6, an atmospheric path radiance model developed by the Air Force Geophysics Laboratory [9], is used to calculate the transmittance for the different channels. On June 16 the values of τ ranged from 0.70 for band 1 to 0.82 for band 5. These results are for the 1.5 km altitude with T_{av} = 292.0 K. The low value for band 1 indicates the increased water vapor absorption at the 8 micron edge of the window as seen in Fig. 1.

3. RESULTS

The central site for the surface measurements program in HAPEX was a large clearing (5 x 5 km) in the Les Landes pine forest in the northern portion of the square near the town of Lubbon. Within this clearing acoustic and radio soundings of the atmosphere were made along a road separating a corn and oat field. There were SAMER flux measurement stations [10] in each of these fields. Images of the TIMS data for this general area were created. From these images it was possible to select the data which corresponded to specific surface conditions.

Averages of the radiant temperature from several hundred to a thousand pixels were obtained and the results are presented in Fig. 2 for 16 June 1986. The surfaces considered were oats, corn, pine forest and bare soil. The oat field was quite mature with full cover while the corn was about 0.7 m high with considerable bare soil showing.

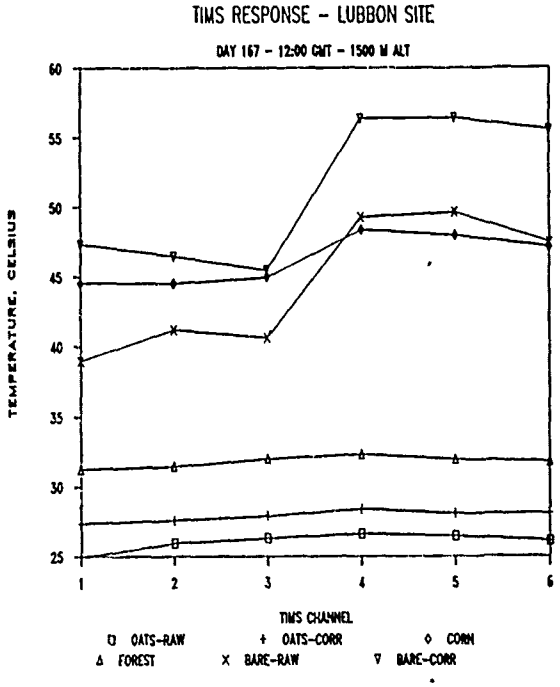


Fig. 2. Spectral variation of the TIMS response for several surface conditions. The data for the corn field and forest are corrected for atmospheric effects, both the raw and corrected values are given for the oat and bare fields.

For the oat and bare fields, both the raw and atmospheric corrected temperatures are presented to demonstrate the magnitude of the atmospheric effects. For the oat field, the correction is between 1.5 and 2.5°C while for the bare field it is much larger, between 5 and 8°C. After correction, the response is rather flat for both oat field and the forest, with about a 1°C spread among the 6 channels. There is a small (~3°C) for the corn field and large (~10°C) shift for the bare soil. The latter implies about a 0.03 shift in emissivity between the lower and upper channels of TIMS. This is comparable to that shown in Fig. 1 for the loess soil.

The results for both 16 and 27 June are summarized in Table 1. The biggest difference between the two days is for the corn field whose surface temperature decreased by 15°C as a result of increased vegetation cover, the height of the plants increased from .7 to 1.7 m during this interval. On 16 June there is a 3 to 4°C difference between the lower and upper TIMS channels and on the 27th the difference has decreased to less than 1°C. While the corn was cooling the oat field was getting warmer as the crop passed into senescense. The air temperatures listed in Table 1 were from the flux measurement stations in each field and it is seen that wasn't much difference between the fields on either day.

Table 1. Average TIMS brightness temperatures in Celsius for Lubbon site

	Oats		Corn		Forest		Bare field	
	Day 167	Day 178	Day 167	Day 178	Day 167	Day 178	Day 167	Day 178
Air temp	27.7	26.2	27.5	25.1	---	---	---	---
Chan 1	27.4	31.3	44.6	30.8	31.3	28.4	47.4	49.8
Chan 2	27.6	31.1	44.5	30.5	31.5	28.2	46.5	48.6
Chan 3	27.9	31.4	45.0	30.5	32.0	28.3	45.5	47.6
Chan 4	28.4	31.6	48.4	31.8	32.4	28.2	56.4	57.3
Chan 5	28.1	31.5	48.0	31.9	31.9	28.4	56.4	56.8
Chan 6	28.1	31.6	47.2	31.9	31.8	28.5	55.5	56.5

4. DISCUSSION

In a recent paper, Price [4] developed an algorithm for estimating land surface temperatures using the split window channels on the NOAA AVHRR sensor. The algorithm involved a correction for the surface emissivity differences between channels 4 and 5 on the AVHRR. The wavelengths for these two channels are approximately the same as channels 5 and 6 of TIMS. The results presented here would indicate that for the HAPEX area the emissivity difference between the channels is essentially zero which reduces the complexity of the surface temperature extraction. However, the absolute value of the emissivity must still be known and we were unable to determine that with these data.

As we noted for the corn field the spectral difference decreased as the crop matured because of the reduced amount of bare soil visible to the sensor. This decrease may then be used as an indicator of the amount of vegetation cover. Gillespie [3] observed this phenomena in flights over a quartz rich rock with varying covers of a pine forest in Sierra Nevada of California. He was able to observe the rock spectral features with up to 80% cover. Thus, the spectral variation of surface emissivity may provide another tool for estimating vegetation cover amounts.

5. REFERENCES

1. Hunt, G.R., (1980). Electromagnetic radiation: the communication link in remote sensing, In Siegal, B.S. and A.R. Gillespie, eds., Remote Sensing in Geology, Wiley, New York, pp. 5-45.

2. Nerry, F., J. Labed, and M. Ph. Stoll, (1988). Emissivity signatures in the thermal infrared band for remote sensing applications: Calibration procedure and method of measurement. Applied Optics.

3. Gillespie, A.R., (1986). Lithologic mapping of silicate rocks using TIMS, in Kahle, A. B. & Abbott, E. (1985) TIMS data user's workshop, June 18-19, 1985, NASA, JPL Publ. 86-38, Pasadena, Nov. 1986.

4. Price, J.C., (1984). Land surface measurements from the split window channels of the NOAA-7 advanced very resolution radiometer, J. Geophys. Res., 89, 7231-7237.

5. Andre, J.C., Goutorbe, J.P. & Perrier, A. (1986) HAPEX-MOBILHY: A hydrologic atmospheric experiment for the study of water budget and evaporation flux at the climatic scale. Bull. Amer. Met. Soc., 67, 138-144.

6. Andre, J.C. & 30 others (1988) Evaporation over land surfaces: First results from HAPEX-MOBILHY special observing period. Annales Geophysicae, 6, 477-492.

7. Palluconi, F., & Meeks, G.R. (1985) Thermal infrared multispectral scanner (TIMS): An investigator's guide to TIMS data; NASA, JPL Publ. 85-32, Pasadena, June 1985.

8. Schmugge, T. & Janssen, L. (1988) Aircraft remote sensing in HAPEX. Proc. 4th Int'l Coll. on Spectral Signatures in Remote Sensing, Aussois, France 18-22 January 1988 (ESA SP-287), 463-467.

9. Kniezys, F.X., Shettle, E.P., Gallery, W.O., Chetwynd, J.H., Abreu, L.W., Selby, J.E.A., Clough S.A. & Fenn, R.W. (1983) Atmospheric Transmittance/Radiance Computer Code LOWTRAN-6, Air Force Geophysics Laboratory Report No. AFGL-TR-83-0187, Hanscom AFB, Massachusetts 01731.

10. Riou, C., (1982) Une expression analytique du flux de chaleur sensible en conditions sur adiabatiques a partir de mesures du vent et de la temperature a deux niveaux. J. Rech. Atmos., 16, 15-22.

SPECTRAL REFLECTANCE CHARACTERISITICS OF COTTON & RICE IN EGYPT

Hashem M. El-Khattib
2 El-Nokali St., El-Nadi St.,
El-Haram, Giza, Egypt.
Tel. .: 720608
Senior Researcher

Fetoh M. Hawela
1 Ahmed El-Wakil St.,
Dar El-Salam, Cairo, Egypt
Tel. : 980137
Senior Researcher

Soil and Water Research Institute, Gamaa St., Giza, Egypt.

The use of portable ground-based sensors for measuring crop reflectance has created a need for comparable and reliable measurements procedures capable of providing caliberated and reproducible canopy reflectance data. It could be said that the measurement of spectral characteristics of the main field crops, using hand-held radiometer, is a suitable means for their identification on satellite images. Also, it will help in understanding their spectral seperability which, in turn, could assist computer-aided classification of crops on satellite digital data. Spectral signature is one of the useful tools in the analysis of LANDSAT digital data as it forms the basis of identification and discrimination between variuos field crops and features on earth. The objective of the current study aims at correlating ground bidirectional spectral reflectance data of cotton and rice field crops throughout their growing season with high altitude remotely sensed data. Using hand-held spectroradiometer having four wavelength bands of the LANDSAT multispectral scanner (MSS), enables us to collect in situ temporal spectral measurements of cotton and rice field crops.

Green, red and infrared reflectances are closely related to the agronomic and physiological properties of both crops such as plant height, crop cover percentage, green leaf biomass and chlorophyll content. Data also reveal that the decreased reflectance values in infrared portion particularly for rice during maturity stage aremainly due to yellowing and dryness of rice plants. The reduction in the infrared reflection for rice canopies is, however, 4 times as that of cotton in band 6; the reduction is almost one third for submerged rice than that of cotton in band 7.

The green/red ratio seems to be ineffective for vegetation discrimination. Significant interrelationships have been established between infrared/red ratio, vegetation index, plant height and crop cover percentage.

Therefore, monitoring the stages of crop development based on vegetation index for cotton and rice have been proposed as well as five distinct and spectrally measurable stages which were defined. The reflectance values reveal that maximum growth for cotton which has a long growing season, about 7 months, where two thirds of this period is represented by vegetative and flowering stages. The highest value of spectral reflectance had been recorded for cotton using the four bands of the radiometer on 15.)88, while rice at the middle of the growing season, 2 months after plantation. However, spectral reflectance values for rice and corp are lower than that of cotton at their maximum growth.

OPERATIONAL TOVS SOUNDINGS USING A PHYSICAL RETRIEVAL APPROACH

Anthony L. Reale
National Oceanic And Atmospheric Administration
National Environmental Satellite, Data, and Information Service

Mitchell D. Goldberg and Jaime M. Daniels
ST Systems Corporation
Lanham, Maryland 20706

ABSTRACT

NOAA/NESDIS produces operational sounding products from the TIROS Operational Vertical Sounder (TOVS) instrument onboard two NOAA polar orbiting satellites. NESDIS provides timely processing and distribution of TOVS sounding products for use in operational numerical weather prediction forecast models.

On September 16, 1988 the TOVS operation was upgraded from a statistical regression to a physical retrieval approach to compute atmospheric sounding products (Fleming, 1988). The physical retrieval approach applies radiative transfer physics to retrieve atmospheric sounding products from the observed TOVS measurements. This paper discusses the assimilation of these new algorithms into the TOVS operational software system and presents results.

Key Words: Sounding, Satellite, Physical, Radiance, Weather

1. TOVS BRIGHTNESS TEMPERATURE COMPUTATION

The TOVS instrument consists of three sounders and 27 sounding channels. The High-resolution InfraRed Sounder(HIRS/2) contains 20 channels in the 15 micron and 4.5 micron ranges with a horizontal field-of-view (fov) of 17.5 km at the sub-satellite point. The Microwave Sounding Unit (MSU) contains 4 channels ranging from 50.3 to 57.9 gHz with a 110 km fov at the sub-satellite point. The Stratospheric Sounding Unit (SSU), provided by the British Meteorological Office, contains 3 channels which measure stratospheric emissions with a 110 km fov at nadir.

The TOVS brightness temperatures for each channel that are used in the retrieval sequence are computed in three principal software modules prior to entering the retrieval sequence. The TOVS raw data first undergo initial processing in the TOVS Preprocessor module. The preprocessing consists of steps to:

* screen and calibrate the incoming TOVS data,
* limb correct the data to the nadir scan view,
* correct the HIRS/2 channels 8, 18, and 19 for water attenuation effects,
* screen the MSU data for precipitation contamination,
* interpolate the MSU data to each HIRS/2 fov, and
* identify mislocated TOVS data.

The preprocessed TOVS data at each HIRS/2 fov are then entered into the TIROS Atmospheric Radiance Module (TARM). The HIRS/2 data are tested for clouds and a cloud clearing algorithm (McMillin, 1982) is applied to remove cloud effects. At the conclusion of TARM processing, a clear-column TOVS brightness temperature profile has been computed for each TOVS "minibox". A minibox is a pre-defined 3 by 3 array of HIRS/2 fov locations (Reale, 1986). Each minibox is labelled as either clear, corrected clear or cloudy; for cloudy miniboxes the HIRS/2 tropospheric channels are suppressed.

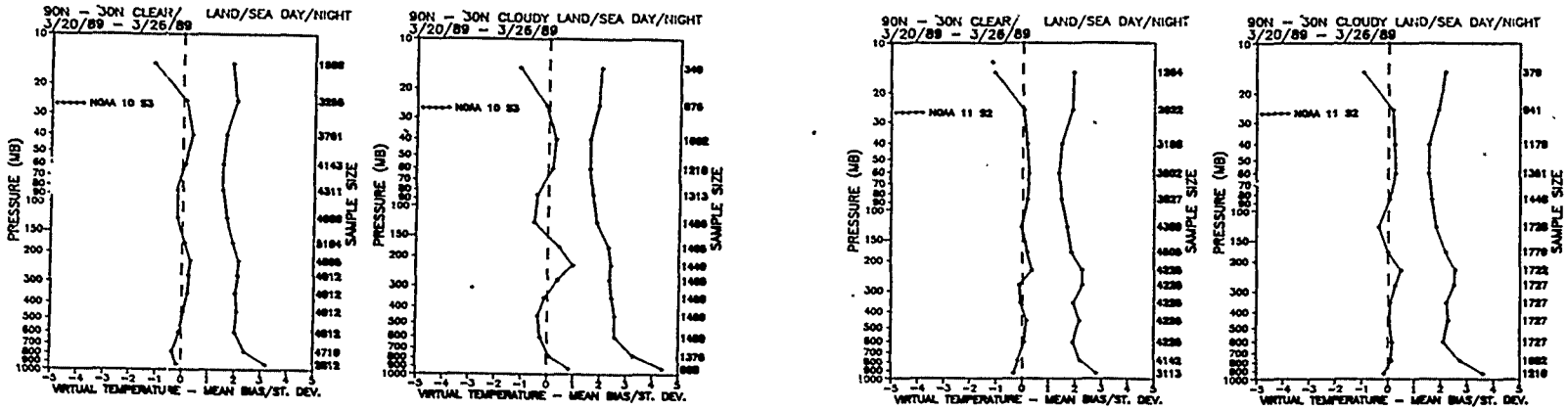
The clear-column profiles for each minibox then enter the TOVS Filter module (Reale, 1986). Objective analysis is applied to the clear and corrected clear profiles and outlier miniboxes are removed. The remaining TOVS miniboxes are then reduced to a "low density" dataset which has a nominal 250 km spacing between miniboxes with a minimum of 100 km spacing in regions of increased brightness temperature gradient.

2. TUNING

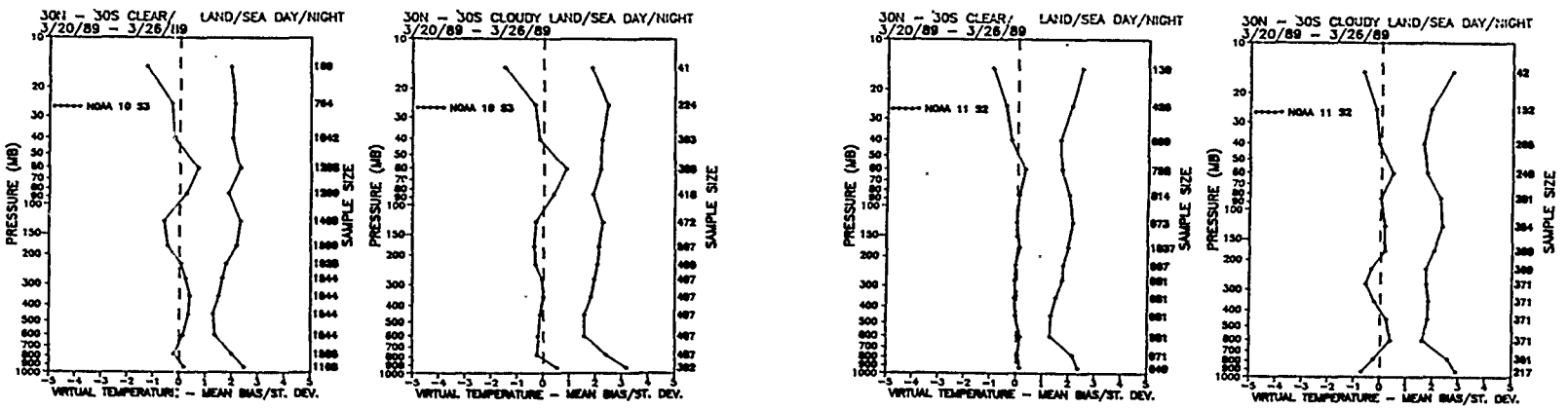
The physical retrieval approach requires knowledge of the atmospheric transmittance functions. Since an accurate knowledge of these functions is not known, the required relationship between temperature and radiance is estimated from "small sample" sets of collocated radiosonde and satellite soundings (Fleming and Crosby, 1986). These are referred to as the tuning data. The procedure to generate tuned datasets for use in the retrieval sequence entails:

* the daily collection of collocated radiosonde and satellite sounding matched pairs,
* the daily updating of a 28-day and 21-day rotating datasets of matched pairs,
* the weekly updating of mean temperature, moisture and TOVS brightness temperature profiles for 27 geographical zones, and
* the weekly updating of the retrieval "operator" constructed for each geographical zone.

Radiosonde and satellite sounding matched pairs are collected daily for each satellite. These matched pairs must be collocated within three hours and at least 300 km. The distance window varies by region based on the availability of timely radiosondes and is less than 50 km in the northern mid-latitudes.



Figures 1 - 4: Vertical accuracy statistics for NOAA-10 and NOAA-11, clear and cloudy soundings, for the latitude zone 90N to 30N.



Figures 5 - 8: Vertical accuracy statistics for NOAA-10 and NOAA-11, clear and cloudy soundings, for the latitude zone 30N to 30S.

Additional checks are also done to insure that the each radiosonde report is complete. A radiosonde report must extend from at least 850 mb to 100 mb, have no large gaps and contain both temperature and moisture data. Because of the time window imposed and the tendency for most radiosondes to report at 0Z and 12Z, the matched pairs collected for each satellite tend to be regional.

The matched pairs collected each day are stored on a 28-day rotating dataset with a separate dataset for each satellite. The latest 21 days of matched pairs provide a principle database from which first guess information is extracted during the retrieval sequence. This is referred to as the 21-day first guess library and is updated daily.

The 28-day rotating database is also used to compute the "small sample" mean temperature, moisture and TOVS brightness temperature profiles. Small sample mean profiles are computed for each of 27 geographical zones and are updated weekly. Each zone represents one of nine latitude regions for sea, land/day or land/night (Goldberg, 1988). The small sample means are based on the most recent matched pairs with a maximum sample size of 100. If a sample size of at least 20 pairs is not reached for a given zone then pairs from neighboring zones are used.

The small sample mean profiles for each zone are used to select a pre-computed covariance matrix, defined as the S- matrix, and to compute linearized weighting functions for each of the TOVS channels, defined as the A-matrix. The S and A matrices are then used to construct a linearized retrieval operator (Fleming, 1986) for each of the of the 27 geographical zones. This is done each week. In all tuning and subsequent retrieval sequence computations the temperature and brightness temperatures have been linearized to 700 cm^{-1}.

There are 108 pre-computed S-matrices from which to choose based on a historical dataset of radiosonde observations sorted over 12 months and 9 latitude belts (Crosby, 1973). Each S-matrix has a mean temperature profile based on the sample of radiosondes used to compute it, these are referred to as the "large sample" means. The S-matrix chosen for a given geographical zone is that for which large sample and small sample mean temperature profiles are closest. The closeness measure is the ordinary Euclidean norm applied between the surface and 50 mb. (Dey, 1989).

The A-matrix is computed for each zone based on the small sample mean temperature and moisture profiles (Fleming and Crosby, 1986). This requires extending the small sample mean temperature profiles to .01 mb which is done using the large sample mean. Simulated TOVS radiances based on the small sample and large sample mean temperature and moisture profiles are also computed and stored for later use. Moisture profiles were not originally available and were estimated and appended to the large sample dataset.

The 28-day matched pair dataset is also used for routine monitoring. Transmittance adjustment or gamma factors for each TOVS channel are monitored by comparing the observed and simulated TOVS radiances for each matched pair. Gamma factors are global and are updated as needed. The accuracy of sounding products is also monitored, this is discussed in Section 4.

3. THE RETRIEVAL SEQUENCE

The physical retrieval algorithm used is the Minimum Variance Simultaneous (MVS) method (Fleming, 1986). The retrieval method is simultaneous in that profiles of temperature and moisture, as well as surface temperature, are retrieved as a single solution vector. The method is physical since it inverts the radiative transfer equation to obtain a temperature and moisture profile from an observed set of radiances. The retrieval sequence can be divided into three parts:

* computation of the first guess,
* selection of the retrieval operator, and
* solving the retrieval.

The first guess temperature, moisture and radiance profiles between the surface and 50 mb. are obtained using a pattern recognition library search technique (Goldberg, 1988). This technique uses discriminate analysis to compute a "closeness measure" between the clear column radiance profile for a given minibox and those contained in the 21-day first guess library. The closeness measure (CM) is shown in equation 1:

$$CM = (R_i - R_{ik})\, B^{-1}\, (R_i - R_{ik})^T \qquad (1)$$

for i=1,... I TOVS channels, and
k=1,... N library radiance profiles.

In equation 1, B denotes a covariance matrix of the clear column radiance profiles in the first guess library, $^{-1}$ denotes the inverse and T the matrix transpose. B is computed using either land or sea profiles depending on the given minibox designation. The TOVS channels (I) used to select the 20 closest pairs are dependent on the sounding type. For clear and corrected clear miniboxes the HIRS/2 channels 2-7, 13-16 and the MSU channels 2-4 are used. For cloudy miniboxes the HIRS/2 channels 2 and 3, the MSU channels 2-4 and the sea surface temperature are used. The twenty closest radiance profiles, along with their matched radiosonce temperature and moisture profiles are then averaged to compute the first guess up to 50 mb. Sample sizes may decrease near 50 mb. due to missing radiosonde data.

The first guess temperature profile is then used to select the MVS retrieval operator. This is done by first selecting the geographic zone for which the small sample and first guess temperature profiles are closest. Only geographic zones in the vicinity of a given minibox are considered. The closeness measure is the ordinary Euclidean norm referenced earlier. Closeness is measured from 1000 to 50 mb. for ocean miniboxes and from 700 to 50 mb. over land. Once the closest zone is selected, the retrieval operator constructed for that zone during "tuning" is used.

Finally, the first guess information above 50 mb. is estimated based on the selected geographic zone. This is done in two steps. Between 50 mb and 30 mb the small sample mean profiles are used. Above 30 mb. where no radiosonde data is available, the large sample mean temperature profile and the TOVS radiances simulated from this profile are used.

The MVS physical retrieval solution (Fleming, 1988) is then applied to compute the sounding products for the minibox. A generalized matrix form of the solution is shown in equation 2:

$$V = VGESS + C(R - RGESS) \qquad (2)$$

In equation 2, V is the sounding products vector, VGESS is the first guess products vector, C is the retrieval operator, R is the clear column radiance vector for the monibox and RGESS is the first guess radiance vector. The TOVS channels used are the same as those used to compute the first guess. The retrieved temperature and moisture profiles are then converted to layer mean virtual temperature and precipitable water contents. These products for the low density dataset are routinely distributed and archived by NESDIS.

4) RESULTS

NESDIS is responsible for the routine monitoring of satellite sounding accuracy. The traditional method has been to compute vertical accuracy statistics based on the matched pair datasets described in Section 2 using the radiosonde as ground truth. Figures 1 to 8 show typical vertical error statistics in the layer mean virtual temperature for NOAA-10 and NOAA-11 operational physical retrievals based on a one week sample during March. Error statistics are shown for clear and cloudy soundings and for two latitude regions, 90N to 30N (top row) and 30N to 30S (bottom row). Typical satellite sounding errors in the troposphere are less than 2K with a bias of less than .5K. The largest errors occur near the surface particularly for the cloudy soundings. Careful interpretating of such statistics is needed since they do not include vast, remote regions of the globe for which radiosondes are not available. Furthermore, the radiosondes are subject to measurement errors. However, experience suggests that vertical error statistics computed in such a manner provide a meaningful estimate of typical satellite sounding errors.

Improved cooperation between NESDIS, the National Meteorological Center (NMC), and international forecast centers such as the European Center for Medium Range Weather Forecasting (ECMWF) has resulted in the use of a new tool for validating satellite sounding products, namely the analyzed fields from weather forecast cycles. Examples of this are illustrated in Figures 9 to 12. Figures 9 and 11

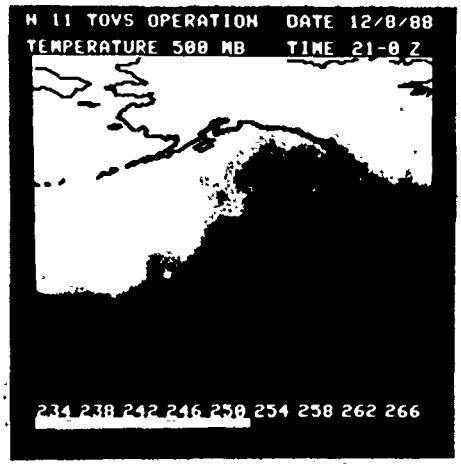


Figure 9: Graphic display of analyzed 500mb temperature field for successive NOAA-11 orbits.

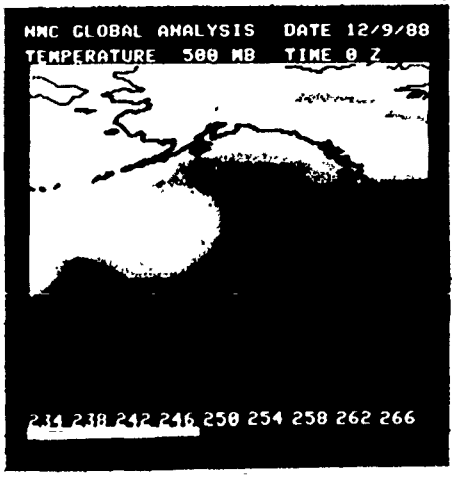


Figure 10: Graphic display of 500mb temperature field for NMC Global Analysis.

Figure 11: Graphic display of analyzed 500mb temperature field for successive NOAA-10 orbits.

Figure 12: Graphic display of 500mb temperature field for NMC 6-hour Forecast.

show graphic displays of analyzed 500 mb. temperature fields of satellite soundings for successive NOAA-10 and NOAA-11 passes over the N. Pacific. Figures 10 and 12 show the corresponding NMC Analysis and 6-hour Forecast analyzed products. Each satellite data field represents a composite analysis over about four hours (i.e., 3 orbits) while the NMC fields are synoptic. Overall, areas of agreement and some differences can be seen. Forecast modelers and satellite meteorologist alike are coming to rely more heavily on such comparisons since they more directly measure model impact. The approval of the TOVS physical retrieval system for operational implementation required that NMC perform a successful model impact study (Dey, 1989).

5. REMARKS

The NESDIS mission provides for the development and implementation of improved soundings generation algorithms. The implementation of a physical retrieval algorithm for TOVS sounding products is one example. Current research to improve the first guess through air mass classification techniques and the use of the NMC 6-hour forecast is ongoing. Improvements in the physical models of atmospheric radiation transfer are also needed and are under investigation. The spirit of cooperation between NESDIS, NMC and international forecast centers who rely on sounding products has greatly improved in recent years. Increased user feedback has become important in identifying problems and determining the priority of research support activities. A better understanding of the impact of satellite soundings on numerical weather forecast models and improved applications of satellite based products in the models are evolving.

6. BIBLIOGRAPHY

1. Crosby, D.S., et. al.: Covariance Matrices and Means of Atmospheric Planck Function Profiles for Application to Temperature Soundings from Satellite Measurements. Journal of Atmospheric Sciences, Vol.30, No. 1, January, 1973

2. Dey, C.H., et. al.: An Examination of NESDIS TOVS Physical Retrievals Using Data Impact Studies. Unpublished Manuscript, NOAA/NMC, 1989

3. Fleming, H.E., et. al.: Minimum Variance Simul taneous Retrieval of Temperature and Water Vapor from Satellite Radiance Measurements. 2nd Con ference on Satellite Meteorology/Remote Sensing and Application, Williamsburg, Va., pp. 20-23, May 13-16, 1986.

4. Fleming, H. E., et. al.: Operational Implemen tation of the Minimum Variance Simultaneous Retrieval Method. 3rd Conference on Satellite Meteorology and Oceanography, Anaheim, Ca., pp. 16-19, Feb. 1-5, 1988.

5. Fleming, H.E., D. S. Crosby, A. C. Neuendorffer: Correction of Satellite Temperature Retrieval Errors Due to Errors in Atmospheric Transmittances. J. Climate Appl. Meteor., 25, 1986.

6. Goldberg, M.D., et. al.: A Method for Obtaining an Improved Initial Approximation for the Tempera ture/ Moisture Retrieval Problem. 3rd Conference on Satellite Meteorology and Oceanography, Anaheim, CA., Feb. 1-5, 1988.

7. McMillin, L.M. and C. Dean: Evaluation of a New Operational Technique for Producing Clear Radiances, Journal of Applied Meteorology, Vol. 12, No.7, July, 1982

8. Reale, A.L., et. al.: Higher Resolution Operational Satellite Retrievals. 2nd Conference on Satellite Meteorology/Remote Sensing and Application, Williamsburg, VA., May 13-16, 1986.

Spectroscopic Measurement of Atmospheric Water Vapor and Schemes for Determination of Evaporation From Land and Water Surfaces Using the Airborne Visible/Infrared Imaging Spectrometer (AVIRIS)

James E. Conel, Robert O. Green, Veronique Carrere, Jack S. Margolis, Gregg Vane, Carol Bruegge, and Ronald Alley

Jet Propulsion Laboratory
California Institute of Technology
Pasadena, California 91109

ABSTRACT

This paper explores use of the Airborne Visible/Infrared Imaging Spectrometer (AVIRIS) to: (1) retrieve atmospheric water vapor distributions using band ratio techniques, (2) validate column abundances obtained by comparison with existing techniques, including infrared and microwave methods, (3) explore retrieval of estimates of surface moisture fluxes via the equation of continuity for water vapor in the atmosphere. AVIRIS provides spectral data to generate precise spatial variations of atmospheric water abundance. The reflective infrared techniques used are precise compared to passive infrared and microwave methods. Validation of the absolute precision of recovery depends upon instrumental factors, upon specification of scattering characteristics of the atmosphere, and upon compensation for variations in surface reflectance from soil or plant moisture or surface mineralogy across the spectral intervals employed. These are currently under way. Studies are also underway to evaluate sensitivity of the discussed techniques to scattering and surface reflectance influences to which they are vulnerable. The paper explores possibilities of retrieving air motions and flux conditions on moisture at the atmosphere-ground interface using column moisture abundances as retrieved by AVIRIS as tracers.

Keywords: Remote Sensing, Imaging spectrometry, atmospheric water vapor, evaporation.

1. INTRODUCTION

This paper analyses atmospheric water vapor in imaging spectrometer data over land surfaces. Atmospheric water vapor is an important reservoir in the global hydrosphere having profound implications for maintenance of fresh water supplies over land areas, for energetics of the atmosphere, and for modelling of the Earth's hydrologic cycle and climate. Measurement of land surface evaporation is difficult, locally or on a regional basis. The present methods, or extensions thereof, offer a good way to determine column water abundance, and in principle regional evaporation. In the present and future studies we want to capitalize on the spectral resolution and imaging capabilities of the Airborne Visible/Infrared Imaging Spectrometer (AVIRIS) to provide multitemporal maps of the column abundances of atmospheric precipitable water over field sites containing wet and dry surfaces at unprecedented high spatial resolution (~20. m) for measurements of this type. From changes in the distributions with time we will explore the possibility of obtaining estimates of average wind speed in the water vapor boundary layer. Estimates of the vertical and horizontal eddy diffusion coefficients may also be derivable by such observations. Water vapor has long been used as an atmospheric tracer. Kleinschmidt (1959) and Danielsen and Reiter (1960) used the conservation of specific humidity to determine the accuracy of isentropic wind trajectory analyses. The most ambitious goal we set, which can be studied regionally with AVIRIS and globally using the High Resolution Imaging Spectrometer (HIRIS) and the Moderate Resolution Imaging Spectrometer (MODIS) both on the Earth Observing System (Eos), is to measure the surface evaporative flux of water. This can be accomplished in principle by evaluating terms of the vertically integrated equation of continuity, suitably time and spatially averaged, to recover surface sources. Another improved product of the proposed atmospheric water analysis will be better maps of the surface distribution of plant (and soil) moisture. These maps, which can in principle be derived independently from analysis of the surface spectrum, will themselves be useful for identification of surface moisture sources, and from spectral properties and surface temperature information, the evapotranspirative/evaporative potential of vegetation/water surfaces. This information is important for localization of indigeneous as opposed to advected moisture components in the image representations.

2. BACKGROUND

Description of AVIRIS. AVIRIS is a whiskbroom-scanner that acquires images in 220 contiguous 10 nm spectral channels throughout the spectral region 400-2450 nm. The instantaneous field-of-view is 10^{-3} radians, representing a surface spot size of 20 m from the NASA U2 flight altitude of 20 km above terrain. A scan angle of $\pm 15^{\circ}$ generates an image of swath width 10.5 km on the ground. The data possess high geometric fidelity and can ordinarily be superposed without correction directly on 7 1/2 minute USGS topographic maps. The status of AVIRIS during the 1987 overflight of Mountain Pass is described by Green et al.(1988).

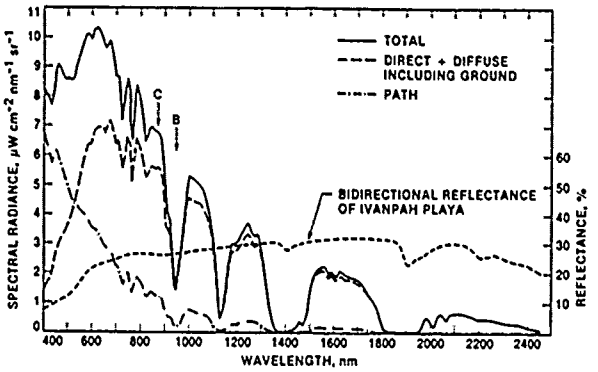


Figure 1. total, direct + diffuse, and path spectral radiance at 10 nm AVIRIS spectral resolution and 20 km observational altitude computed from LOWTRAN 7 standard MLS model for field-determined Ivanpah Playa reflectance.

Spectral method of water vapor measurement. In contrast to mid-infrared and microwave techniques (Isaacs, et al., 1986) that depend upon analysis of emitted radiation for retrieval of column water vapor abundances, the present method relies upon measurement of reflected solar radiation and the determination of water absorption band depth by a band ratioing technique. This method, employing the 1400 nm water absorption was the experimental basis for water detection and measurement on Mars (Farmer, et al., 1977). It has heretofore been suggested for terrestrial satellite applications by Gorodetsky and Syachinov (1983). The band ratio method indexes absorption band depth by forming a ratio between radiance in adjacent continuum. Figure 1 shows the spectral radiance at AVIRIS resolution for an altitude of 20 km according to the standard midlatitude summer model of LOWTRAN 7 (Kneizys, et al., 1989), depicting the atmospheric water absorption and continuum bands of interest in this work. Band ratios are calibrated for path water abundance using the LOWTRAN 7 model (including atmospheric scattering) to provide water vapor curves of growth, particularized for illumination and site elevation. The LOWTRAN 7 model contains an improved absorption line list for these bands compared to LOWTRAN 6, constructed by the same authors, and in addition calculates the contribution of multiple aerosol scattering in the emergent flux. All such curves of growth for the bands of Figure 1 are accurately (r^2 = 0.98 or better) found by curve fitting to be of the form RATIO = Aexp(-B/w), where w is defined as the spectroscopic abundance of water in precipitable cm. Reflectance techniques employing the present bands for recovery of column water abundance are insensitive to temperature distribution in the atmosphere. Adjustments for pressure are incorporated on a site specific basis using the LOWTRAN 7 model. The method benefits in sensitivity from a double path, downward slant and upward vertical, through the atmosphere. The precision of the technique for recovery of path water abundance can be determined from the ratio curve of growth formula as $\sigma_w/w = (2/B/w)(\sigma_R/R)$, where σ_w/w is the fractional uncertainty in the water determination and σ_R/R is the fractional uncertainty in the ratio determination. For an instrumental signal-to-noise ratio of 100 (1% determination of radiance) the ratio is determined to 2%. Using the 1125 nm band for which B = 0.92, and w = 3 pr. cm, gives σ_w = 0.07 pr. cm or a fractional uncertainty of 2.4%. The technique thus offers the definite possibility of improvement in column abundance determinations over mid-infrared or microwave recovery techniques (see Isaacs, et al. (1986), for a discussion of precision and accuracy of these methods). The present method is influenced by atmospheric scattering, by atmospheric temperature structure, and by the surface (background) reflectance. Water bands in vegetation and in standing water with suspended soil load contribute extra apparent atmospheric moisture. These bands must be analyzed independently using AVIRIS radiances outside water bands of atmospheric origin.

Example of Water Mapping with AVIRIS. Conel et al. (1988) analyzed AVIRIS image data from a desert site at Mountain Pass in eastern California 50 miles (80 km) south of Las Vegas, Nevada, using the 940 nm atmospheric water band. The result is shown in Figure 2. The abundances recovered decreased sharply with elevation within the site, which varies from 2600 feet (793 m) to 6800 feet (2073 m). Close agreement (~10%) was obtained between column abundances determined by the AVIRIS spectroscopic technique, and by a spectral hygrometric method also employing the 940 nm water band, and which is (by proxy) calibrated against vertically integrated water abundances obtained from radiosonde data. Direct comparison of the AVIRIS determined and radiosonde-measured is also planned.

3. INTERPRETATION OF AVIRIS-GENERATED ATMOSPHERIC WATER MAPS

The atmospheric water maps generated from AVIRIS imagery represent instantaneous views of the spatial variation of total column water abundance. These distributions represent (1) contributions and losses from local sources and sinks of water at the surface, (2) sources and sinks in the atmosphere itself by chemical reactions, and condensation and evaporation of clouds, and advection of water vapor by winds across vertical surfaces bounding the image area. In the following section we discuss possibilities for separation of some of these components from the images themselves, utilizing the high spatial and spectral resolution aspects of the data.

Advection-diffusion model for atmospheric distribution over changes in surface conditions; concentration and flux boundary conditions. What changes in atmospheric moisture distribution are to be expected as a partially saturated air mass moves from a uniform dry land surface over a water surface and back over dry surface? Brutsaert (1982) has provided a summary of previous studies of this problem. Here we will exploit a simple two-dimensional steady state advection-diffusion model of the atmosphere with dynamically uniform conditions, a source- and sink free atmosphere, and uniform surface boundary conditions over "wet" and "dry" regions. We illustrate the atmospheric water vapor distributions expected on the models together with changes in column abundances arising from discontinuities at the surface separating wet and dry regions. Three types of surface boundary conditions are possible (Sutton, 1953): (1) those representing constant specific humidity at the surface within a given region, (2) those representing constant flux, and a linear combination of (1) and (2), this latter being similar to the linearized radiation condition of the theory of heat conduction (Carslaw and Jaeger, 1963). The first two of these possibilities will be discussed here, and comparisons made between column abundance

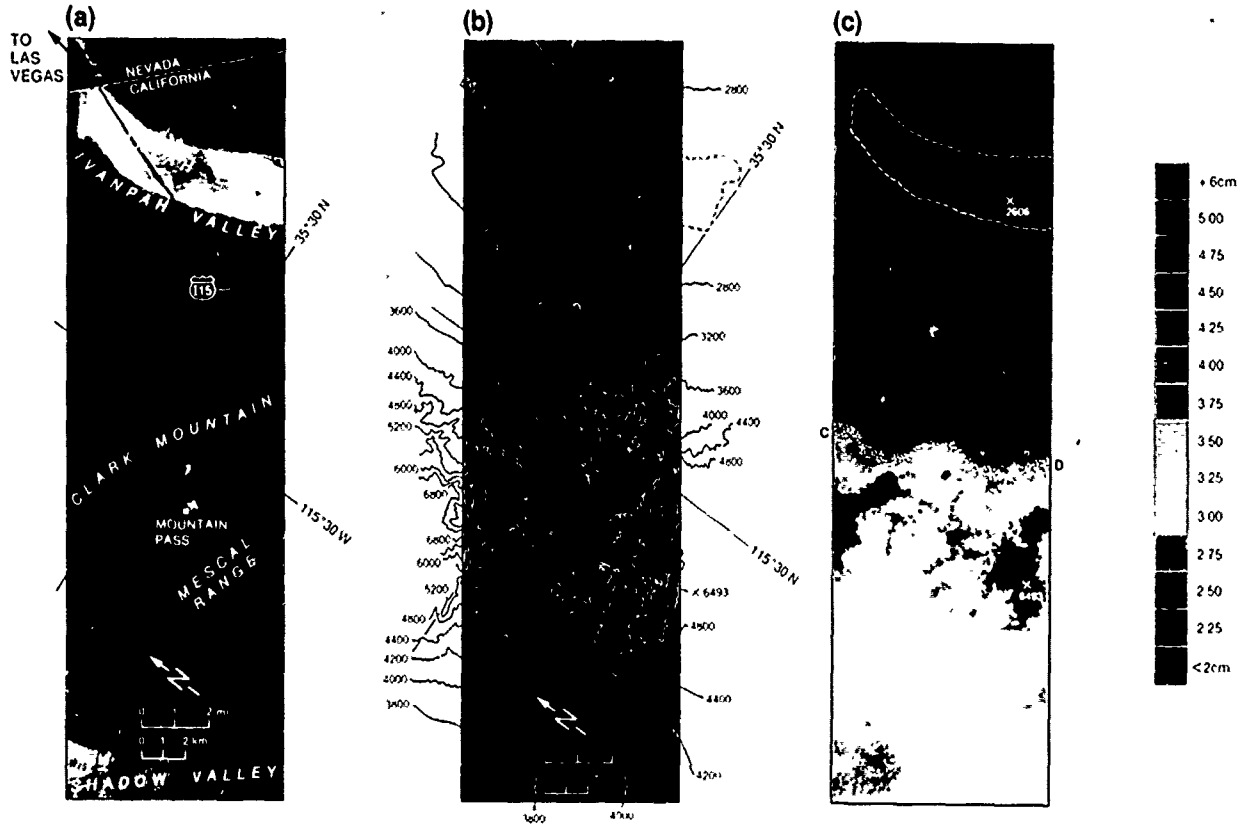


Figure 2. (a) AVIRIS image of Mountain Pass site. Point A locates atmospheric measurement station, B, playa target used for inflight radiometric calibration of AVIRIS. (b) generalized topographic map of site, (c) map of precipitable water distribution determined using the 940 nm atmospheric water band and the radiance ratio L(940)/L(870). Stippled region CD marks possible edge of water vapor tapered against topography. For details see Conel et al. (1988).

(CA) patterns resulting from each assumption. Condition (1) of constant specific humidity may be identified with the saturation value at the surface, which depends upon surface temperature and the nature of the evaporating liquid (Sutton, 1953). Condition (2) specifies the local rate of evaporation.

The two dimensional equation of continuity for the mean specific humidity $\bar{q}$ (in gm H_2O/gm moist air) is

$$\bar{u}(z)\frac{\partial \bar{q}}{\partial x} = \frac{\partial}{\partial z}\left(K_v(z)\frac{\partial \bar{q}}{\partial z}\right) \qquad (1)$$

where the +x-axis is horizontal and directed to the right, parallel to the ground surface, the +z-axis vertical upward, normal to the right. The y-axis is perpendicular to the x,z-plane. $\bar{u}(z)$ is the mean velocity parallel to the x-axis in cm/sec positive to the right, and $K_v(z)$ is the vapor eddy diffusivity in cm^2/sec (ref. 3). The mean wind speed and vertical vapor eddy diffusivity can be represented by

$$\bar{u}(z) = az^m \qquad (2)$$

$$K_v(z) = bz^n \qquad (3)$$

with $n = 1 - m$. For uniform vertical conditions in (2) and (3), $m = n = 0$. The following boundary conditions are adopted to describe consequences of changes in surface conditions.

$$\bar{q}(x,z) \to 0 \qquad x \to \pm\infty,\ z \to \infty \qquad (4a)$$

$$\bar{q}(0,z) = 0 \qquad 0 < z < \infty \qquad (4b)$$

$$\bar{q}(x,0) = \bar{q}_o \qquad 0 < x < L \qquad (4c)$$

$$\bar{q}(x,0) = 0 \qquad x > L \qquad (4d)$$

Condition (4a) determines that the solution remains finite at infinity. Condition (4b) determines that the water vapor concentration on $x = 0$ is equal to zero, i.e., that no water enters the region via transport from the region $x < 0$ for $z > 0$. Conditions (4c) and (4d) describe supply of moisture over a strip of width L extending indefinitely in the direction $\pm y$. The quantity $\bar{q}_o$ is the equilibrium specific humidity of water vapor that is in contact with the surface (water body or land surface) at a specified temperature T_o.

Alternatively, the concentration boundary conditions (4c) and (4d) can be replaced by conditions on boundary fluxes, namely.

$$-\rho b\frac{\partial \bar{q}}{\partial z} = f_o, \quad z = 0, \quad 0 < x < L \qquad (5a)$$

$$-\rho b\frac{\partial \bar{q}}{\partial z} = 0, \quad z = 0, \quad x > L \qquad (5b)$$

where f_o is in gm $cm^{-2}sec^{-1}$, and ρ is atmospheric density in gm cm^{-3}. Condition (5b) states that neither supply nor deposition occurs over the surface specified.

The solution of Equation (1) and (2) and (3) subject to (4a) - (4d) and L large (Sutton's problem for a large wet surface) is (Brutsaert, 1982):

$$\bar{q}/\bar{q}_o = 1 - P(\nu,\xi) \qquad (6)$$

where $\nu = (1-n)/(2 + m - n)$ and

$$\xi = \frac{a}{b(2 + m - n)^2} \frac{z^{2+m-n}}{x}$$

$P(\nu,\xi)$ is the incomplete gamma function. If the wet region has fetch L, with boundary conditions (4c) and 4d) then

$$\bar{q}/\bar{q}_o = 1 - P(\nu,\xi) \qquad 0 < x < L, \quad z > 0 \qquad (8)$$

$$= P(\nu,\xi_1) - P(\nu,\xi) \qquad x > L \qquad (9)$$

where ξ is given by Equation 7, and

$$\xi_1 = \left(\frac{a}{b}\right) \frac{z^{2+m-n}}{(2+m-n)^2 (x-L)} \qquad (10)$$

For the case $m = n = 0$, i.e., uniform velocity and eddy diffusivity, and boundary conditions (4), the solution to (1) reduces to:

$$\bar{q}/\bar{q}_o = \text{erfc}\left(\frac{z}{2\sqrt{(bx/a)}}\right) \qquad 0 < x < L \qquad (11)$$

and

$$\bar{q}/\bar{q}_o = \text{erfc}\left(\frac{z}{2\sqrt{(bx/a)}}\right) - \text{erfc}\left(\frac{z}{2\sqrt{[b(x-L)/a]}}\right), \qquad x > L \qquad (12)$$

where erfc(x) = 1 - erf(x) and erf(x) is the error function.

For the flux boundary conditions (5a) and (5b) and $m = n = 0$, the solution for the distribution $\bar{q}(x,z)$ may be written down from an analogous problem in heat conduction (Carslaw and Jaeger, 1959):

$$\bar{q} = \frac{2F_o}{\rho b} \left[\sqrt{(bx/a)}\,\text{ierfc}\left(\frac{z}{2\sqrt{(bx/a)}}\right)\right] \qquad 0 < x < L \qquad (13)$$

and

$$\bar{q} = \frac{2F_o}{\rho b} \left[\sqrt{(bx/a)}\,\text{ierfc}\left(\frac{z}{2\sqrt{(bx/a)}}\right) - \sqrt{[b(x-L)/a]}\,\text{ierfc}\left(\frac{z}{2\sqrt{[b(x-L)/a]}}\right)\right], \qquad x > L \qquad (14)$$

and ierfc(x) is the first repeated integral of the error function.

The column abundance CA is defined as

$$CA = \int_0^\infty \bar{q}(x,z)\,dz \qquad (15)$$

and may be evaluated easily for the distributions represented by (11) and (12) and (13) and (14).

For the concentration boundary conditions (4c) and (4d),

$$CA = 1.128\bar{q}_o\sqrt{(bx/a)} \qquad 0 < x < L \qquad (16)$$

and

$$= 1.128\bar{q}_o \left(\sqrt{(bx/a)} - \sqrt{[b(x - L)/a]}\right) \qquad x > L \qquad (17)$$

For the flux boundary conditions (5a) and (5b),

$$CA = (F_o/a)x \qquad 0 < x < L \qquad (18)$$

and

$$= F_o/a \qquad x > L \qquad (19)$$

The solutions (11) and (12) and (13) and (14) are plotted in Figure 3 with $a = 200$ cm sec^{-1}, $b = 10^5$ cm^2 sec^{-1} and width L of the the wet boundary along the wind direction equal to 10 km. The column abundances given by (16) and (17) and (18) and (19) are also depicted. Variations of CA with horizontal distance reflect the nature of these conditions and changes therein with position. The fall-off in concentration CA for x > 10 km results from removal of moisture at the surface to maintain $\bar{q}(x,0)$ equal

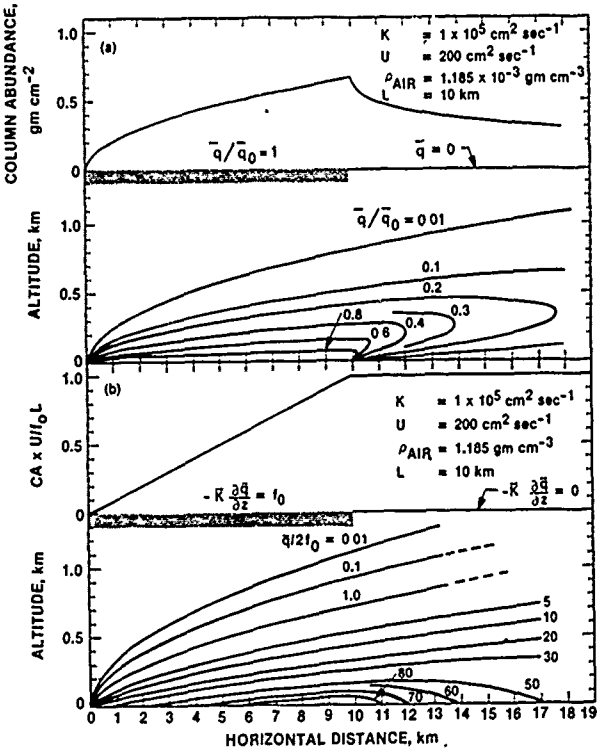


Figure 3. Atmospheric advection-diffusion model calculations for simple changes in surface boundary conditions. (a) step function in surface specific humidity, (b) step function in surface normal flux. Top curve of each panel is water vapor column abundance, the quantity retrieved from band-ratio analysis of AVIRIS radiance data.

to zero. The condition of zero normal flux for x > 10 km in the second example determines that the flux CA is constant, since water vapor is neither added to nor extracted from the atmosphere in this region. Changes in CA reflect changes in surface sources; this would be expected to hold for transient (i.e., $\partial\bar{q}/\partial t \neq 0$) conditions as well.

Isolation of surface fluxes from averaging water column abundances over time. An alternative means of estimation of regional land surface evaporation or evapotranspiration may be available if a sufficiently long series of atmospheric observations is available. Acquisition of a suitable series may not be feasible

with AVIRIS, but an improved observational situation can be visualized for HIRIS and MODIS during the operational phase of the Earth Observing System (Eos). Water vapor column abundances may be analyzed taking into account the characteristics of atmospheric dispersion (Csanady, 1973; Pasquill and Smith 1983; Pielke, 1984). Consider again the simple two-dimensional case, this time including time dependence. Following Peixoto (1973) and Peixoto and Oort (1983), the conservation equation for atmospheric water is

$$\frac{\partial \overline{W}}{\partial t} + \frac{\partial \overline{Q}}{\partial x} = -\left[K_v(z)\frac{\partial \overline{q}}{\partial z}\right]_{z=0} = f_o \qquad (20)$$

where

$$W = \int_o^\infty \overline{q}(x,z,t)dz \qquad (21)$$

and

$$Q = \int_o^\infty \overline{u}(z)\overline{q}(x,z,t)dz \qquad (22)$$

The right hand side of (20) represents the net vertical flux of moisture at the lower boundary z = 0. In writing (20) it has also been assumed that the vertical specific humidity gradient $\partial\overline{q}/\partial z \to 0$ as $z \to \infty$. To arrive at (20) the atmospheric source and sink terms of Peixoto and Oort (1983) representing the difference between evaporation and precipitation evaluated at the surface, have been expressed in terms of a net flux of moisture at the lower boundary. This is the term of interest here. Observations with AVIRIS from high altitude provide W(x,t), after adjustment for the double pass of light through the atmosphere. From a multitemporal data set a value of $\partial W/\partial t$ may be obtained. The interpretation of Q(x,t) in Equation 22 is more problematical observationally with AVIRIS. Lacking data on the form of $\overline{u}(z)$, which will in general be the case, it is still possible to define a convection velocity U_c by

$$Q = U_c \int_o^\infty \overline{q}(x,z,t)dz = U_c W \qquad (23)$$

For a continuous source and steady conditions without phase changes, U_c is the bulk velocity of the material at x (see Csanady, 1973, p. 119). Probably the best that can be achieved with AVIRIS, without support by surface or other data on winds, is to estimate this factor from changes in W(x,t) with position and time contained in the data. In the absence of outside (advected) sources of moisture, the surface flux f_o could be evaluated numerically from equation (20) using (23) provided U_c were known. Note that for determination of the flux by this approach, evaluation of neither $K_v(0)$ nor $\partial\overline{q}/\partial z = 0$ is required.

To progress with extraction of the surface flux in more general terms, it appears necessary to resort to time and space averaging. Denoting time averages by overbars and space averages by < >, gives for (20),

$$\left\langle \frac{\partial \overline{W}}{\partial t} \right\rangle + \left\langle \frac{\partial \overline{Q}}{\partial x} \right\rangle = \langle \overline{f}_o \rangle \qquad (24)$$

where $\overline{Q} = \overline{U}_c\overline{W}$. Except for short times associated with severe storms, $\langle \partial\overline{W}/\partial t \rangle$ is very small compared to the other terms (Peixoto and Oort, 1983, p. 9). Thus $\langle \overline{f}_o \rangle$ can be evaluated from suitable regional averages of $\overline{Q}$.

Estimation of wind speed and dispersion charactristics in the water vapor boundary layer. Modelling requires specification of wind speed, diffusivities, and surface parameters, all as functions of position and time. Retrieval of wind speed in the water vapor boundary layer may be possible by analysis of time varying patterns of column abundance distributions mapped in multitemporal AVIRIS data sets. The use of water vapor as an atmospheric tracer was employed by Kleinschmidt (1959) and by Danielsen and Reiter (1960). Eloranta et al. (1975) used lidar to measure motions of aerosol density inhomogenities, and with cross-correlation techniques estimated the horizontal radial wind component from their data. Movement of water vapor column density fluctuations for estimation of bulk wind speed are feasible provided data sets closely spaced in time (~ 15 minute are feasible with AVIRIS) are available. In the case of AVIRIS a measure of the average speed projected onto a plane defined by the AVIRIS image can be obtained.

4. SUMMARY

We report observations of the spatial variation of atmospheric precipitable water using the Airborne Visible/Infrared Imaging Spectrometer (AVIRIS), 450-2500 nm spectral range, 10 nm spectral sampling interval, 20 m ground resolution) over the Mountain Pass area in eastern California. maps are generated using a band ratio method and the 940 atmospheric water band and 870 nm continuum radiances. The ratios yield total path precipitable water from curves of growth, supplied by the LOWTRAN 7 atmospheric model, and include multiple scattering contributions. An independent validation of the AVIRIS-derived water column abundance is supplied through atmospheric optical depth measurements at 940 nm and by measurements with a ratioing spectral hygrometer that uses the 940 nm band and 870 nm continuum radiances. The spectral hygrometer is used as a proxy instrument to connect with a radiosonde scale of vertical atmospheric moisture distribution. Agreement to within about 10% is found between the spectroscopic (LOWTRAN 7) and radiosonde determinations. The precision of recovery of column abundances is superior to that provided by presently operational infrared and microwave sounding techniques. The atmospheric water values found at Mountain Pass conform to topography, and the amount of precipitable water decreases with elevation. The edge of the water vapor boundary layer appears to be present, truncated against the topography.

Regional evaporation and evapotranspiration are difficult to estimate observationally. The availability of maps of the column water abundance, and in multitemporal experiments the changes in abundances with position and time, suggest the possibility of recovering local bulk convective velocities of transport and values for the eddy diffusion coefficient in the water vapor boundary layer. With suitable time and space averaging of such observations the surface evaporative or evapotranspirative flux may be recoverable by an application of the equation of continuity for water vapor in the atmosphere. The high spatial resolution of AVIRIS data (20 m) offers the possibility of detailed examination of local evaporative and evapotranspirative sources and thereby the potential for separation in the imagery of externally-derived (i.e., advected) versus locally-derived (i.e. evaporative) sources of atmospheric water.

5. REFERENCES

1. Brutsaert, W., Evaporation into the atmosphere, D. Reidel, Dordrecht, 299 pp., 1982.

2. Conel, J.E., R.O. Green, V. Carrere, J.S. Margolis, R.E. Alley, G. Vane, C.J. Bruegge, and B.L. Gary, Atmospheric water mapping with the Airborne Visible/Infrared Imaging Spectrometer (AVIRIS), Mountain Pass, CA., in Proceedings of the Airborne Viaible/Infrared Imaging Spectrometer (AVIRIS) Performance Evaluation Workshop (G. Vane, ed.). JPL Pub. 88-38, 21-26, 1988b.

3. Csanady, G.T., Turbulent Diffusion in the Environment, D. Reidel, Dordrecht, 248 pp., 1973.

4. Danielsen, E.F., and E.R. Reiter, Bemerkungen en E. Kleinschmidt: Nicht-adiabatische Abkuhlung im Bereich "des jet stream", Beitr. Phys. Atmos., 32(3 and 4), 265-273.

5. Eloranta, E.W., J.M. King, and J.A. Weinman, The determination of wind speeds in the boundary layer by monostaticlidar, J. Appl. Met., 14, 1485-1489, 1975.

6. Farmer, C.B., D.W. Davies, A.L. Holland, D.D. LaPorte, and P.E. Doms, Mars: water vapor observations from the Viking orbiters, J. Geophys. Res., 82(28), 4225-4248, 1977.

7. Gorodetsky, A.K., and V.I. Syachinov, Remote sensing of atmospheric water content from satellites, in Variations in the Global Water Budget (A. Street-Perrott, M. Beran, and R. Ratcliffe eds.), D. Reidel, Dordrecht, 81-87, 1983.

8. Green, R.O., G. Vane, and J.E. Conel, Determination of aspects of the inflight spectral, radiometric, spatial and signal-to-noise performance of the Airborne Visible/Infrared Imaging Spectrometer over Mountain Pass, Ca., in Proceedings of the Airborne Visible/Infrared Imaging Spectrometer (AVIRIS) Performance Evaluation Workshop, JPl Pub. 88-38, 162-184, 1988.

9. Isaacs, R.G., R.N. Hoffman, and L.D. Kaplan, Satellite remote sensing of meteorological parameters for global numerical weather prediction, Rev. Geophys., 24(4), 701-743, 1986.

10. Kleinschmidt, E., Jr., Nicht-adiabatische abkuhlung im bereich der jetstream, Beitr. Phys. Atmos., 32(1-2), 824-836.

11. Kneizys, F.X., E.P. Shettle, G.P. Anderson, L.W. Abrew, J.H. Chetwynd, J.E.A. Shelby, and W.O. Gallery, Atmospheric Transmittance/Radiance; Computer Code LOWTRAN 7, (in press), AFGL Hanscom AFB, MA., 1989.

12. Pasquill, F., and F.B. Smith, Atmospheric Diffusion, Ellis Horwood Ltd., 437 pp., 1983.

13. Peixoto, J.P., Atmospheric vapor flux computations for hydrological purposes, WMO Report No. 20, World Meteorological Organization, Geneva, 83, pp., 1973.

14. Peixoto, J.P., and A.H. Oort, The atmospheric branch of the hydrological cycle, in Variations in the Global Water Budget, (A. Street-Perrott, M. Beran, and R. Ratcliffe, eds.), D. Reidel, Dordrecht, 5-65, 1983.

15. Pielke, R.A., Mesoscale meteorological modelling, Academic Press, New York, 1984.

6. ACKNOWLEDGMENTS

Discussions with Professor Fred Shair of California Institute of Technology about atmospheric diffusion are most gratefully acknowledged. This research was carried out at Jet Propulsion Laboratory, California Institute of Technology, under a contract with che National Aeronautics and Space Administration.

Column Atmospheric Water Vapor Retrievals from Airborne Imaging Spectrometer Data

Bo-Cai Gao and Alexander F. H. Goetz
Center for the Study of Earth from Space/CIRES
University of Colorado
Boulder, Colorado 80309- 0449

ABSTRACT

High spatial resolution column atmospheric water vapor amounts were derived from spectral data collected by the Airborne Visible Infrared Imaging Spectrometer (AVIRIS). The quantitative derivation is made by curve fitting observed spectra with calculated spectra in the 1.14 μm and 0.94 μm water vapor band absorption regions with a non-linear least squares technique. The precision of the retrieved column water vapor amounts is approximately 5%. The derived column water vapor amounts are independent of the absolute surface reflectance. Curve fitting of spectra near 1 μm from areas covered with vegetation indicates that both the amount of atmospheric water vapor and the moisture content of vegetation can be retrieved simultaneously. It now appears feasible to derive high spatial resolution column water vapor amounts over land areas from satellite altitude with the proposed High Resolution Imaging Spectrometer (HIRIS) and possibly with the proposed Moderate Resolution Imaging Spectrometer (MODIS).

Key Words: Column water vapor, AVIRIS, HIRIS, Imaging Spectrometry

INTRODUCTION

High spatial resolution, column water vapor amount can be useful in the studying of the rate of energy exchange between the earth's surface and the atmosphere and self-calibration of imaging spectrometer radiance measurements. In this paper, the integrated water vapor content from ground to space in a vertical path is defined as the column water vapor amount in units of cm.

We propose a method for quantitative retrievals of high spatial resolution column atmospheric water vapor from spectra obtained with the Airborne Visible Infrared Imaging Spectrometer (AVIRIS) (Vane, 1987) over land areas on clear days. The method consists of curve fitting observed spectra with simulated spectra in the 1.14 μm or the 0.94 μm water vapor band absorption region. The derived water vapor values can be used to remove other water vapor absorption bands in the 0.4-2.5 μm region.

REMOTE SENSING OF ATMOSPHERIC WATER VAPOR

Atmospheric water vapor profiles and column amounts are obtained from orbit by measurement of atmospheric emission in the 6-12 μm infrared region or near 22.2 GHz in the microwave region. Examples are the HIRS2 (High Resolution Infrared Sounder) and MSU (Microwave Sounding Unit) onboard TIROS-N (Susskind, et al., 1984). Experience has shown that, below the 500 mb level, water vapor profiles derived from IR emission spectra have large uncertainties because of the contributions to the observed radiance from the surface and because of the insensitivity of the radiance to changes in humidity. The HIRS2 column water vapor measurement accuracies over land are only ±25% (Reuter et al., 1988) while microwave measurements are even less accurate.

IMAGING SPECTROMETRY

Imaging spectrometers are being developed for remote sensing of the land and coastal waters (Goetz et al., 1985). In the mid-1990s, NASA expects to carry, among others, two imaging spectrometers, the Moderate Resolution Imaging Spectrometer (MODIS) (Solomonson et al., 1989) and the High Resolution Imaging Spectrometer (HIRIS) (Goetz and Herring, 1989) aboard the Earth Observing System (EOS) Polar Platform (Butler et al., 1987). The precursor to HIRIS now under development is the Airborne Visible/Infrared Imaging Spectrometer (AVIRIS) (Vane, 1987). This instrument images the earth's surface in 224 spectral bands approximately 10 nm wide, covering the region 0.4-2.5 μm, from an ERS-2 aircraft at an altitude of 20 km. The ground instantaneous field of view (GIFOV) is 20 x 20 meters. Figure 1 shows an example of an AVIRIS spectrum. The main water vapor absorption features centered at approximately 0.94, 1.14, 1.4 and 1.9 μm are clearly visible. These water vapor features provide the opportunity for derivation of high spatial resolution column water vapor from AVIRIS spectra.

PRINCIPLES FOR RETRIEVAL OF COLUMN WATER VAPOR AND MOISTURE CONTENT OF VEGETATION FROM AVIRIS SPECTRA

According to Esaias et al., (1986), in simplified form, the radiance at a downward looking aircraft sensor can be written as

$$L_{Sensor}(\lambda) = L_{Sun}(\lambda)\,T(\lambda)\,R(\lambda) + L_{Path}(\lambda), \qquad (1)$$

where λ is the wavelength, $L_{Sensor}(\lambda)$ is the radiance at the imaging spectrometer, $L_{Sun}(\lambda)$ is the solar radiance above the atmosphere, $T(\lambda)$ is the total atmospheric transmittance, which is equal to the product of the atmospheric transmittance from the sun to the earth's surface and that from the surface to the aircraft, $R(\lambda)$ is the surface reflectance at the observational geometry, and $L_{Path}(\lambda)$ is the path scattered radiance, including effects of single scattering and multiple scattering. All these must be taken into account in order to derive the atmospheric water vapor amount.

The Choice of Water Vapor Bands to Be Fitted

In order to derive the column water vapor with high

precision from a water vapor absorption band in a measured spectrum, the transmittances of the band must be sensitive to the change in the number of water vapor molecules in the line of sight. The monthly means of column water vapor in different parts of the United States typically range between 0 5 and 4 3 cm (Iqbal, 1983) Figure 2 shows calculated atmospheric water vapor transmittance spectra between 0 6 and 2.8 μm for water vapor amounts of 0.63, 1.3, 2.5 and 5.0 cm in the line of sight These transmittance spectra have a spectral resolution of approximately 20 nm. The spectra were calculated with a spectral band model described below. These curves indicate that, for the typical atmospheric conditions, the 0.94 and 1.14 μm bands are sensitive to the changes in the amount of water vapor, while the 1.4 and 1 9 μm bands are relatively insensitive to water vapor amount changes. The other bands centered at about 0.72, 0.82 and 2 18 μm are too weak to allow accurate retrieval of column water vapor from measured spectra. The water vapor transmittances (not including continuous absorptions) in narrow spectral regions near 0.86, 1.04 and 1.24 μm are greater than 99%. These regions are often referred to as the "atmospheric windows."

A Band Model for Atmospheric Gaseous Absorption

In order to derive column water vapor amounts from AVIRIS spectra, it is necessary to compare calculated atmospheric spectra with observed spectra. The calculation requires both spectral absorption and atmospheric models.

Each AVIRIS image contains approximately 250,000 spectra, necessitating a fast algorithm for efficient data analysis. A narrow-band spectral absorption model approach is most appropriate for this purpose. The Malkmus (1967) model is used in our spectral calculations. The band parameters are calculated from the spectral line parameters compiled in the HITRAN Data Base (Rothman et al., 1987). An extension of the Curtis-Godson scaling approximation (Rodgers and Walshaw, 1965) is used to take account of the pressure and temperature variations along a typical atmospheric path.

The Reflectances of Soil, Rocks, and Vegetation

Changes in surface spectral reflectance within a water vapor absorption band will affect the retrieval accuracy. Surface reflectances of many natural targets in the visible and near infrared regions have been compiled by Bowker et al. (1985). The reflectance curves of soil and rocks have the common characteristic that the reflectances vary nearly linearly with wavelength in the 0.94 μm and the 1.14 μm water vapor band absorption region This property allows the quantitative derivation of column water vapor from an AVIRIS spectrum over soil or rocks, as discussed below. The reflectance spectrum of vegetation exhibits weak liquid water absorption features centered at approximately 1 0 and 1.2 μm The centers of the liquid water bands are shifted approximately 0.06 μm to longer wavelengths of the corresponding water vapor band centers. These shifts allow the simultaneous retrievals of column atmospheric water vapor amounts and the moisture contents of vegetation from AVIRIS spectra over vegetated areas, as described below.

Column Water Vapor Retrievals from AVIRIS Spectra over Non- Vegetated Areas Ignoring Scattering

If atmospheric scattering near 1 μm is neglected, the column water vapor amount over a non-vegetated area can be derived according to the principles described below.

The solid line in the top plot of Figure 3 shows a portion of a ratioed AVIRIS spectrum measured over Cuprite, NV against a solar radiance curve above the atmosphere. Neglecting scattering and based on the above discussions on surface reflectances and atmospheric transmittances, we may assume the 1.04 and 1.24 μm regions correspond to the 100% atmospheric gaseous transmission levels, and we may assume that the 100% transmission levels between these two windows can be joined with a straight line. These assumptions imply that the absolute surface reflectances and radiance values are not required to derive column water vapor amounts from AVIRIS spectra. Only a relative band-to-band calibration of AVIRIS data over the wavelength region used is required. Curve fitting the ratioed spectrum with a simulated atmospheric gaseous transmission spectrum with appropriate atmospheric and spectral models yields the quantitative amount of water vapor in the line of sight. This line of sight amount can then be converted to the vertical column amounts based on the observational geometry.

The non-linear least squares curve fitting technique described by Russell et al. (1988) is used in our retrievals. The dotted line in top plot of Figure 3 is the best fit to the ratioed spectrum. The residuals (observed/calculated) expressed as a percentage of the 100% transmission levels are shown in the bottom plot of Figure 3. These residuals have an rms of approximately 5%.

Column Water Vapor and Equivalent Liquid Water Thickness Retrievals from AVIRIS Spectra over Vegetated Areas, Ignoring Scattering

The top plot of Figure 4 shows an example of AVIRIS radiances over vegetation divided by solar radiances above the atmosphere. The liquid water absorption features are not discernable in the spectrum dominated by the 0.94 and 1.14 μm atmospheric water vapor bands. The bottom plot of Figure 4 shows the radiances over the same vegetated area but divided by the radiances from a nearby airport runway having nearly constant spectral reflectance around 1.0 μm. The liquid water absorption features near 1.0 and 1.2 μm along with the relative shifts of band centers are clearly seen.

We have assumed that the vegetational liquid water absorption features between 0.8 and 1.4 μm can be described by the absorption of an equivalent amount of liquid water. The calculation of transmittances is made by using the liquid water absorption coefficients compiled by Palmer and Williams (1974). By modeling the shape of the vegetation spectra with the atmospheric band model and the liquid water absorption model, both the water vapor amount and the equivalent liquid water thickness can be derived.

Path Radiance

Typically, the atmospheric aerosols are mostly concentrated in the lower 1-2 km of the atmosphere, as are the water vapor molecules. Near 1.0 μm the Rayleigh scattering is negligible and the main contribution to the path radiance is the scattering by aerosols. When the aerosol concentrations are small (visibilities of 20 km or greater), the scattered radiation near 1.0 μm is typically of the order of a tenth of the direct reflected solar radiation (Iqbal, 1983).

We have found that AVIRIS spectra measured over bodies of water on clear days show the 0.94 and 1.14 μm atmospheric water vapor band absorption features. The radiation near 1 μm over water did not result from any radiation emerging from beneath the water surface, because any solar radiation at this wavelength is completely absorbed, but rather from atmospheric scattering. Sunglint may also contribute to the observed radiances. The water vapor absorption features in spectra from the water areas are attributed to aerosol scattering along with water vapor absorption within the lower few kilometers of the troposphere.

Comparisons of spectra over the water areas and over the nearby solid surface areas show that the shapes of both sets of spectra in the 0.94 and 1.14 μm water vapor band absorption regions are similar. Therefore, to a first order approximation,

the path scattered radiation near 1 μm can be treated as a fraction of direct reflected radiation. This means that when the aerosol concentration is small, column water vapor and equivalent liquid water thickness can still be obtained with the simple absorption models described above.

RESULTS

In order to test the ability to derive column water vapor and equivalent liquid water thickness of vegetation from AVIRIS spectra with our spectral curve fitting technique, retrievals were made from data collected over several sites. Results retrieved from data measured over the Cuprite mining district in southwest Nevada, Moffett Field near San Francisco Bay in California, and Rogers Dry Lake are described below.

Retrievals over the first two sites are made by curve fitting of the 1.14 μm water vapor band in 4-by-4 pixel averaged spectra, and retrievals over Rogers Dry Lake are made by curve-fitting both the 1.14 μm and the 0.94 μm water vapor bands in averaged spectra. During the 1987 AVIRIS measurements over Cuprite and Moffett Field there were no simultaneous measurements of atmospheric temperature, pressure and water vapor VMR profiles. Column atmospheric water vapor amounts over these sites were retrieved with the "mid-latitude summer" atmospheric model in LOWTRAN 6 (Kneizys et al., 1983). During the two 1988 AVIRIS measurements over Rogers Dry Lake, rawinsonde measurements of atmospheric temperature, pressure and water vapor VMR profiles were available. Column atmospheric water vapor amounts from the Rogers Dry Lake spectra were retrieved with the atmospheric profiles of the rawinsonde measurements and also with the "mid-atitude summer" model of LOWTRAN 6.

Cuprite, Nevada

The surface reflectances within the AVIRIS scene over Cuprite vary by more than a factor of 2. Figure 5a shows an image of the site. The bright regions represent areas with high reflectances at 1.037 μm the maximum being approximately 0.6. The dark regions represent areas with low reflectances at 1.037 μm the minimum being approximately 0.25. Within the scene, there is less than 5% surface vegetation cover. More than 85% of the areas within the scene have less than 150 meters of relief. There are no sources of water vapor within the scene or in the nearby areas. Therefore, it was assumed that the water vapor distribution within the scene was horizontally homogeneous. The purpose for choosing the Cuprite site was to test the ability to derive column water vapor from surfaces with very different albedos.

The mean of the column water vapor amounts over the entire scene is 0.97 cm, and the standard deviation 4.8%. An image processed from column water vapor values over the site is shown in Figure 5b. In the vicinities of major reflectance contrast, the derived column water values remain approximately the same. This demonstrates that the derived column water vapor values are relatively insensitive to the absolute surface reflectance values.

A topographic map of the scene is shown in Figure 5c. The small variations in column water vapor values over the scene are related to the surface topography rather than to the surface albedo differences.

Moffett Field, California

The AVIRIS scene over Moffett Field contains areas with bodies of water, an airport runway, and vegetation. The column water vapor amounts derived from spectra over vegetated areas are approximately equal to those derived from vegetation-free areas. The retrieved equivalent liquid water thicknesses of vegetation over different vegetated areas within the scene range from 0.05 to 0.3 cm.

Rogers Dry Lake, California

As mentioned earlier, two sets of AVIRIS data were collected over Rogers Dry Lake. Because the retrievals from both sets are similar, only the retrievals from the data measured on August 31, 1988 are described below.

The variation of surface elevation over the entire scene is less than 50 meters. Column water vapor values were derived by curve fitting the 0.94 μm band absorption feature and the 1.14 μm band absorption feature, respectively, by using the atmospheric temperature, pressure and water vapor VMR profiles measured from a rawinsonde that was released approximately one hour prior to the AVIRIS overflight. The mean column water vapor values of 2.796 cm ±3.06%, and 2.773 cm ±3.30%, agree very well with the 2.8 cm of integrated column water vapor from the rawinsonde measurement. The good agreement may be fortuitous since scattering is not included in our calculations and calculated transmittances with the Malkmus-band model are slightly dependent on the widths of the model parameters (J. Kiehl, private communication). Neverthess, the good agreement is encouraging. Similar retrievals were also made by using the "Mid- Lititude Summer Model." The mean difference in column water vapor values derived with the two atmospheric models is 1.1%.

MODEL SENSITIVITIES

The precision of the retrieved column water vapor amounts described above are controlled by the finite signal-to-noise ratio of the AVIRIS spectra and by the errors in the assumption of linear surface reflectances with wavelength in the water vapor band absorption regions. In addition to these, the accuracy of column water retrievals are affected by other factors, including errors in assumed atmospheric temperature, pressure and water vapor volume mixing ratio profiles, knowledge of the surface elevation at a GIFOV of 20 m, and errors in the instrument resolution. These errors have been investigated and their cumulative contributions to errors in retrieved column water vapor amounts are estimated to be less than 5%.

SUMMARY AND CONCLUSIONS

We have developed a method for the quantitative derivation, at high spatial resolution, of column atmospheric water vapor amounts from imaging spectrometer measurements of reflected solar radiation on clear days and when the atmospheric aerosol concentrations are small. The method makes use of the facts that: (1) the reflectances of many ground targets vary approximately linearly with wavelength in the 0.94 and 1.14 μm water vapor band absorption regions, (2) the scattered radiation near 1 μm is small compared to the direct reflected radiation when the atmospheric aerosol concentrations are small, and (3) the scattered radiation also contains the water vapor absorption features because both scattering and absorption processes occur within the troposphere. The quantitative derivation of column water vapor amounts is made by curve fitting of observed spectra with calculated spectra in the 1.14 and 0.94 μm water vapor band absorption regions with an atmospheric model, a narrow-band spectral model, and a non-linear least squares fitting technique.

The precision of the derived column water vapor amounts from sets of AVIRIS data is about 5% rms. The retrieved column water vapor amounts are independent of the surface reflectance. Therefore, the absolute surface reflectances and radiance values are not required for column water vapor retrievals when the concentrations of atmospheric aerosols are low.

Our method for column water vapor retrievals is expected to be valid for typical atmospheric conditions with visibilities of 20 km, or greater. This method is not directly applicable to situations when the atmospheric aerosol concentrations are

high, particularly when large amounts of stratospheric aerosols are present.

Simultaneous retrievals of column atmospheric water vapor amounts and the equivalent liquid water thicknesses of surface vegetation have also been made. The derived column water vapor amounts from spectra over vegetation agree with those derived from spectra over non-vegetated areas. The derived equivalent liquid water thicknesses is related to the number of liquid water molecules that interacted with the incident solar radiation, and therefore related to the moisture content of vegetation.

Because the observational nadir angle and the solar zenith angle are not zero, a spatial smearing effect is present in our derived column water vapor amounts. The magnitude of the spatial smearing is estimated to be 500 meters or greater and depends on the three dimensional water vapor distribution and on the solar and observational geometry. If over a few kilometer range the water vapor distribution is nearly uniform horizontally, the smearing effect is irrelevant.

Acknowledgments

Bo-Cai Gao was supported by a Cooperative Institute for Research in Environmental Sciences (CIRES) Visiting Fellowship and this study was partially supported under Contract #958039 from NASA/JPL.

References

1. Bowker, D. E., R. E. Davis, D. L. Myrick, K. Stacy, and W. T. Jones, Spectral Reflectances of Natural Targets for Use in Remote Sensing Studies, NASA Reference Publication 1139, 1985.
2. Butler, D. M., et al., From Pattern to Process: The Strategy of the Earth Observing System, NASA Earth Observing System, Vol. II, 1987.
3. Esaias, W., et al., Moderate Resolution Imaging Spectrometer, Earth Observing System, Vol. IIb, Instrument Panel Report, NASA, 49-53, 1986.
4. Goetz, A. F. H., G. Vane, J. Solomon, and B. N. Rock, Imaging spectrometry for Earth remote sensing, Science, 228, 1147, 1985.
5. Goetz, A. F. H., and M. Herring, The High Resolution Imaging Spectrometer (HIRIS) for Eos, IEEE Transactions on Geoscience and Remote Sensing, 27, 136, 1989.
6. Iqbal, M., An Introduction to Solar Radiation, Academic Press, 43-95, 1983.
7. Kneizys, F. X., E. P. Shettle, W. O. Gallery, J. H. Chetwynd, L. W. Abreu, J. E. A. Selby, S. A. Clough, and R. W. Fenn, Atmospheric Transmittance /Radiance: Computer Code LOWTRAN 6, AFGL-TR-83-0187, 1983.
8. Malkmus, W., Random Lorentz Band Model with Exponential-Tailed S Line Intensity Distribution Function, J. Opt. Soc. of Amer., 57, 323, 1967.
9. Palmer, K. F., and D. Williams, Optical properties of water in the near infrared, J. Opt. Soc. Amer., 64, 1107, 1974.
10. Reuter, D., J. Susskind, and A. Pursch, First-Guess Dependence of a Physically Based Set of Temperature-Humidity Retrievals from HIRS2/MSU Data, J. Atmos. and Oceanic Technology, 70-83, 1988.
11. Rodgers, C. D., and C. D. Walshaw, The computation of infra-red cooling rate in planetary atmosphere, Quart. J. Roy. Meteorol. Soc., 94, 99, 1966.
12. Rothman, L. S., R. R. Gamache, A. Goldman, L. R. Brown, R. A Toth, H. M. Pickett, R. L. Poynter, J.-M. Flaud, C. Camy-peyret, A. Barbe, N. Husson, C. P. Rinsland, and M. A. H. Smith, The HITRAN database: 1986 edition, Appl. Opt., 26, 4058, 1987.
13. Russell, J. M., III, C. B. Farmer, C. P. Rinsland, R. Zander, L. Froidevaux, G. C. Toon, B. C. Gao, and J. H. Shaw, Measurements of odd nitrogen in the stratosphere by the ATMOS experiment on Spacelab 3, J. Geophys. Res., 93, 1718, 1988.
14. Solomonson, V. V., W. L. Barnes, P. W. Maymon, H. E. Montgomery, and H. Ostrow, MODIS. Advanced Facility Instrument for Studies of the Earth as a System, IEEE Transactions on Geoscience and Remote Sensing, 27, 145, 1989.
15. Susskind, J., J. Rosenfield, and D. Reuter, Remote Sensing of Weather and Climate Parameters From HIRS2/MSU on TIROS-N, J. Geophys. Res., 89, 4677, 1984.
16. Vane, G., Ed., Airborne Visible/Infrared Imaging Spectrometer (AVIRIS), JPL Publication 87-38, 1987.

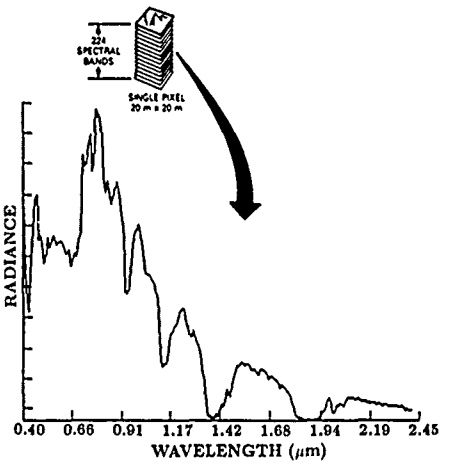


Figure 1 Spectrum from single AVIRIS pixel (from Vane, 1987).

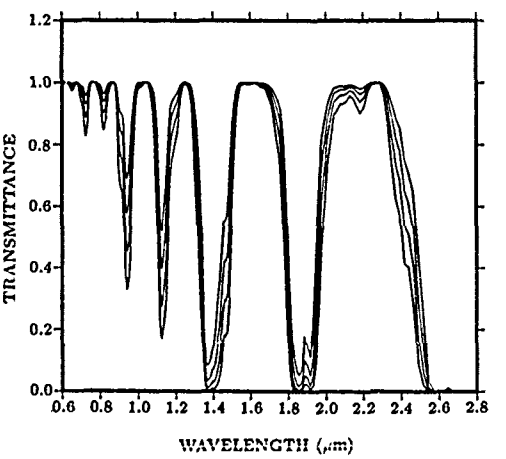


Figure 2 Vertical atmospheric transmittance vs wavelength at different water vapor amounts. The four curves from the top to the bottom correspond to column water vapor amounts of 0.63, 1.3, 2.5 and 5.0 cm. The calculations were made by using the spectral model described in Section 4.3 and the Tropical Model of LOWTRAN6 (Kneizys, et al., 1983).

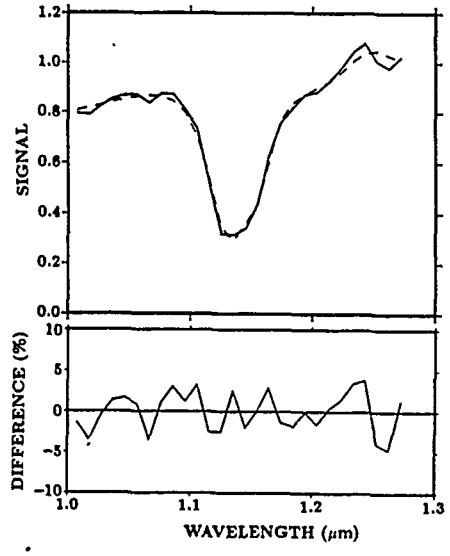


Figure 3 An example of curve fitting of spectra measured over a vegetation-free surface. The top plot shows the observed spectrum (solid line) and the fitted spectrum (dashed line). The bottom plot shows the percent differences between the observed and the fitted spectra.

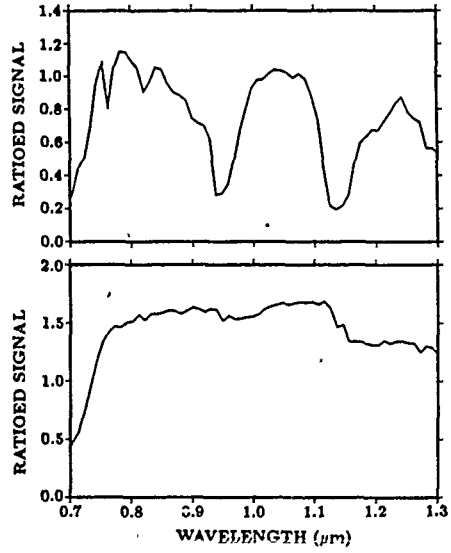


Figure 4 An AVIRIS spectrum of vegetation divided by the solar spectrum above the atmosphere (top plot), and the same AVIRIS spectrum of vegetation divided by a spectrum of an airport runway (bottom plot). The two features centered at approximately at 1.0 and 1.20 μm in the bottom plot are the absorption features of liquid water in vegetation.

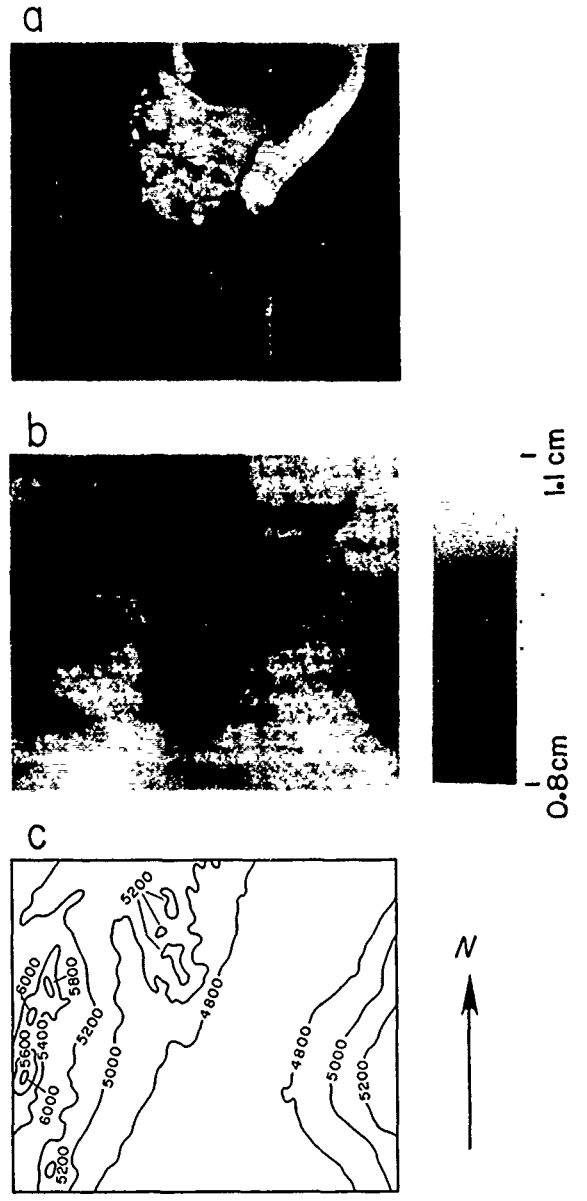


Figure 5 Column water vapor retrievals from AVIRIS data measured over Cuprite, NV on June 25, 1987. (a) an image of the scene processed from radiances of one channel centered at 1.037 μm; (b) a column water vapor image over the scene retrieved by curve fitting the 1.14 μm water vapor band absorption region; and (c) a topographic map of the scene. The elevations in (c) are in units of feet.

REMOTE SENSING OF THE THERMAL FORCING ON THE TROPICAL PACIFIC

W. Timothy Liu[1] and Catherine Gautier[2]

[1] Jet Propulsion Laboratory, California Institute of Technology, Pasadena, CA 91109
[2] California Space Institute, Scripps Institute of Oceanography, La Jolla, CA 92093

ABSTRACT

Monthly fields of shortwave radiation (SR) and latent heat flux (LE) over the central and eastern tropical Pacific between 1980 and 1983 have been computed using satellite data. They are the dominant variable components of surface thermal forcing on the ocean in this time scale. During the 1982-1983 El Nino Southern Oscillation episode, surface wind convergence and cloudiness associated with the displacement of equatorial organized convection caused a reduction in both the SR into the ocean and the LE out of the ocean. The lag correlation coefficients between the forcing (SR-LE) and the sea surface temperature is found to be significantly high outside the equatorial region, showing that surface thermal forcing is the dominant factor in sea surface temperature change. In the narrow equatorial wave guide, ocean dynamics play a more important role, and surface heat flux is a consequence rather than the cause of sea surface temperature change.

Key Words: Oceanography, Air-sea Interaction, Climate, Remote-sensing

1. INTRODUCTION

One of the most prominent climatic signals with time scales from a few month to a few years is the El Nino Southern Oscillation (ENSO). It is manifested by the ocean surface warming in the central and eastern equatorial Pacific. Understanding the interannual variability of sea surface temperature in the tropical ocean is a major objective of the decade long Tropical Ocean and Global Atmosphere (TOGA) program. The role of surface thermal forcing on upper ocean heat balance and circulation is poorly known. In the past, the forcing was derived from meteorological reports from merchants ships and fishing vessels whose distribution is inadequate to describe ENSO variability. Satellite sensors have the potential of providing repeated and uniform coverage. Shortwave radiation (SR) and latent heat flux (LE) are the dominant variable components of surface thermal forcing on the tropical Pacific in these time scales. The TOGA Heat Exchange Project (THEP) is a research effort initiated to derive SR and LE using a combination of satellite data (Liu and Niiler, 1985), and this study is part of the effort.

2. LATENT HEAT FLUX

Of the three parameters required to compute LE, satellite sensors can measure sea surface temperature (T_s) and surface wind speed (U), but cannot measure surface humidity (Q). It was found that at the monthly time scale, atmospheric water has a single dominant mode of variability and Q can be derived from the columnar water vapor (W) measured by microwave radiometers. A statistical Q-W relation was established using 17 years of radiosonde reports from mid-ocean meteorological stations (Liu, 1986). This relation was found to be adequate in describing the seasonal and interannual variations over global oceans except in high latitudes during summer, with accuracy estimated to be 0.4-0.8 g/kg. With this relation, monthly fields of LE from 1980 to 1983 in the tropical Pacific were computed using data from SMMR (Scanning Multichannel Microwave Radiometer) on Nimbus-7. In comparison with monthly data from equatorial moorings and atolls, the scatters were found to be 0.6 m/s in U, 0.8°C in T_s and 0.4 g/kg in Q. The random error in LE is estimated to be 26 W/m^2 (Liu, 1988).

3. SHORTWAVE RADIATION

The annual and interannual variabilities of SR depends largely on the variability of cloud-cover. The technique for computing SR developed by Gautier et al. (1980) capitalizes on the high repetition and high spatial resolution of VISSR (Visible Infrared Spin Scan Radiometer) on geostationary satellite to monitor cloud changes. In this study, the radiance observed by VISSR on GOES-W was first calibrated against a standard target on Earth (White Sands). Surface shortwave radiation is then computed by inverting radiation transfer equations separately for cloudy and clear pixels. The SR computed has been compared with in situ measurements during a number of field experiments in the tropical oceans (e.g., Gautier, 1982) and in extratropical oceans (e.g., Gautier and Katsaros, 1984). The statistics indicate that the accuracy of daily estimates is approximately 15 W/m^2.

4. VARIABILITY IN THE TROPICAL PACIFIC

Fig. 1 shows the time-longitude distribution of Ts, U, W, LE, SR and net flux (SR-LE), centered on the equator between 90°W and the date-line. The 1982-83 ENSO episode is envisioned as an apparent eastward migration of the warm water pool marked by the 28°C isotherm starting in June 1982. This results in a reverse of zonal T_s gradient near the date-line. The organized deep convection marked by high W also moves east from the date-line starting June 1982, leaving dry air behind. The seasonal cycle of U is disrupted by the eastward migration of low wind center representing surface convergence associated with the organized convection at the eastern terminal of anomalous westerlies. During April 1983, zonal belts of high T_s, high W and low U stretch across the entire equatorial Pacific. Detailed evolution of these three parameters during the episode is described in Liu (1989). Despite the warm water, LE is below normal due mainly to the low U near the convergence center. The annual October high does not reach the expected level. The semi-annual frequency of SR is disrupted during the episode by a eastward migration of low value following the displacement of organized convection and cloud system. Although the change of SR is partly compensated by the the change in LE, the net flux (SR-LE) into the ocean is reduced. Monthly maps of the net flux are presented by Liu et al. (1989).

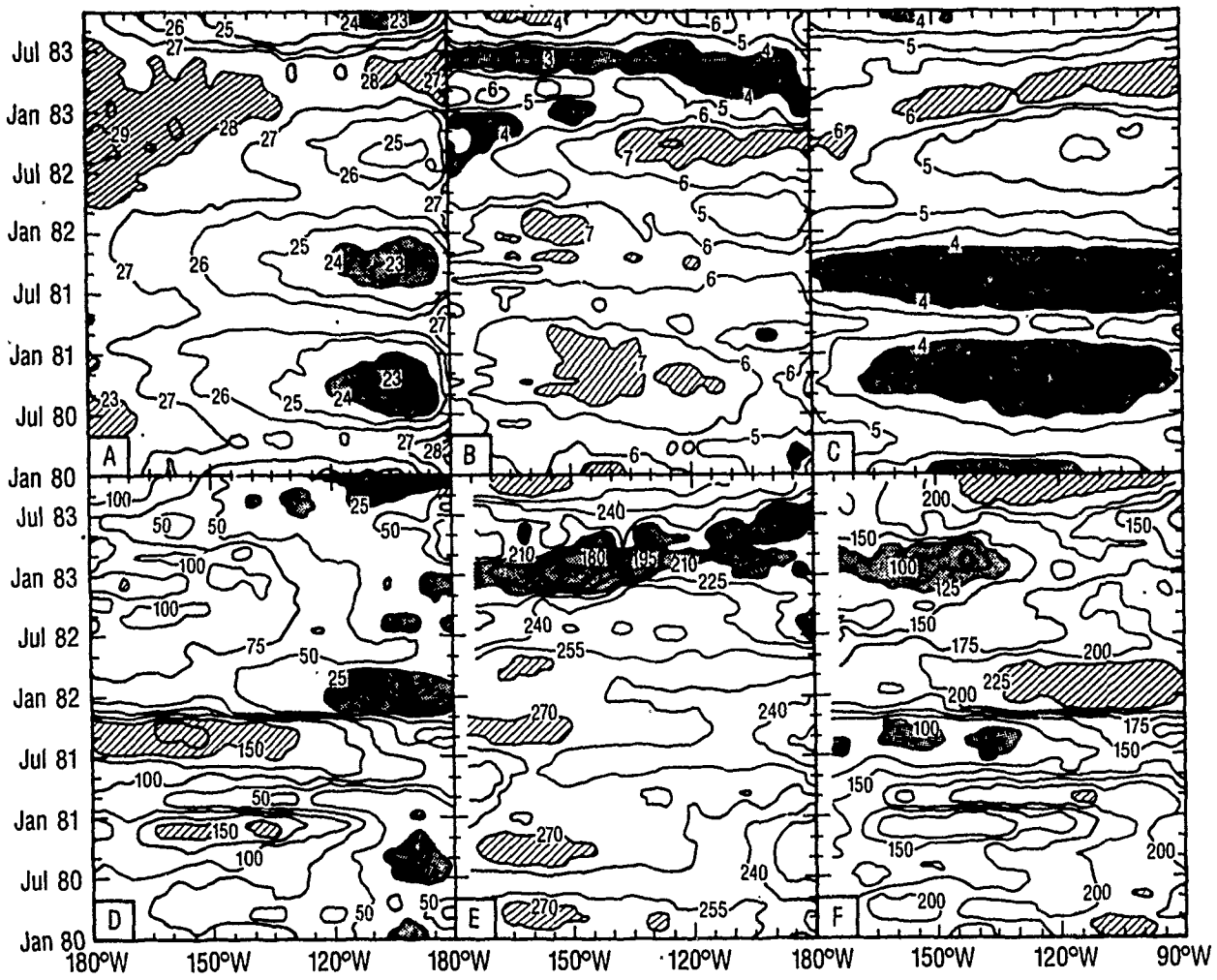


Fig. 1 Time-longitude variation, centered on the equator, of (A) sea surface temperature, (B) surface wind speed, (C) columnar water vapor, (D) latent heat flux, (E) shortwave radiation, and (F) the difference between shortwave radiation and latent heat flux. The intervals between isolines are 1°C, 1 m/s, 0.5 g/cm^2, 25 W/m^2, 15 W/m^2, and 25 W/m^2 respectively.

The change of heat storage in the upper ocean is governed by the balance heat gain from the surface and the loss through ocean dynamics. Fig. 2 shows the distribution of the lag correlation coefficient, at 2° latitude by 2° longitude grids, between the time series of the net flux and T_s. The time series of net flux starts from March 1982 and ends at August 1983. The flux leads T_s by two months. Away from the equatorial region, the correlation is significant, demonstrating that surface thermal forcing is an important factor for upper ocean heat balance and T_s change. In the equatorial wave guide, ocean dynamics plays a more dominant role and surface heat flux is a consequence rather than the cause of T_s change.

5 FUTURE POTENTIAL

The SSMI (Special Sensor Microwave Imager) launched in 1987 on the operational spacecraft of DMSP (Defense Meteorological Space Program) is an improvement in design over SMMR but it does not have the low frequency channels which are sensitive to sea surface temperature variation. Future computation of LE has to rely on Ts derived from operational infrared radiometer such as the AVHRR (Advanced Very High Resolution Radiometer). The methodology of computing LE can be extended to other tropical oceans. With access of GMS (Japanese Geostationary Satellite) data through ISCCP (International Cloud Climatology Project) and other means, computation of SR over the entire tropical Pacific is feasible.

ACKNOWLEGMENTS

The study of W.T. Liu was performed at the Jet Propulsion Laboratory, California Institute of Technology under contract with the National Aeronautics and Space Administration. The work of C. Gautier was performed at the California Space Institute and the Scripps Institute of Oceanography with the support of the National Oceanic and Atmospheric Administration. This is part of an interagency supported research program, the TOGA Heat Exchange Program.

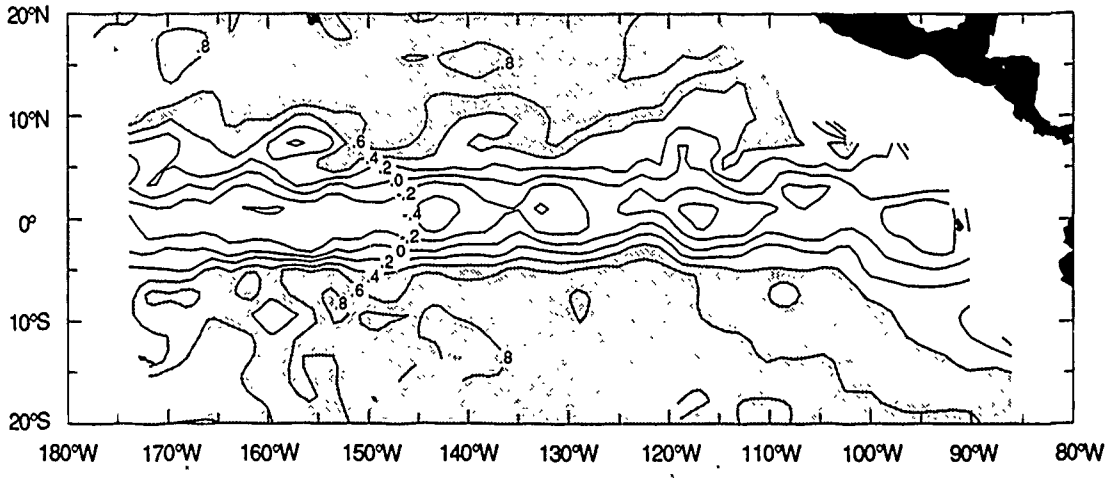


Fig. 2 Distribution of the lag correlation coefficients between ocean surface thermal forcing (shortwave radiation minus latent heat) and sea surface temperature. The forcing (2/82-8/83) leads sea surface temperature by two months.

REFERENCES

Gautier, C., "Satellite measurements of insolation over the tropical oceans", *Tropical Ocean-Atmosphere Newsletter, 12*, 5-6. 1982.

Gautier, C., G. Diak, and S. Mase, "A simple physical model to estimate incident solar radiation at the surface from GOES satellite data", *J. Appl. Meteor., 219*, 1005-1012, 1980.

Gautier, C. and K. B. Katsaros, "Insolation during STREX 1: comparisons between surface measurements and satellite estimates", *J. Geophys. Res., 89*, 11779-11788, 1984.

Liu, W.T., "Statistical relation between monthly precipitable water and surface-level humidity over global oceans", *Mon. Wea. Rev., 114*, 1591-1602, 1986.

Liu, W.T., "Moisture and latent heat flux variabilities in the tropical Pacific derived from satellite data", *J. Geophys. Res., 93*, 6749-6760, 6965-6968, 1988.

Liu, W.T., "The annual and interannual variabilities of precipitable water, surface wind speed, and sea surface temperature, over the tropical Pacific", *Ocean Air Interaction*, in press, 1989.

Liu, W.T., C. Gautier, and R. Tagett, "*1982-1983 El Nino Atlas - Heat Flux*", JPL Publication, in preparation, 1989.

Liu, W.T. and P.P. Niiler, *"Tropical Ocean and Global Atmosphere (TOGA) Heat Exchange Project - A summary report,"* JPL Publ. 85-49, Jet Propul. Lab., Pasadena, CA, 1985.

Physical Retrieval of Precipitation Cell Parameters Using Passive 118-GHz Observations

A.J. Gasiewski
School of Electrical Engineering
Georgia Institute of Technology, Atlanta, GA 30332-0250

D.H. Staelin
Research Laboratory of Electronics
Massachusetts Institute of Technology, Cambridge, MA 02139

Abstract- The physical retrieval of precipitation parameters from 118-GHz spectra is investigated using an iterative planar-stratified numerical radiative transfer model. Liquid and frozen hydrometeors are modelled as spherical Marshall-Palmer (MP) and Sekhon-Srivastava (SS) distributed Mie-scattering polydispersions (respectively) with Henyey-Greenstein phase functions. A comparison of 118-GHz rain cell spectral perturbations observed during COHMEX with model calculations based on coincident radar data yield ±10 percent agreement over the convective core region, although a mean ice size 50% larger than that given by the SS distribution is required for agreement over the anvil region.

A rain cell model parameterized by top-altitude and total water density suggests that the dominant 118-GHz spectral modes can be used to retrieve top-layer altitude with RMS errors of $\sim 1-1.5$ km, consistent with statistical retrieval results. Model calculations also suggest that cell density is not observable using 118-GHz spectra alone. However, cell density might be retrieved from coincident observations using similar clear-air channels at 53- and 118-GHz, although such a retrieval would be sensitive to the mean ice particle size.

The sensitivity of radiometric channels to temperature at various levels in the troposphere is exhibited, in the presence of precipitation, through the perturbed temperature weighting functions. Calculations using the perturbed and clear-air weighting functions suggest 118-GHz transparent-channel cell-top reflectivities of almost 60 percent for high-altitude convective cores. Precipitation cell probing depth is also revealed using the perturbed weighting functions. At 118-GHz, cells with water densities greater than 0.5 g/m^3 are found to be radiometrically opaque, while cells with less than 0.1 g/m^3 density are transparent.

Key Words- radiative transfer, scattering, precipitation, retrieval.

1. Introduction

The sensitivity of microwave brightness observations to precipitation is predominanalty due to absorption by liquid hydrometeors and to scattering of the cold cosmic background radiation by the ice canopy commonly found at altitudes above the freezing level. Radiometric observations near the 118-GHz O_2 line are characterized by temperature weighting functions that peak at successively higher altitudes, and thus are sensitive to the height of the precipitation, as well as to the strength of the scattering and absorption. In order to optimally utilize 118-GHz radiometric data in precipitation retrievals, the forward radiative transfer problem for a scattering and absorbing distribution of hydrometeors is solved using an iterative planar-stratified numerical model, appropriate for passive microwave O_2 observations.

The model has been verified by observations made using the Millimeter-wave Temperature Sounder (MTS) [1] aboard the NASA ER-2 high-altitude aircraft during COHMEX, 1986. The MTS is a Dicke-switched scanning spectrometer which senses atmospheric emission spectra using 8 double-sideband (DSB) channels located from 500 to 2000 MHz around the 118.75-GHz O_2 resonance, and one DSB channel in the 5-mm O_2 band centered at 53.65 GHz.

For the case of hydrometeor profile retrievals using single-spot, multichannel radiometric data, the most appropriate observables are the statistically significant modes in the brightness temperature spectrum. The physical retrieval of precipitation parameters using the 118-GHz observables is analyzed using the forward-transfer model in conjunction with a rain cell model parametrized by cell-top altitude and density. Consideration is given to the sensitivity of retrievals to instrument noise and model uncertainties, particularly in the ice particle size spectrum. Due to the frequency dependence of hydrometeor scattering and absorption, additional observable degrees of freedom can be obtained by inclusion of coincident 5 mm-band passive O_2 data.

2. Numerical Radiative Transfer Model

The radiative transfer model assumes that precipitation cells are composed of tenuous horizontally-stratified polydispersions of spherical hydrometeors situated over a specularly reflecting surface. By azimuthal symmetry, only the vertically and horizontally polarized Stokes' parameters are required to describe the brightness temperature field :

$$\overline{T}_B(h,\theta,\nu) = \begin{pmatrix} T_{Bv}(h,\theta,\nu) \\ T_{Bh}(h,\theta,\nu) \end{pmatrix} \tag{1}$$

which obeys the differential equation of radiative transfer (DRTE) [2] :

$$\cos(\theta)\frac{d\overline{T}_B(h,\theta)}{dh} = -\kappa_e \overline{T}_B(h,\theta) + \kappa_a \begin{pmatrix} T(h) \\ T(h) \end{pmatrix} + \kappa_s \int_0^{\pi} \overline{p}'(h,\theta,\theta') \cdot \overline{T}_B(h,\pi - \theta') \sin(\theta')\, d\theta' \tag{2}$$

where h is the height, θ is the zenith angle ν is the frequency, T is the temperature, κ_s, κ_a, and $\kappa_e = \kappa_s + \kappa_a$ are the total scattering, absorption, and extinction coefficients (respectively), and the Rayleigh-Jeans approximation is used. The Henyey-Greenstein (HG) phase function is used to model the hydrometeor scattering anisotropy. In Eq. 2, the scattering and absorption coefficients, κ_s and κ_a, are computed as the sum of contributions from both gaseous O_2 [3] and

H_2O [4] absorbers and Mie dielectric spheres, using the millimeter-wave dielectric constants of liquid water and ice [5,6]. Aggregate hydrometeor scattering, absorption, and phase function asymmetry are derived by integration over Marshall-Palmer (MP) liquid or Sekhon-Srivastave (SS) ice size distributions $N_D = N_0 e^{-\Lambda D}$, where N_D is the number density of particles with equivalent liquid diameters between D and $D + dD$.

The temperature weighting function set $\{\overline{W}(h',h,\theta), \overline{W}_S(h,\theta), \overline{W}_{CB}(h,\theta)\}$ defines the relationship between the brightness temperature vector $\overline{T}_B(h,\theta)$ and the atmospheric temperature profile $T(h)$, the surface temperature T_S, and the cosmic background temperature $T_{CB} = 2.7$ K :

$$\overline{T}_B(h,\theta) = \int_0^\infty T(h')\overline{W}(h',h,\theta)\,dh' + \overline{W}_S(h,\theta)T_S + \overline{W}_{CB}(h,\theta)T_{CB} \tag{3}$$

The iterative, or "perturbation" method was used to solve the DRTE for the temperature weighting functions and brightness temperatures. The series expansion for the temperature weighting function set becomes :

$$\begin{aligned} \overline{W}(h',h,\theta) &= \sum_{n=0}^{\infty} \overline{W}^{(n)}(h',h,\theta) \\ \overline{W}_S(h,\theta) &= \sum_{n=0}^{\infty} \overline{W}_S^{(n)}(h,\theta) \\ \overline{W}_{CB}(h,\theta) &= \sum_{n=0}^{\infty} \overline{W}_{CB}^{(n)}(h,\theta) \end{aligned} \tag{4}$$

An iterative procedure to compute the weighting function corrections is [7] :

$$\begin{aligned} \overline{W}^{(0)}(h',h,\theta) &= [1-\varpi(h)]\,\overline{W}^{(\epsilon)}(h',h,\theta) \\ \overline{W}_S^{(0)}(h,\theta) &= \overline{W}_S^{(\epsilon)}(h,\theta) \\ \overline{W}_{CB}^{(0)}(h,\theta) &= \overline{W}_{CB}^{(\epsilon)}(h,\theta) \end{aligned} \tag{5}$$

$$\overline{W}^{(n+1)}(h',h,\theta) =$$

$$\int_0^\infty \varpi(h'') \int_0^\pi \overline{p}'(h'',\theta,\theta') \cdot \overline{W}^{(n)}(h',h'',\pi-\theta')\sin(\theta')d\theta'\,\overline{W}^{(\epsilon)}(h'',h,\theta)\,dh''$$

$$\overline{W}_S^{(n+1)}(h,\theta) =$$

$$\int_0^\infty \varpi(h'') \int_0^\pi \overline{p}'(h'',\theta,\theta') \cdot \overline{W}_S^{(n)}(h'',\pi-\theta')\sin(\theta')d\theta'\,\overline{W}^{(\epsilon)}(h'',h,\theta)\,dh''$$

$$\overline{W}_{CB}^{(n+1)}(h,\theta) =$$

$$\int_0^\infty \varpi(h'') \int_0^\pi \overline{p}'(h'',\theta,\theta') \cdot \overline{W}_{CB}^{(n)}(h'',\pi-\theta')\sin(\theta')d\theta'\,\overline{W}^{(\epsilon)}(h'',h,\theta)\,dh'' \tag{6}$$

where $\varpi = \kappa_s/\kappa_\epsilon$ is the total single-scattering albedo and $\overline{W}^{(\epsilon)}$ is a non-scattering weighting function, but calculated using the extinction (rather than absorption) coefficient as the measure of differential opacity. The forward-transfer computation was implemented by trapezoidal quadrature over 40-50 height levels and Gaussian quadrature over 6-8 elevation angles, and typically required up to 95 iterations for convergence.

3. Experimental Verification of the Numerical Model

Computed brightnesses for a summertime mature convective cell couplet were compared with coincident 118-GHz brightness imagery made using the MTS. The couplet was observed ~ 100 km SE of Huntsville, AL by the CP-2 (10 cm, H-polarized) weather radar on July 11, 1986 during COHMEX; rain rates up to ~ 85 mm/hr were exhibited in the main cell core. A vertical reflectivity cross-section of the couplet coincided with the ER-2 flight track to within ~ 2 km, with worst-case temporal coincidence of 4 minutes.

Hydrometeor size distribution and phase profiles were derived from equivalent volume radar reflectivity (Z_e) measurements. Hydrometeor classification (liquid, ice, or mixed-phase) at each level depends upon the hydrometeor temperature. Above freezing, reflectivity measurements were assumed to correspond to rain, and the MP relations are used. For temperatures below that of ice nucleation (taken to be -35° C), measurements were assumed to be from ice, and the SS relations (corrected by 6.5 dB to account for the lower refrectivity of ice) were used. For temperatures between the freezing and ice nucleation extremes, the ice-to-liquid mixing ratio was assumed to increase linearly from 0% at 0° C to 100% at -35° C , with the total equivalant reflectivity fixed by Z_e. Temperature profiles were derived from a nearby radiosonde, relative humidity was modified to be 100% within regions of non-zero hydrometeor density, and surface reflectivity was assumed to be 5%.

Monochromatic nadiral 118-GHz brightness spectra were computed from 14 vertical reflectivity profiles and one hydrometeor-free profile used as a reference (Fig. 1. The brightness temperature perturbations (relative to the hydrometeor-free case) exhibit ±10% agreement with those observed in the convective regions (B and C) of the storm. However, the residual perturbation discrepancies in the anvil region (A) indicated significantly less scattering than suggested by the model. To resolve the discrepancy, the effect of decreasing the 118-GHz anvil ice scattering coefficient by a factor of 5 while holding the $\lambda = 10$ cm reflectivity constant was investigated. Indeed, Z_e remains unchanged if the estimated frozen hydrometeor density M is reduced by a factor of $(1.5)^3 = 3.375$, and the mean particle diameter $1/\Lambda$ is simultaneously increased by 1.5.

The computed transparent channel brightness temperatures using the modified anvil ice size relations agree with observations to within 10% of the channel perturbation amplitude. The modified ice density is also closer to the estimated liquid density below the anvil than is the original SS ice density. These results suggest that larger ice particle sizes than those given by the SS distribution may be present in anvil regions. An unusually large mean ice particle size could be due to accretion of smaller ice particles, which can be expected to be significant in view of the maturity of the anvil region (~ 30 – 60 minutes) and the moderate descent rates (~ 0.1 – 1.0 m/sec) of the ice particles. Both MP and Heymsfield-Platt size relations increase the 118-GHz ice scattering coefficient, hence were not used.

The computed 118-GHz perturbed weighting functions (Fig. 2) can be used to calculate both the probing depth of the 118-GHz channels, defined here as the distance into the cell over which 86% of the total cell emission has originated. Over the main cell core, the transparent MTS channel weighting function (120.68 GHz) exhibits some sensitivity in the liquid-phase regions located below ~ 9.5 km altitude, but direct physical retrieval of rain rate is improbable using the 118-GHz channels alone. Due to strong O_2 absorption, the most opaque MTS channel (119.80 GHz) responds only to the glaciated top of the main cell core. The cell-top reflectivity is the fraction of the perturbed weighting function that senses altitudes greater than the cell-top altitude due to scattering from the cell top. From the computed weighting functions, the nadiral cell-top reflectivity for the storm couplet approaches 50% for the transparent channel in the main cell core, and does not exceed 6% in the anvil region.

4. Physical Retrieval of Rain Cell Parameters from 118-GHz Spectra

Physical retrieval of precipitation parameters from the 118-GHz observables was investigated using a two-parameter cell model. The altitude-density (A-D) model assumes precipitation cells to be of constant total water (liquid and ice) density from a base altitude of 1 km up to the cell-top altitude. The hydrometeor phase is assumed to be liquid below the freezing altitude, frozen above the altitude of ice nucleation (-30° C), and linearly mixed in between. Nadiral 118-GHz weighting functions and corresponding brightness spectra were computed for a set of cell-top altitudes ranging from 2 to 16 km (2 km increments) and for total water densities from 0.01 to 10.0 g/m^3 (two values per decade).

Only the first three Karhunen-Loève (KL) modes in the 118-GHz rain cell brightness perturbation spectra exhibit amplitudes significantly greater than the sensitivity of practical earth-observing spectrometers. These principal component amplitudes (k_1, k_2, and k_3), constitute the radiometrically observable space, and are determined by the rank-ordered eigenvectors of the observed rain cell brightness perturbation covariance matrix [8]. The first two KL coefficients (k_1 and k_2) were subsequently computed from the A-D spectra (Fig 3). The range of the mapping from A-D to KL space agrees well with the observed (empirical) range of coefficients derived from an ensemble of 297 convective cell cores [8].

A radiometrically opaque cell exhibits a small transparent-channel probing depth relative to its top altitude. A-D model calculations suggest that for the most transparent MTS channel, precipitation cells are radiometrically opaque for densities greater than 0.5 g/m^3 , and transparent for densities less than 0.1 g/m^3 . The maximum cell-top reflectivity for the transparent channels reaches a maximum of nearly 60% for a cell top of 16 km, and somewhat less for opaque channels.

For purely liquid precipitation (i.e. tops below $\sim$ 6 km), the cell-top reflectivity near 118-GHz is only several percent, consistent with the low single-scattering albedo of liquid precipitation at microwave frequencies. The cell-top altitude for radiometrically opaque liquid cells (i.e., < 8 mm/hr rain rate, or $\sim$ 0.5 g/m^3 MP liquid density) can be retrieved over land using 118-GHz spectra, with an expected accuracy of less than a kilometer. For highly transparent cumulus clouds (< 0.01 g/m^3) the integrated water density would be retrievable; however, since k_1 and k_2 are not independent (implying only one observable degree of freedom), cell-top altitude and density cannot be unambiguously retrieved using 118-GHz spectra alone.

The transparent 118-GHz channels are always more strongly perturbed by glaciation than the opaque channels, regardless of the call-top altitude. The A-D model calculations suggest that the presence of a glaciated cell top is unambigously detectable by comparing k_1 with a threshold value of -20 to -25 K. For values of k_1 lower than this threshold, the presence of ice is ascertained. For values of k_1 greater than this threshold, the existance of high-altitude, radiometrically-transparent ice cover can be ascertained by testing for unusually high values of k_2.

The sensitivities of the coefficients k_1 and k_2 to cell-top altitude for ice densities greater than 0.5 g/m^3 provide justification for the statistical retrieval of the cel- top altitude. The separation of the "hooked" regions of the constant-altitude curves in Fig 3 suggest that the minimum attainable RMS error from 118-GHz physical retrievals of radiometrically-opaque cell-top altitude will be 1-1.5 km. Variations in the hydrometeor size distribution have little effect on the estimated altitude, although variations in the altitudes of freezing and ice nucleation will shift the A-D curves slightly, particularly for radiometrically thin ice canopies. The use of k_3 in altitude retrievals is not justified due to forward-transfer modelling uncertainties.

Contrary to the altitude retrieval capability of k_1 and k_2, all three coefficients k_1, k_2, and k_3 are relatively insensitive to density variations within the radiometrically opaque range (0.5 to 10.0 g/m^3). In addition, the forward-transfer relation presents an an ambiguity in density retrieval over this density range. Thus, physical retrieval of radiometrically-opaque density using 118-GHz spectra alone is not viable. Sensitivity of the two most dominant 118-GHz spectral modes to hydrometeor density is exhibited, however, for high-altitude (> 8 km) partially transparent (< 0.3 g/m^3) cloud cover. Such scenarios will exhibit unusually high values of k_2, and thus are expected to be discernable from opaque cumulonimbus clouds. However, 118-GHz observations will also be sensitive to the ice particle size distribution in this domain.

5. Meteorological Information from Coincident 60- and 118-GHz Observations

Nadiral observations of brightness temperature over convective precipitation cells during GALE and COHMEX using "similar weighting function" (SWF) O_2 channels within the 60-GHz (5-mm wavelength) and 118-GHz absorption bands have been used to investigate whether additional precipitation information can be obtained from such coincident multiband observations. In the context of passive sounding, SWF channels are defined to have nearly identical temperature weighting functions over a hydrometeor-free atmosphere. Owing to the frequency dependence of hydrometeor scattering and absorption, the response of SWF 60- and 118-GHz channels to liquid and frozen precipitation is markedly different. This can be seen in Fig. 4 for SWF 53.65- and 118.75 $\pm$ 1.47 GHz nadiral MTS observations during two passes over strong convection on June 29, 1986 during COHMEX. Constant altitude curves computed using the A-D model are overlayed.

Small negative perturbations ($\leq$35 K) in the 118-GHz brightness are 2-3 times as large as coincident SWF 53-GHz perturbations. Such frequency scaling is commonly seen in the MTS GALE and COHMEX observations over cumulus and weak cumulonimbus clouds (i.e., 118-GHz radiometrically transparent clouds, or < 0.1 g/m^3), and is due to the Rayleigh frequency dependence of scattering hydrometeors. The additional degree of freedom can be used by a corrector to to improve temperature soundings over radiometrically transparent, low-density (< 0.1 g/m^3) cloud cover.

Over heavier cumulonimbus clouds and convective precipitation, negative 118-GHz perturbations track the coincident 53-GHz perturbations more closely. This near-equality of the 53/118-GHz SWF perturbations is due to Mie scattering, which dominates for hydrometeor densities greater than 0.1 g/m^3 . Since the distribution-averaged scattering and absorption efficiencies are nearly constant within the Mie region, the scattering and absorption coefficients are nearly frequency independent from 53 to 118 GHz.

The curves of constant altitude and density in the SWF brightness plots form a curvilinear grid, particularly over the opaque range (0.5 to 10.0 g/m^3) and above 7 km in altitude. covering the region of 118-GHz density ambiguity. Thus, SWF O_2 channels will be useful for retrieval of radiometrically-opaque cell-top density. However, variations in the size distribution will shift the the curvilinear grid. By adopting the MP (rather than SS) ice size distribution, the perturbation ratios $\Delta T_{118}/\Delta T_{53}$ tend to increase for all cell top altitudes, consistent with the smaller mean size of MP-distributed particles.

6. Discussion

Many simplifying approximations have been made in this numerical model, such as the assumption of exponentially decaying

polydispersions of Mie spheres with Henyey-Greenstein scattering phase functions. A more complete analysis would assume a gamma distribution of sizes along with the presence of prolate and oblate ice scatterers, modeled by their analytic phase matrices. Consideration should also be given to hydrometeors with entrained air or snowflake habits, and a comparison of the HG phase function with the exact Mie phase matrix. Since the numerical model verification was limited to a single case study, comparisons for different types of precipitation as well as within other microwave bands are desireable. However, the mapping from A-D to KL space is not likely to be significantly altered.

Over precipitation for which the probing depth is small compared to the horizontal extent of the cell, the planar-stratified model is appropriate. However, where the probing depth is large relative to the size of the cell, the applicability of the planar-stratified model is questionable.

Acknowledgements-The authors would like to thank Drs. R.A. Adler, G.M. Heymsfield, M. Yeh, and T.T. Wilheit of the NASA Goddard Space Flight Center for helpful discussions, Dr. R.W. Spencer and R.E. Hood of the NASA Marshall Space Flight Center for the July 11 CP-2 radar data, and J.W. Barrett, P.G. Bonanni, and Dr. P.W. Rosenkranz of the Massachusetts Institute of Technology for assistance in the field experiment. This work was suported by NASA grant NAG 5-10.

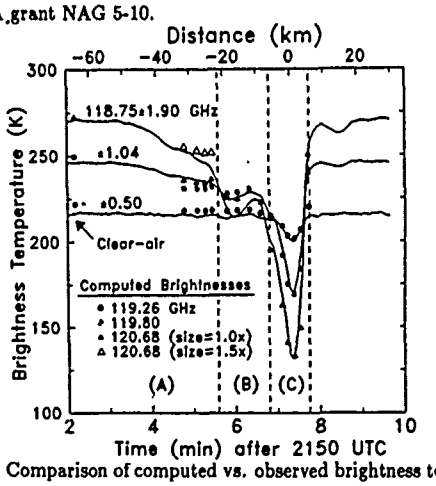


Figure 1: Comparison of computed vs. observed brightness temperatures a storm cell couplet on July 11, 1986 during COHMEX for the two extreme and one intermediate opacity MTS channel. Solid points are values computed using hydrometeor profiles derived from the SS relations, while the open points (anvil, region A) use a modified ice size distribution.

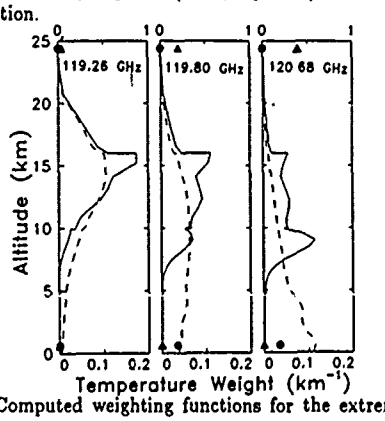


Figure 2: Computed weighting functions for the extreme and intermediate opacity MTS channels over the July 11 convective cell core, assuming MP liquid and SS frozen hydrometeor profiles. The dashed line is the hydrometeor-free weighting function. The symbols show the surface and cosmic background weights (bottom and top of graph, respectively) for the hydrometeor-free (●) and perturbed (▲) profiles.

References

[1] Gasiewski, A.J., J.W. Barrett, P.G. Bonanni and D.H. Staelin, Aircraft-Based Radiometric Imaging of Tropospheric Temperature Profiles and Precipitation Using the 118.75-GHz Oxygen Resonance, submitted to *J. Appl. Met.*

[2] Tsang, L., J.A. Kong and R.T. Shin, *Theory of Microwave Remote Sensing*, pp. 119-311, New York : John Wiley and Sons, 1985.

[3] Rosenkranz, P.W., Interference Coefficients for Overlapping Oxygen Lines in Air, *J. Quant. Spectrosc. Radiat. Transfer, 39*, 287, 1988.

[4] Liebe, H.J., An updated model for Millimeter Wave Propagation in Moist Air, *Radio Science, 20*, 1069, 1985.

[5] Ray, P.S., Broadband Complex Refractive Indices of Ice and Water, *Appl. Opt., 11*, 1836, 1972.

[6] Warren, S.G., Optical Constants of Ice from the Ultraviolet to the Microwave, *Appl. Opt., 23*, 1206, 1984.

[7] Gasiewski, A.J., Atmospheric Temperature Sounding and Precipitation Cell Parameter Estimation Using Passive 118-GHz O_2 Observations, Ph.D. thesis, Department of Electrical Engineering and Computer Science, Massachusetts Institute of Technology, Cambridge, MA, 1988.

[8] Gasiewski, A.J. and D.H. Staelin, Statistical Precipitation Cell Parameter Estimation Using Passive 118-GHz O_2 Observations, submitted to *J. Geo. Res.*.

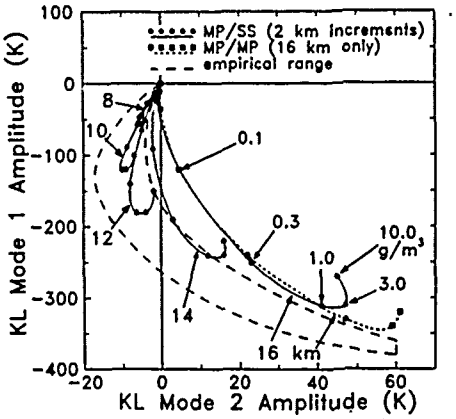


Figure 3: The mapping from altitude-density (A-D) space to the $k_1 - k_2$ (KL coefficient) subspace. Shown are constant altitude curves for 8, 10, 12, 14, and 16 km cell tops assuming a Sekhon-Srivastava ice size distribution (solid) and a Marshall-Palmer ice size distribution (dashed, 16 km only). The empirical range is derived from a set of 279 precipitation cell core observations.

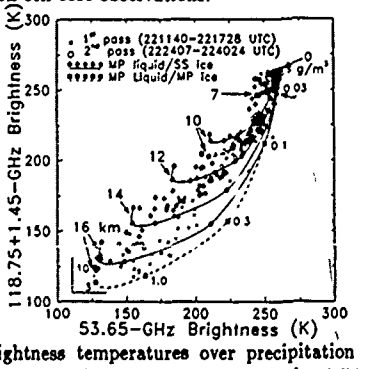


Figure 4: Brightness temperatures over precipitation for similar weighting-function O_2 channels at 53.65 GHz and 118.75±1.47 GHz. The points are MTS data observed during two passes over strong convection on June 29, 1986 during COHMEX. The solid lines are computed constant-altitude curves, parametrized by cell density. Computations used the A-D model with a MP liquid size distribution and a SS (solid) or MP (dashed) ice size distribution.

DETECTION AND ANALYSIS OF POLAR LOWS USING SATELLITE IMAGERY AND SOUNDER DATA

John Turner

British Antarctic Survey, Natural Environment Research Council,

High Cross, Madingley Road, Cambridge, CB3 0ED, UK

ABSTRACT

The intense mesoscale atmospheric vortices often referred to as polar lows are a major hazard to shipping and gas/oil exploration and operations in the polar regions. These systems develop very rapidly and frequently have surface winds in excess of 30 m s^{-1} as well as heavy precipitation, usually in the form of snow. Despite recent advances in numerical weather prediction, the skill in forecasting polar low developments is still poor and satellite data are the principal means by which such system are located. Satellite imagery has been used extensively for some years to detect polar lows, however, the manual analysis of imagery is a time consuming and imprecise process and the signs of developing systems can easily be missed. However, new forms of objective satellite data are becoming available to help in the detection of polar lows including atmospheric sounder profiles and surface winds from microwave scatterometers. The British Antarctic Survey is investigating the value of TIROS Operational Vertical Sounder (TOVS) data in polar meteorological research and has recently carried out a study into its use in investigating the structure and evolution of polar lows. One of the main problems in using TOVS data is that few validation studies have been carried out on the accuracy of the temperature retrievals. This is largely due to the fact that there is so little data available in the polar regions, especially on the mesoscale. However, the first instrumented aircraft flight through a polar low carried out by Shapiro et al (1986) has provided the most complete data set of the thermal structure of a polar low yet assembled. A comparison of the TOVS temperature profiles with the aircraft data will be presented and the horizontal resolution of the satellite retrievals will be discussed. It will be shown that the satellite data are capable of resolving many of the sub-synoptic-scale features of a polar low including the warm core found at the heart of many system. As an operational tool the TOVS thermal fields are able to identifying the strong low level baroclinic zones on which polar lows frequently form and to resolve the upper level cold core vortices which often provide the trigger for developments (Rasmussen, 1985). Examples of the possible role of TOVS data in operational detection and analysis of polar lows will also be presented.

1. INTRODUCTION

In recent years there has been a growing interest in polar meteorology due to operations concerned with the exploitation of natural resources at high latitudes and the realisation of the important role the polar regions play in the global climate system. Studies of the synoptic climatology of the polar regions (Streten and Troup, 1973; Rasmussen, 1983) have highlighted the wide range of atmospheric systems that occur on the synoptic and mesoscale, however, our knowledge of the mechanisms behind their formation and consequently our ability to model then numerically is still poor. Some of the most vigorous systems and therefore the most dangerous to polar operations are the mesoscale vortices usually referred to as polar lows. Although polar lows have been traditionally associated with the Northern Hemisphere, recent studies using satellite data have found similar systems around the coast of the Antarctic. This paper examines the role that satellites can play in detecting polar lows in their early stages and looks at how new forms of satellite data can be used in diagnostic studies.

2. THE NATURE OF MESOSCALE POLAR VORTICES

Mesoscale vortices develop and evolve over the high latitude ocean areas of both hemispheres. In the Northern Hemisphere they occur in a number of preferred areas, such as the Norwegian and Barents Seas and the Gulf of Alaska. These are the regions in which the 'classic' polar low developments occur and in which the greatest number of

case studies have been performed. Polar lows frequently occur in these areas because of the poleward transport of warm mid-latitude water by the ocean currents, such as the Gulf Stream. When off-ice atmospheric flow brings extremely cold Arctic air over the very warm water the air-sea temperature difference can be as large as 40° C which produces a very unstable vertical temperature structure. In these conditions polar lows occur with considerable cumulo-nimbus convection (Rasmussen, 1985). In the North Atlantic, over the sea areas South of Iceland, the nature of the polar lows is considerably different and the systems have more resemblance to small, mid-latitude cyclones. Here baroclinicity plays a greater role in the formation of the lows (Harrold and Browning, 1969) and they frequently have front-like structures on which the precipitation is organised. Systems forming in the Iceland area often track over the UK and give some of the heaviest snowfall to occur in the country. In the Southern Hemisphere the distribution of land masses is such that the poleward transport of warm water is much reduced compared to the Northern Hemisphere, so polar lows of the strongly convective type do not occur. The systems found in the Antarctic coastal region have many of the characteristics of the Northern Hemispheric, baroclinic type of polar low with the characteristic cloud spiral rather than the cluster of cumulo-nimbus towers. As in the Northern Hemisphere the fluxes of moisture and heat from the ocean are felt to be important as the vortices usually dissipate rapidly after making landfall.

3. DETECTION OF POLAR LOWS

With so few surface observations over the high latitude ocean areas satellites are crucial for locating the developing polar lows. In the very unstable conditions near the ice edge there is usually extensive cloud, often in the form of plumes and cloud streets and detecting small polar lows within this cloud is very difficult. In the northernmost areas, such as the Barents Sea, many deep cumulus and cumulo-nimbus clouds can also be present and any of these may be the precursor of a polar low development. Fig. 1 shows a mature polar low of this convective type in the Barents Sea at 04:11 GMT on 14 December 1982.

Upper level short wave troughs have been shown to act as the trigger for a number of polar low developments and the combination of satellite imagery and analysed fields of upper air data from numerical models may provide one means of determining which cloud clusters will develop. Fig. 2 shows the cloud associated with a developing polar low of the baroclinic type which formed south of Jan Mayen in response to synoptic-scale baroclinic forcing.

This case illustrates the difficulties of identifying the location of possible polar low developments. During the previous few hours three separate cloud features had been present in an East-West line to the South of Jan Mayen. All had the potential to develop into an active system but, in the hours preceding the spin-up of the system, it provided impossible to determine which cloud structure would develop. In the end it was the easternmost cloud which spawned the low and which then moved eastwards. However, this was clearly a very active region and further systems continued to develop during the following day.

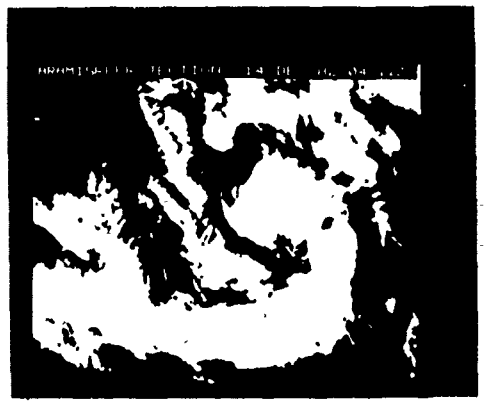


Fig 1 A polar low and deep cumulo-nimbus cloud in the Barents Sea at 04:11 GMT 14 December 1982.

Fig. 2 A developing polar low South of Jan Mayen at 08:25 GMT 27 February 1984.

As discovered in the late 1960s low level baroclinic zones close to the ice edge are often favourable areas for polar low developments. Although numerical models have improved considerable in recent years it is still difficult to pin point such areas due to their limited horizontal resolution. Satellite sounding data, however, can provide an indication of horizontal thermal gradients where the acquisition of upper air data is impossible by other means. As global temperature sounding data is available on an operational basis from the TIROS-N/NOAA series of polar orbiting satellites (Schwalb, 1978) the use of such data for

determining the possible locations of polar low developments could be carried out on a routine basis.

4. SATELLITE DATA AS A RESEARCH TOOL

Polar lows research has been gaining momentum since the 1960s, however, it is only in the last ten years that the full range of satellite data have been employed in these studies. During the 1970s high resolution satellite imagery was used to examine the cloud structures associated with polar lows and to determine cloud top temperatures. These data also allowed the production of the first climatologies of meso- and synoptic-scale disturbances in the polar regions (Streten and Troup, 1973; Businger, 1985). In the 1980s the first studies were carried out on the applications of sounder data within polar lows research (Steffenson and Rasmussen, 1986; Turner and Warren, 1988a). These indicated that despite the poor vertical resolution of the satellite temperature profiles the data could provide valuable information for use in diagnostic case studies when in-situ upper air temperature data were not available. The data were found to be particularly useful in the Antarctic coastal region where the temperature gradient between the very cold air masses over the interior of the continent and the relatively mild air over the ocean is large. Here the analysed fields of TOVS data, when used in conjunction with high resolution imagery, were found to be able to identify ice edge baroclinic regions and cold fronts descending from the Antarctic plateau (Turner and Warren, 1988b).

One of the difficulties in using sounder data is that few detailed validation studies have been carried to compare the satellite profiles against in-situ observations. The error characteristics of these data in the polar regions are therefore not well understood and we cannot easily determine the limits of horizontal resolution of the data. This is mainly due to the poor observing network at high latitudes and the lack of major data gathering campaigns. However, one of the projects of the international Arctic Cyclone Expedition involved the study of a polar low with an instrumented aircraft (Shapiro et al, 1987) and this provided the most detailed information yet on the thermal structure of a mature polar low. Passes of satellite sounding data have recently been validated against the aircraft observations to examine the capabilities of the sounding data. This particular polar low has been described in detail by Shapiro and only a brief description of the system will be provided here. The low developed South of Jan Mayen during the 27 February 1984 and was of the baroclinic type of system. It had surface winds in excess of 30 m s^{-1} and heavy precipitation on a mesoscale front extending Southwestwards from the centre of the system. At the heart of the low there was a warm core which coincided with the cloud free area at the centre of the cloud spiral. A number of passes of sounder data were processed covering the different stages of the polar low's development. Fig. 3 shows the analysed TOVS 1000-500 mb thicknesses at 08:20 GMT, when the low was still forming. The approximate position of the cloud associated with the system, as determined from the imagery, has been superimposed. This figure shows the strong baroclinic zone in which the low developed extending from Southwest to Northeast. There appears to be no thermal structure associated with the cloud vortex itself and the thermal gradient is fairly uniform across the system. Fig 4 shows similar data for the system at 13:40 GMT, which corresponds to the time when the aircraft made its observations within the low. The TOVS thicknesses at this time show far more thermal structure associated with the polar low and a break down of the previous Southwest-Northeast thermal gradient. In the centre of the system the satellite soundings have resolved the warm core and, although less intense than in the aircraft data, is correctly positioned in the centre of the cloud spiral. To the Southeast of the warm core is a pool of cold air which is believed to have been advected around the base of the low and this represents the leading edge of the cold arctic air being drawn into the system. Although the aircraft found a warm pool associated the front to the southwest of the centre there is no indication of this in the satellite thickness fields. In the aircraft data this was very shallow and on a scale comparable to the sounding measurements themselves.

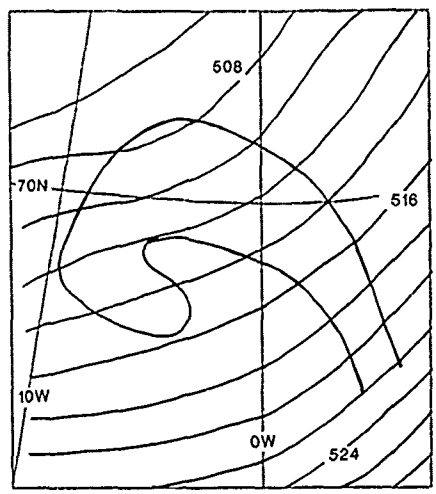


Fig 3. Analysed TOVS 1000-500 mb thickness values over a polar low at 08:20 GMT 27 February 1984.

One of the strengths of the satellite data is that passes are available every few hours in the polar regions so that thermal features can be followed during the lifetime of the mesoscale systems. In this case several passes of data were lost by the satellite operators during the late morning of the 27 February so a full analysis could not be performed. Nevertheless, with the data available it was possible to show that the warm core observed by the aircraft was a short lived feature which lasted only a few

hours. The sequence of data also support the idea proposed by Shapiro et al that the warm core represented the remains of warm mid-latitude air drawn into the heart of the system and trapped in a 'seclusion' process.

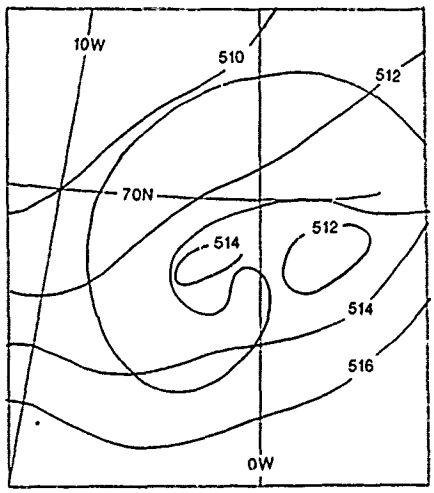


Fig 4. Analysed TOVS 1000-500 mb thickness values over a polar low at 13:40 GMT 27 February 1984.

5. NEW SATELLITE DATA

Over the next decade a number of polar orbiting satellite will be flying wind scatterometers which will provide data of value for polar lows research and operational forecasting. The European ERS-1 satellite will be launched in the early 1990s and provide global surface wind vectors at a resolution of 50 km. This is rather coarser than would be liked for studying the mesoscale wind field of vortices, nevertheless, it will certainly provide a useful tool for detecting the low level circulations associated with developing polar lows. Difficulties may be encountered due to ice contamination of the radar backscatter values and this will be a problem in certain areas of the Antarctic coastal region.

6. CONCLUSIONS

Satellite imagery is an important tool for the forecaster monitoring polar low developments at high latitudes. Many factors play a role in triggering the formation of these systems and a combination of imagery, satellite sounding profiles and model output is needed in this work. In research, further studies of the value of sounder data are required, however, the first case studies have indicated that these data can show many mesoscale features not resolved by the conventional observing network. A comparison of TOVS soundings against aircraft observations within a polar low shows good agreement on scales greater than 100 km.

REFERENCES

1. Shapiro, M. A., Fedor, L. S. and Hampel T "Research aircraft measurements of a polar low over the Norwegian Sea", Tellus, 39A, pp272-306, 1987.

2. Rasmussen, E., "A case study of a polar low development over the Barents Sea", Tellus, 37A, pp407-418, 1985.

3. Streten. N. A. and Troup. A. J., "A synoptic climatology of satellite observed cloud vortices over the Southern Hemisphere", Quart. J. R. Met. Soc., 99, pp56-72, 1973.

4. Rasmussen, E., "A review of meso-scale disturbances in cold air masses", In Mesoscale meteorology-theories, observations and models, ed. D. K. Lilly and T. Gal-Chen. D. Reidel Publishing Company, Boston, pp247-283, 1983.

5. Harrold, T. W. and Browning, K. A., "The polar low as a baroclinic disturbance", Quart. J. R. Met. Soc., 95, pp719-730, 1969.

6. Schwalb, A., "The TIROS-N/NOAA A-G satellite series", NOAA/NESS Tech. Memo. 95, 1978.

7. Businger, S., "The synoptic climatology of polar-low outbreaks", Tellus, 39A, pp419-432, 1985.

8. Steffensen, M. and Rasmussen, E., "An investigation of the use of TOVS data in polar low research", Polar lows project, Technical Report No. 25, Norwegion Meteorological Institute, 1986.

9. Turner, J. and Warren, D. E., "The structure of sub-synoptic-scale vortices in polar airstreams from AVHRR and TOVS data", Proceedings of the Second International Conference on Polar Meteorology and Oceanography. Madison, USA. March 1988.

10. Turner, J. and Warren, D. E., "High resolution temperature sounding of the polar atmosphere", Proc. of IGARSS '88. Edinburgh, 1988.

Statistical Retrieval of Precipitation Cell-Top Altitude Using Passive 118-GHz Observations

A.J. Gasiewski
School of Electrical Engineering
Georgia Institute of Technology, Atlanta, GA 30332-0250

D.H. Staelin
Research Laboratory of Electronics
Massachusetts Institute of Technology, Cambridge, MA 02139

Abstract- A non-linear statistical retrieval operator for precipitation cell-top altitude using high spatial-resolution passive 118-GHz O_2 brightness spectra is demonstrated. The retrieval operator consists of a Karhunen-Loève (KL) transformation followed by a rank reduction, a linearization, and a linear minimum mean-square-error estimator. Information from the 118-GHz data on the ambient atmospheric temperature profile and the precipitation cell size is also incorporated into the linear stage of the retrieval operator. The RMS retrieval error is 1.5 km for cumulus-stage cells with tops ranging from 1.5 to 16 km. The sensitivity of nadiral 118-GHz spectra to the cell-top altitude is predominantly due to the scattering and absorption of radiation originating from low, warm atmospheric levels by colder liquid and frozen precipitation. This effect causes cold perturbations in the brightness spectrum, which typically become stronger with increasing cell-top altitude. The different peaking-altitudes of the clear-air 118-GHz weighting functions provide additional "altitude-slicing" sensitivity.

Key Words- precipitation, retrieval, statistical, non-linear.

1. Introduction

Satellite-based observations of precipitation provide a means for monitoring the global hydrologic cycle over inaccessible regions, as well as giving advanced warning of inclement weather near populated regions. Precipitation cell-top altitude retrievals have previously been demonstrated using passive measurements of the infrared radiance emitted at the cloud top [1]. However, due to Mie scattering and absorption of infrared radiation by sub-millimeter sized hydrometeors, passive IR observations cannot directly probe beneath non-precipitating cloud canopies. At microwave frequencies such small hydrometeors behave instead as Rayleigh scatterers and absorbers, and thus cause significantly less extinction. This feature suggests the possibility of using passive microwave observations to directly probe the larger precipitating particles (i.e., those of radius $\gtrsim$200 μm) located beneath non-precipitating canopies.

The statistical retrieval of cell-top altitude has been investigated using high-resolution passive 118-GHz multichannel precipitation cell imagery. The images were produced by the Millimeter-wave Temperature Sounder (MTS) instrument [2] aboard the NASA ER-2 high-altitude aircraft during GALE (February, 1986) [3] and COHMEX (June-July, 1986) [4]. Presented here are the results of a non-linear statistical 118-GHz cell-top altitude retrieval technique, illustrating a systematic method of parameter estimation from non linearly dependent observables. The method is realized within a framework that allows straightforward incorporation of auxiliary information from both weakly dependent and strongly linearly dependent observables.

The embedding of cell-top altitude information in 118-GHz observations occurs via two mechanisms. First, a statistical dependence exists between the cell-top altitude and the brightness temperature of the cell top. In the case of cells with tops below the freezing level, higher cell-top altitudes are associated with increased precipitation, and hence increased absorption, which produces decreases in brightness for land backgrounds. In the case of cells extending above the freezing level increasing quantities of ice occur in the cell top. The presence of ice causes strong scattering of the cold cosmic background radiation, producing large negative perturbations in brightness temperature. Second, the altitude distribution of atmospheric water in the troposphere and lower stratosphere can be probed by virtue of the successively higher peaking altitudes of the 118-GHz clear-air temperature weighting functions [5]. The 118-GHz clear-air weighting functions peak at altitudes ranging from the surface, for transparent frequencies located ~2.5 GHz from the 118.750-GHz line center, to ~35 km (well above most clouds) for frequencies at the line center.

2. Precipitation Cell Brightness Spectra

In order to evaluate the precipitation cell parameter retrieval capability of 118-GHz observations, an ensemble of 279 independent near-nadiral spectra for precipitation cell cores was compiled, consisting of spectra from MTS observations during GALE and COHMEX. The observed precipitation cells were distributed throughout the southeastern United States, with most of the observations occurring during summer over the Huntsville, AL area.

The optical cell-top altitude a_i of each cell top was estimated by stereoscopy, using the MTS video images and the known altitude and speed of the aircraft. Generally, a_i is slightly higher than the corresponding microwave cell-top altitude due to coverage by optically-opaque cirrus shields which are nearly transparent to microwaves. A noise n_{ai} with an estimated RMS value of ~1 km is superimposed on a_i due to errors in the stereoscopic altitude estimation process. The size s_i of each cell was taken to be the distance along the flight track of the aircraft over which the MTS transparent channel brightness perturbation decreased to half its maximum value. The optical cell-top altitudes ranged from 2 to 16 km and the sizes ranged from 1.5 km (the resolution of the 118-GHz scanning spectrometer) to 200 km. In addition to cell-top altitude and

size, spectral observations were further classified according to one of two types, as determined primarily by the MTS video imagery. Observations of convective precipitation cells that appeared to be in their early stages of convection were designated to be of the cumulus (C) type. Observations of convective cells that exhibited anvils were considered to be mature or dissipating (M).

The spectrum for the i^{th} rain cell is denoted by a 8-vector of brightness temperature observations :

$$\overline{T}_{Bi} = \begin{pmatrix} T_{B(118.75\pm0.50)i} \\ T_{B(118.75\pm0.66)i} \\ \vdots \\ T_{B(118.75\pm1.90)i} \end{pmatrix} \quad (1)$$

where the component subscripts give the channel center frequencies (in GHz), arranged in order of decreasing opacity. Each rain cell spectrum is an average of several adjacent single-spot observations. The number of averaged spots (typically 4 to 8) for a single observation was limited by the size and structure of the rain cell core. The sampling region was typically between 5 and 10 km along the flight track Rain cell spectra were considered independent if they occurred over distinct cells.

For each rain cell spectrum, a corresponding clear-air reference spectrum $\hat{\overline{T}}_{Bi}$ was estimated from MTS observations in the vicinity of the cell. The perturbation brightness spectrum $\Delta\overline{T}_{Bi}$ for the i^{th} rain cell is :

$$\Delta\overline{T}_{Bi} = \overline{T}_{Bi} - \hat{\overline{T}}_{Bi} + \overline{n}_i \quad (2)$$

where $\overline{n}_i$ is a zero-mean Gaussian noise process, independent among channels and observations, and with standard deviation $\sigma_n = 0.5$ K for each channel.

3. Non-linear Statistical Retrieval of Precipitation Cell-Top Altitude

The retrieval of cell-top altitude from passive 118-GHz observations is performed by a non-linear statistical estimator on the perturbation spectra. The altitude estimator $\hat{a}_i$ consists of an orthogonal Karhunen-Loève (KL) transformation [6], followed by a rank reduction operation, a non-linear operator, and a linear-statistical estimator (Fig. 1).

The complexity reduction operation consists of a Karhunen-Loève transformation $\overline{\overline{E}}$, which rotates the 8-dimensional perturbation spectra space into the 8-dimensional KL basis, followed by a rank reduction operation $\overline{\overline{P}}$, in which only those KL components exhibiting a signal-to-noise ratio (SNR) significantly greater than unity are retained :

$$\overline{k}_i = \overline{\overline{P}}\,\overline{\overline{E}}\,\Delta\overline{T}_{Bi} \quad . \quad (3)$$

Using this procedure, only the statistically significant principal components of the perturbation spectra are used in the retrieval of cell-top height Noisy modes, which contain no geophysical information, are discarded.

The KL transformation $\overline{\overline{E}}$ is obtained by diagonalization of the 8 × 8 covariance matrix $\overline{\overline{R}}_{\Delta\overline{T}_B\Delta\overline{T}_B}$ of the perturbation spectra .

$$\overline{\overline{R}}_{\Delta\overline{T}_B\Delta\overline{T}_B} = \overline{\overline{E}}^T \begin{bmatrix} \lambda_1 & & & 0 \\ & \lambda_2 & & \\ & & \ddots & \\ 0 & & & \lambda_8 \end{bmatrix} \overline{\overline{E}} \quad (4)$$

where each eigenvalue λ_j is the variance of the j^{th} component of the decomposed spectra. The matrix $\overline{\overline{E}}$ is a row matrix consisting of the eigenvectors of $\overline{\overline{R}}_{\Delta\overline{T}_B\Delta\overline{T}_B}$. An estimate $\hat{\overline{\overline{R}}}_{\Delta\overline{T}_B\Delta\overline{T}_B}$ of the covariance matrix $\overline{\overline{R}}_{\Delta\overline{T}_B\Delta\overline{T}_B}$ was obtained from the cell core observation ensemble. The subset consists of $N = 197$ observations of cumulus (C) and mature or dissipating (M) cells over land with sizes s_i greater than 5 km :

$$\hat{\overline{\overline{R}}}_{\Delta\overline{T}_B\Delta\overline{T}_B} = \frac{\sum_{i=1}^{N} \Delta\overline{T}_{Bi} \left(\Delta\overline{T}_{Bi}\right)^T}{N-1} \quad . \quad (5)$$

All perturbation spectra channels were assumed to be zero mean; although this constraint is artificial, it does not affect the overall estimator significantly.

The rank-reducing projection operator $\overline{\overline{P}}$ passes only those coefficients with SNR = $(\lambda_j/\sigma_n^2) - 1$ greater than 5 ($\approx$ 7 dB), which is the criterion adopted in this analysis for statistical significance. Using this criterion, only the first three rank-ordered components k_1,k_2, and k_3 of the KL-transformed spectra are retained. The eigenvectors for these three modes (Fig. 2) filter primarily . (1) the overall (frequency-weighted) amplitude of the perturbation spectrum, (2) a linear combination of the slope of the perturbation spectrum near the 118.75±1.47 GHz channel and the amplitude of the perturbation spectrum in the 4 most opaque channels (118.75±0.50 to 1.26 GHz), and (3) the curvature of the perturbation spectrum (respectively). Standard deviations for the three components are 129 K, 10 K, and 1.4 K, respectively, corresponding to SNR's of 48.2, 26.1, and 8.3 dB. All other components displayed standard deviations commensurate with the instrument noise σ_n, and hence are discarded.

The two most dominant KL coefficients are subsequently linearized with respect to the optical cell-top altitude. The non-linear operator

$$\overline{l}(\overline{k}) = \begin{pmatrix} l_1(\overline{k}) \\ l_2(\overline{k}) \end{pmatrix} \quad (6)$$

was derived by fitting an appropriate curve to the scatter plot of the optical cell-top altitude estimate a_i vs. the KL coefficient k_{ji} for each channel j. The criterion used in constructing the non-linear mapping was that the scatter plots of a_i vs. l_{ji}, when interpreted as two-dimensional probability distributions p_{al}, for optical cell-top altitude and j^{th} linearized coefficient, were separable and maximally correlated.

The linearization mapping $l_1(\overline{k})$ is defined by spline curve fits to a series of selected points. Since the first KL coefficient k_1 follows a trajectory that is monotonic with respect to optical cell-top altitude, the component l_1 is a single-valued function of only k_1. The spline fit for k_1 displays monotonically increasing sensitivity with respect to height. This is explained by the onset of glaciation, which commences above the freezing level and is complete above the level of ice nucleation. The glaciated cell tops scatter microwave radiation from the cold cosmic background, which decreases the observed brightness temperatures to values well below the lowest atmospheric temperature The cell-top altitude retrieval capability of the 118-GHz channels is also exhibited in the k_2. However, unlike the first KL coefficient, the k_2 follows a trajectory that is multiple-valued in optical cell-top altitude To linearize k_2, it is necessary to determine the more appropriate of two possible spline-fit branches by hypothesis testing [6]. The coefficient k_3 exhibited no significant dependence on cell-top altitude, and was ignored.

The final stage in Fig. 1 consists of a linear minimum mean-square error (LMMSE) estimator $\overline{D} = \overline{R}_{a\tilde{l}} \overline{\overline{R}}_{\tilde{l}\tilde{l}}^{-1}$:

$$\hat{a}_i = \overline{D}\,\Delta\tilde{l}_i + \langle a \rangle \qquad (7)$$

where $\hat{a}_i$ is the 118-GHz cell-top altitude estimate for the i^{th} cell, $\Delta\tilde{l}_i = (\tilde{l}_i - \langle\tilde{l}\rangle)$, $\langle\tilde{l}\rangle$ is the mean linearized KL coefficient, $\langle a \rangle$ is the mean cell-top altitude for the ensemble, and $\overline{R}_{a\tilde{l}}$ and $\overline{\overline{R}}_{\tilde{l}\tilde{l}}$ are covariance matrices for the subscripted quantities.

The root-mean-square (RMS) discrepancy σ_e for the retrieval over the observation set consisting of cumulus, mature and dissipating cells is 1.97 km and the correlation coefficient between $\hat{a}_i$ and a_i is 0.86. A significant portion of σ_e is attributed to the RMS uncertainty of ~1 km in the optical cell-top altitude a_i. Some of the residual is attributed to the difference between the optical and the 118-GHz cell-top altitudes. Since the visible part of the spectrum is highly sensitive to thin cirrus cover, the estimated optical cell-top altitude is often higher than the retrieved 118-GHz cell-top altitude. This hypothesis was tested by performing the retrieval on a reduced perturbation spectra set consisting of only cumulus (C) convective cells observed over land. For this reduced set, the average deviation between optically estimated cell-top height and the height of the topmost radiometrically opaque precipitating layer is smaller than if mature or dissipating cells are included. The RMS error σ_e for the cumulus set was reduced to 1.6 km, and the correlation coefficient was increased to 0.91.

The retrieved cell-top altitude using the 118-GHz non-linear statistical method is plotted in Fig. 3 along with 5-, 10-, 20- and 30-dBZ radar reflectivity contours of a convective cell observed during COHMEX. Within the cell core (region C), the retrieved altitude follows the 20 dBZ reflectivity contour while in an adjacent smaller cell (region B), the retrieved altitude follows more closely the 5- to 10-dBZ contour. The deviation of the retrieved altitude from the 10-dBZ reflectivity contour is typically within 1 km.

4. Inclusion of Auxiliary Observations in the Retrieval of Cell-Top Altitude

Mechanisms contributing to the retrieval discrepancy include variations in the altitudes of freezing and ice nucleation. Since the 118-GHz clear-air reference spectra $\tilde{\overline{T}}_{Bi}$ contain information on the atmospheric temperature profile, they can be used to correct cell-top altitude retrievals for such variations. To perform this correction, $\tilde{\overline{T}}_B$ is incorporated into the LMMSE operator $\overline{D}$ as an additional input :

$$\hat{a}_i = \overline{D} \begin{pmatrix} \Delta\tilde{l}_i \\ \Delta\tilde{\overline{T}}_{Bi} \end{pmatrix} + \langle a \rangle \quad . \qquad (8)$$

where $\Delta\tilde{\overline{T}}_{Bi} = \tilde{\overline{T}}_{Bi} - \langle\tilde{\overline{T}}_B\rangle$ and $\overline{D}$ is computed from :

$$\overline{D} = \overline{R}_{a(\tilde{l},\tilde{\overline{T}}_B)} \overline{\overline{R}}^{-1}_{(\tilde{l},\tilde{\overline{T}}_B)(\tilde{l},\tilde{\overline{T}}_B)} \quad . \qquad (9)$$

Morphological precipitation cell features observable by a satelli based imaging 118-GHz sounder include cell size s, the logarithm of which is statistically dependent on cell-top altitude with a correlation coefficient of 0.83. Further improvements in the cell-top altitude retrieval resulted from incorporating $\ln(s)$ (where s is in km) into the LMMSE estimator :

$$\hat{a}_i = \overline{D} \begin{pmatrix} \Delta\tilde{l}_i \\ \Delta\ln(s_i) \end{pmatrix} + \langle a \rangle \qquad (10)$$

where $\Delta\ln(s_i) = \ln(s_i) - \langle\ln(s)\rangle$ and :

$$\overline{D} = \overline{R}_{a(\tilde{l},\ln(s))} \overline{\overline{R}}^{-1}_{(\tilde{l},\ln(s))(\tilde{l},\ln(s))} \quad . \qquad (11)$$

The reduction in the retrieval discrepancy σ_e by inclusion of various 118-GHz observations (both primary and auxiliary) is summarized in Table 1 for the cumulus-only cell set. The simultaneous use of all available 118-GHz information yields the lowest retrieval discrepancy ($\sigma_e = 1.5$ km).

5. Discussion

The 118-GHz retrievals of cell-top altitude are the first performed using microwave O_2 channels. Since the 118-GHz cell-top altitude is closely related to the altitude of the topmost layer of precipitation, it is a potentially more useful meteorological parameter for precipitation monitoring than cloud altitude estimates derived from optical or infrared measurements. Since the RMS error in the optical cell-top stereoscopic retrieval is estimated to be as large as ~ 1 km for cumulus cells, a lower bound on the cumulus cell-top altitude retrieval accuracy of as low as 1.1 km RMS for 118-GHz observations alone is suggested. Although the non-linear statistical estimator was constructed specifically for convective rain cell core observations over land, the techniques of complexity reduction, linearization, linear estimation, and incorporation of auxiliary observations may be applied to retrieval operators for cell observations over oceans, as well as for the retrieval of geophysical parameters in other areas of passive and active remote sensing.

Acknowledgements- The authors acknowledge the contributions of J.W. Barrett, P.G. Bonanni and Dr. P.W. Rosenkranz of the Massachusetts Institute of Technology. This work was supported by NASA grant NAG 5-10.

References

[1] Smith, W.L. and C.M.R. Platt, Comparison of Satellite Deduced Cloud Heights with Indications from Radiosonde and Ground-Based Laser Measurements, *J. Appl. Met., 17*, 1796, 1978.

[2] Gasiewski, A.J., J.W. Barrett, P.G. Bonanni and D.H. Staelin, Aircraft-Based Radiometric Imaging of Tropospheric Temperature and Precipitation Using the 118.75-GHz Oxygen Resonance, Submitted for publication to *J. Appl. Met.*, 1988.

[3] Dirks, R.A., J.P. Kuettner, and J.A. Moore, Genesis of Atlantic Lows Experiment (GALE) : An Overview, *Bull. Am. Met. Soc., 69*, 148, 1988.

[4] Williams, S.F., H.M. Goodman, K.R. Knupp, and J.E. Arnold, *SPACE/COHMEX Data Inventory Document*. NASA report TM-4006, U.S. Government Printing Office, 1987.

[5] Ali, A.D.S., P.W. Rosenkranz, and D.H. Staelin, Atmospheric Sounding Near 118-GHz, *J. Appl. Met., 19*, 1234, 1980.

[6] Van Trees, H.L., *Detection, Estimation and Modulation Theory, Part I*, John Wiley and Sons, New York, 1968.

[7] Kailath, T., *Lectures on Wiener and Kalman Filtering*, pp. 25-26, Springer-Verlag, New York, 1981.

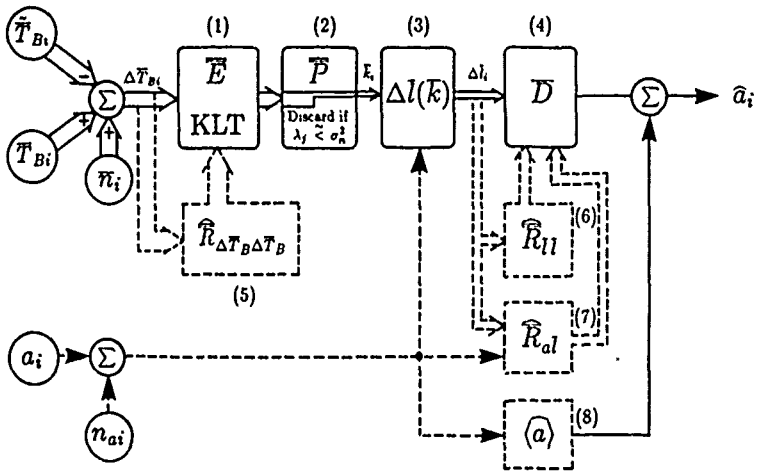


Figure 1: Block diagram illustrating the non-linear statistical retrieval method. (1) Karhunen-Loève transformation, (2) rank reduction, (3) non-linear mapping, (4) LMMSE estimator, (5) received signal covariance estimate, (6) linearized coefficient covariance estimate, (7) optical cell-top altitude and linearized coefficient cross-covariance estimate, (8) cell-top altitude mean estimate. The dashed lines represent off-line operations.

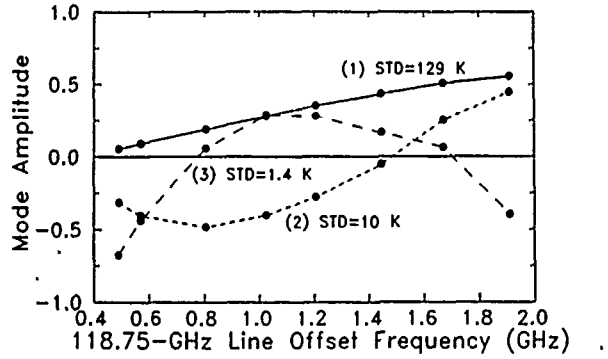


Figure 2: The three most significant components (along with standard deviations) in the MTS 118-GHz cell-core perturbation spectra.

Observations included	RMS discrepancy σ_ϵ (km)	Correlation coefficient
(*a-priori*)	3.85	-
b	3.65	0.324
c	2.16	0.828
b,c	2.04	0.848
a	1.63	0.906
a,c	1.61	0.909
a,b	1.53	0.917
a,b,c	1.50	0.921
118-GHz observations		
a	Linearized rain cell perturbation spectra	l
b	Clear-air reference spectra	$\tilde{T}_B$
c	Logarithm of cell size	$\ln(s)$

Table 1: RMS discrepancies σ_ϵ and correlation coefficients of retrieved cell-top altitude $\hat{a}_i$ vs. optical cell top altitude a_i for inclusion of various auxiliary 118-GHz observations in the LMMSE estimator.

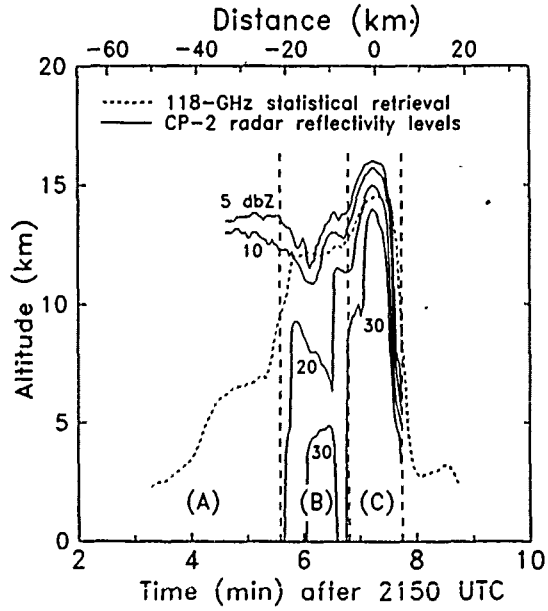


Figure 3. 118-GHz non-linear cell-top altitude retrieval compared to coincident radar observations of a convective precipitation cell couplet. The three regions (A), (B), and (C) denote the anvil, an adjacent cell, and the main cell core, respectively (COHMEX July 11, 1986, 2152 to 2200 UTC).

SENSITIVITY OF SATELLITE SIGNALS TO AEROSOL POLLUTANTS

O'Neill, N. T.[1], Royer, A.[1], Hubert, L.[1], J. Freemantle[2]

[1]CARTEL, Université de Sherbrooke, Sherbrooke Qué Canada
[2]CRESS, York University, 4700 Keele St., Toronto, Canada, M3J 1P3

Abstract

Radiative transfer simulations over a variety of environmental and operational parameters were performed in order to evaluate the sensitivity of satellite signals to the presence of polluting aerosols such as sulfates and carbonaceous soot. We present various examples of critical parameterizations which are essential towards an understanding of the feasibility of a sulfate inversion methodology.

Résumé

L'objectif de cette recherche est la détermination de la sensibilité des luminances satellitaires à la pollution des aérosols tels que les sulfates ou la suie. Nous avons effectué une serie de simulation théorique du transfert radiatif dans une large gamme de paramètres environnementaux et de contraintes opérationnelles. Des exemples de paramétérisation sont montrés et discutés. Ce types d'analyse est un préalable à la comprehension et au développement des méthodes d'inversion des teneurs en sulfate.

Introduction

The objective of the present study was to simulate the atmospherically scattered signal received by a satellite sensor and hence to evaluate the sensitivity of this signal to atmospheric aerosols in general and to atmospheric pollutants in particular. The emphasis of the investigations was on broadband scattering and continuum absorption effects (i.e. the level of information which can be extracted from typical remote sensing satellites) rather than the high spectral resolution features characteristic of structured molecular absorption phenomena.

The simulations were performed as a function of environmental, geometrical and operational parameter constraints. The environmental modelling included the effects of urban and rural type aerosols where the absolute and relative concentration of water soluble (sulfate) particles was allowed to vary and where particle growth and refractive index effects of relative humidity were incorporated. Combined aerosol modes which included fine particle (sulfate based), coarse particle (dust like), and absorptive aerosols (carbon based) were incorporated as external (independent) aerosol mixtures.

Numerous studies have shown that ground level scattering coefficients and measures of visibility are strongly correlated with the sulfate component over urban and rural sites (O'Neill et al., 1989). Other researchers have observed that scattering by carbonaceous products and ammonium nitrate can be important at least in urban areas (Sloane, 1983) while absorption by carbonaceous soot may significantly affect the light scattering budget in rural areas (Japar et al., 1986) and in particular over urban sites (Rosen et al., 1978).

These types of optical correlations have inspired a few reseachers in the remote sensing field to investigate the applicability of satellite optical imaging as a means of monitoring at least the most optically important atmospheric pollutants. The principal application of satellite detection methodologies to date has involved the detection of massive aerosol dust clouds where the optical effects are actually dominated by supermicron dust particles. In terms of the fine particle mode a number of researchers have attempted to extract indicators related to sulfate loading from satellite images (Lyons, 1980; Fraser and Kaufman, 1984). These investigations were primarily empirical in nature and sought to compare satellite derived estimates of sulfate content with a comparatively few ground based measurements. The simulation approach employed in this study permits a more flexible and fundamental understanding of the important variables in the error budget for extracting sulfate content.

General Approach

Aggregate aerosol models consisting of three externally mixed (independent) particle size modes (carbonaceous soot, water soluble and dust like) were used as standard inputs throughout the simulations. The hygroscopic component of these aggregate aerosol models was assumed to be primarily due to the narrow fine particle water soluble mode which itself was assumed to consist primarily of sulfate particles. The effects of relative humidity were then incorporated by applying a particle growth model to the sulfate component of the fine particle mode and computing the modified size distribution (O'Neill et al., 1989).

Four aggregate models which could be characterized as either urban (high carbonaceous soot concentration) or continental were defined . Standard and "variant" versions of both the urban and continental models were defined such that in the former case the sulfate component covaried with the other components while in the latter case the sulfate component varied independently of the other components.

Assuming spherical particles, a Mie scattering code was utilized to compute the scattering and absorbing parameters characterizing the aerosol components in a point volume. The component parameters were then optically averaged to obtain the aggregate aerosol optical (point volume) parameters and subsequently integrated over altitude assuming that the aggregate aerosol particle size distribution was itself independent of altitude. The columnar optical parameters so obtained were then entered into a radiative transfer model to simulate the signal received by a satellite sensor over a variety of operational and environmental parameters. These parameters included columnar sulfate concentration, sensor passband, surface reflectance, aggregate aerosol model, carbonaceous soot absorption and relative humidity.

Optical Parameter Variations

Figure 1 shows Mie based calculations of the variation of the *sulfate mass scattering efficiency* as a function of (dry) effective radius (O'Neill et al., 1989) for a number of different relative humidities. This parameter which describes the rate of increase of the aerosol scattering coefficient with the dry mass concentration of sulfates is the critical optical parameter for describing non cumulative optical effects of aerosols. In particular it is essential for discriminating sulfates from the other fine particle constituents. It's value depends on (i) the presence and covariance of other fine particle constituents (ii) relative humidity and to a lesser degree on (iii) the dimensions of the dry water soluble particles (c.f. Figure 1).

Specification of the sulfate mass scattering efficiency plus the integrated sulfate mass concentration is generally sufficient to quantitatively describe the variation of exo-atmospheric apparent reflectance. It's value must accordingly be measured or estimated for a given scene.

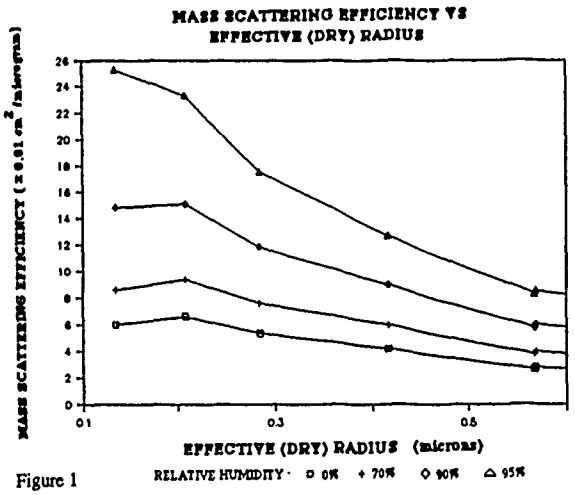


Figure 1

Radiative Transfer Model

The model employed for the radiative transfer calculations was the 5S model (Tanré et al., 1985). This model is a multiple scattering analytical solution to the radiative transfer problem posed in a plane parallel vertically homogeneous atmosphere. It's principal feature is an impressively rapid execution time suited to the data intensive requirements of remote sensing (Royer et al, 1988). Investigations into the optimal model which best combined the requirements of speed and accuracy indicated that the 5S model with certain qualifications was the best compromise for purposes of satellite data inversion. These investigations included an analysis into the effects of vertical inhomogenity in the atmosphere.

Variation of aggregate aerosol model

Figures 2a and 2b are plots of apparent reflectance versus equivalent sulfate surface concentration for the four number density models (urban standard and variant and continental standard and variant) surface reflectances of 0.0 and 0.5, an assumed scale height of 1 km and for two different relative humidities.

In the case of zero surface reflectance the dominance of sulfate scattering over the contributions of the other particles explains in part the lack of difference between the apparent reflectances. A second more subtle reason derives from the fact that the main difference between the models is the variation in the optical effects due to the soot component. The addition of soot particles (as occurs in changing from the continental standard to the urban standard atmosphere) causes a roughly equivalent increase in both the scattering and absorption volume coefficients of this component. The consequence of this is that the change in aggregate aerosol model is nearly transparent to a satellite measurement of apparent reflectance.

Such is not the case for a large surface reflectance where the attenuating properties of the atmosphere dominate the apparent reflectance measurement. Here the absorbtive and scattering increases due to the soot particles are actually combined since scattering in this context plays an attenuating role. The net effect is a decrease in apparent reflectance which is characteristically different for the larger soot concentration (urban model).

Ground albedo effects

The most notable effect of variations in surface reflectance is clearly the change in slope from postive to negative as the surface reflectance is increased (Figures 2 to 4). This phenomenon which has serious implications in terms of inversion algorithms for sulfate concentration is due to the competing influences of atmospheric attenuation and absorption. The reflectance for which the slope is actually zero (Chartier et al., 1989) is indeed a singularity at which the two influences balance (independently of variations in atmospheric optical thickness). For surface reflectances below this *critical surface reflectance* the satellite signal is dominated by scattering effects and accordingly increases with an increase in the number of scattering particles (sulfate particles). If the actual surface reflectance is greater than the critical surface reflectance then attenuation effects of surface reflected photons dominate and the satellite signal actually decreases with an increase in the number of scattering particles. An inversion scheme operating on pixels near the critical reflectance value is clearly fraught with instability.

Relative humidity effects

Referring to Figure 3, one can observe the significant effects of relative humidity on the computed value of apparent reflectance over a range of relative humidities. At a surface reflectance of zero (Figure 3a) the increase in apparent reflectance with increasing sulfate is most significant when the relative humidity is high and the scattering cross sections of a given concentration of sulfate particles are maximum. At large reflectances above the critical reflectance (Figure 3b) the effects of attenuation dominate and the larger cross sections associated with higher relative humidities effect a decrease in apparent reflectance as the relative humidity increases. At values of reflectance near the critical reflectance one obtains a flat or nearly flat response curve for apparent reflectance versus sulfate concentration.

Dependence on satellite bands

A more relevant description of the apparent reflectance variation seen in Figure 3 is in terms of signal to noise ratios (Table 1). The apparent reflectance variations (90% relative humidity, surface reflectance = 0) between 0 and 50 $\mu g/m^3$ were integrated over the passbands of a number of satellite sensors and normalized to the respective noise plus radiometric resolution figures for each sensor.

Table 1 Apparent reflectance variation corresponding to an SO4 variation from 0 to 50 ug.*m-3 (normalized to the total satellite sensor noise error)**

Satellite	Sensor	Band	Band width (micrometers)	Bits/ pixel	Signal/Noise
Landsat -5	TM	1	.45 - 52	8	4,8
		2	.52 - .60	8	3,2
		3	63 - 69	8	3,0
		4	76 - 90	8	4,2
		5	1 55 - 1.75	8	0,4
		7	2 08 - 2 35	8	0,1
	MSS	1	5 - 6	7	3,1
		2	6 - 7	7	2,7
		3	7 - 8	7	2,2
		4	8 - 1 1	7	1,9
SPOT	HRV	XS1	50 - .59	8	5,0
		XS2	61 - 69	8	
		XS3	79 - .90	8	3,9
NOAA	AVHRR	1	58 - 68	10	10,3
	(9,10,11)	2	715 - .982	10	6,4
METEOSAT			4 - 1 1	6	6,7
GOES	VAS/VISSER		.54 - 75	6	10,3
Landsat-6	SEAWIFS	1	443 - 453	10	58,8
		2	490 - 510	10	50,3
		3	555 - 575	10	56,0
		4	655 - 675	10	69,2
		5	745 - .785	10	41,7
		6	843 - 887	10	107,0

It is clear that, for small surface reflectance, only those sensors with 10 bit resolution (or for which the dynamic range straddles the low reflectance range characteristic of atmospheric phenomena) can yield sufficient information on the signal variations due to sulfates. Of particular note are the (nominal) S/N features for the SEAWIFs sensor.

Viewing Geometry

Figure 4 shows the variation of the apparent reflectance as a function of sulfate concentration for different values of viewing zenith angle. The left hand side of each graph corresponds to look directions on the anti solar side (observers back

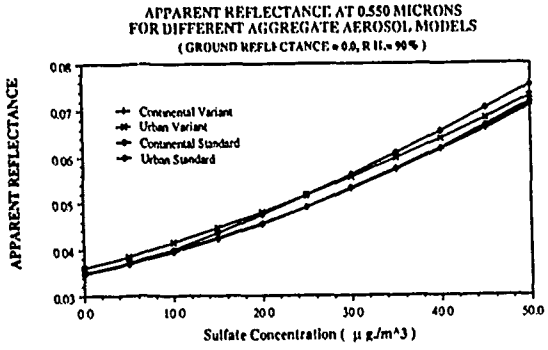


Fig. 2a

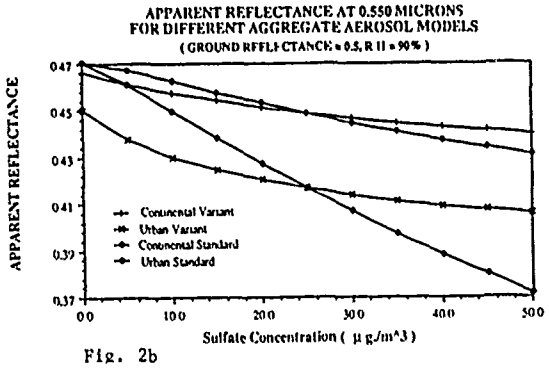


Fig. 2b

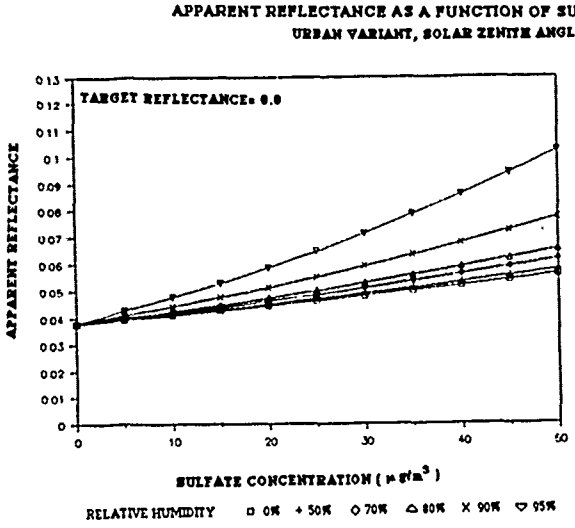


Fig. 3a

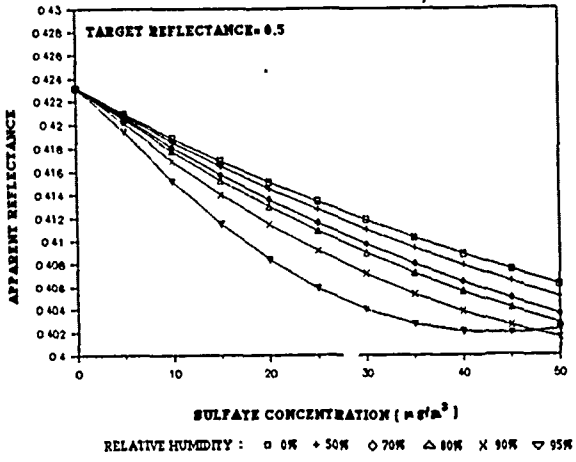


Fig. 3b

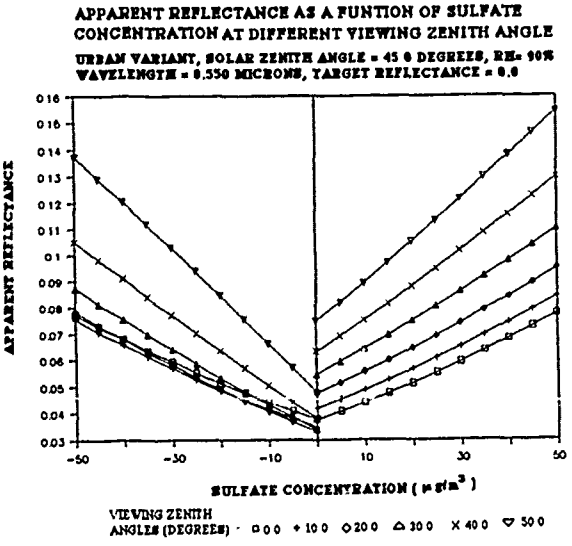


Fig. 4a

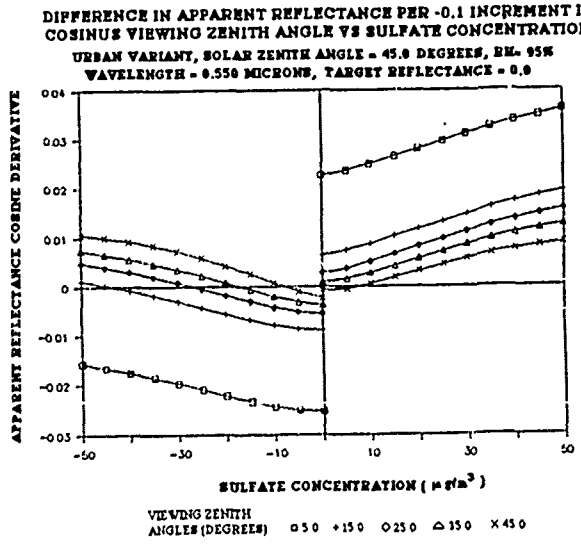


Fig. 4b

to the sun) while the right hand side of each graph represents the solar side.

For zero surface albedo the apparent reflectance dependence on viewer geometry is predominantly a function of viewing zenith angle and can be taken as being ~ 1/cosθ Second order effects are induced by the angular form of the scattering phase function.

At large surface reflectance the apparent reflectance response curves are somewhat complicated by the competing effects of scattering and attenuation. The non linearities observable in Figure 3b are in fact further exaggerated at large zenith angles where scattering effects become increasingly more important. Figure 4b shows that in the case of large surface reflectance a more promising variation to consider is the derivative of the apparent reflectance with respect to the cosine of the observer's zenith angle. It should be noted however that the reasonably monotonic increase of apparent reflectance with increasing sulfate concentration is not large and accordingly implies a fine signal sensitivity requirement. As well, the underlying assumption of diffuse Lambertian reflection which was employed in the simulations clearly has some impact on the results.

Sulfate/Dust Discrimination

One of the most critical elements in the development of an inversion methodology specific to sulfates is clearly how to distinguish the radiative transfer effects of sulfate from those of other aerosol contributors. The most readily accomplished discrimination is effected between the dust like and sulfate components. Figure 5 shows some calculations of apparent reflectance as a function of wavelength for a fixed total aerosol content (defined in terms of the total aerosol optical depth at 0.55 μm). In order to better isolate the relative influence of dust and sulfate particles the separate curves have been calculated for pure samples of each type of particle.

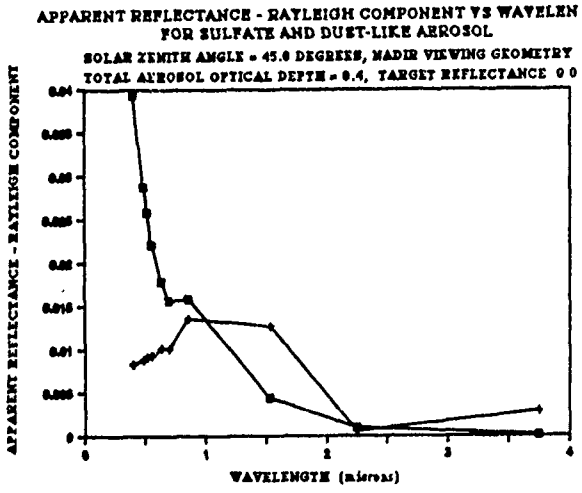


Figure 5 ◻ Sulfate + Dust Like

As the wavelength increases from the visible spectral region to the near IR one can characterize a transition of sulfate particles from optically small ($1 < 2\pi r_{eff}/\lambda < 10$) to optically very small ($2\pi r_{eff}/\lambda < 1$) and dust like particles from optically large ($2\pi r_{eff}/\lambda > 10$) to optically small. The net effect of these shifts in effective optical size is to produce a dominance of sulfate scattering in the visible and a dominance of dust scattering in the near IR.

This example illustrates well that spectral data can be employed to discriminate the difference between particle modes. It is however difficult to envisage a spectral inversion scheme which could successfully separate the effects of fine particle scattering constituents such as nitrates and sulfates. In this case one must rely on independently determined values of sulfate mass scattering efficiency to separate the vertically integrated mass concentration effects of other fine particle mode constituents.

Summary and Conclusions

The variability of the simulated signals indicate that information on the vertically integrated contribution of submicron scattering and absorbing aerosols can be usefully extracted. Inasfar as one can assume invariant aerosol scale heights this extracted information can be correlated with ground level concentrations.

Fine mode particles can be distinguished from coarse particles by exploiting the spectral dependence of the apparent reflectance. Sulfate discrimination within the fine particle mode is more a function of the precision with which the sulfate mass scattering efficiency can be measured or computed.

Inversion algorithms should be restricted to reflectances well below and above the Kaufman critical surface reflectance. For the aggregate aerosol models employed, we found that the satellite signal was fairly insensitive to the presence of carbonaceous soot over small surface reflectances (scattering atmosphere) but conversely was strongly dependent over high surface reflectances. This suggests that an inversion methodology would be best served by formulating the sulfate inversion around low reflectance pixels in the image and the soot inversion around high reflectance pixels.

The extraction of mass loading parameters from the total signal received by the satellite sensor implies some apriori knowledge of the ground surface reflectance in the satellite image. To a certain degree this requirement can be circumvented by restricting measurements to spectral regions and scene surfaces whose reflectance is small. However to fully exploit the potential of the satellite data requires the establishment of reflectance calibration sites and the employment of multi-angle and multi-temporal techniques.

The Landsat SEAWIFS sensor, due to be launched in 1991, offers a spectral resolution, repetition rate (2 days), an image overlap capability and a radiometric resolution which more closely approximate the requirements for passive atmospheric remote sensing. Within the near future the HIRIS and MODIS sensors possibly coupled with vertical structure data from the laser sounder (LASA) represent the realization of a sensor package for which one of the major design constraints was the remote sensing of atmospheric constituents. The latter sensors will be part of the EOS space station package to be launched in 1995.

This project was supported by the Ontario Ministry of the Environment (Project No. 349G) and NSERC (grant No.s A8643 and A1765).

References

1. Chartier, L., O'Neill, N. T., A. Royer, (1989), Analyse des aérosols à partir des images AVHRR-NOAA 9 et 10. Application à la cartographie des sulfates, IGARSS'89, Vancouver, B.C.
2. Fraser R. S. , Kaufman, Y. J., (1984), Satellite Measurements of Aerosol Mass and Transport, Atmospheric Environment, 18, pp. 2577-2584.
3. Japar, S. M., et al., (1986), The Contribution of Elemental Carbon to the Optical Properties of Rural Atmospheric Aerosols, Atmospheric Environment, 20, pp. 1281-1289.
4. Lyons, W. A., (1980), Evidence of Tansport of Hazy Air Masses From Satellite Imagery, Annals of New York Academy of Sciences, pp. 418-433.
5. O'Neill, N. T., et al., Critical Evaluation of Atmospheric Pollutant Parameterization from Satellite Imagery, Ministry of the Environment of Ontario, Internal Report (project No. 349G).
6. Rosen, et al., (1978), Identification of the Optically Absorbing Component in Urban Aerosols, Applied Optics, 17, 3859 3861.
7. Royer, A , et al., (1988), Comparison of Radiative Transfer Models Used to Determine Atmospheric Optical Parameters from Space, SPIE, 928, Critical Reviews of Optical Science and Technology, Modelling of the Atmosphere.
8. Sloane, C S (1983), Optical Properties of Aerosols - Comparison of Measurements with Model calculations, Atmospheric Environment, 17, pp. 409-416.
9. Tanré, D , et al., (1985), Simulation of the Satellite Signal in the Solar Spectrum (5S), Proc. of the 3rd Int. Col. on Spectral Signatures of Objects in Rem. Sens., Les Arcs, France, ESA SP-247, 315-319.

Implications from LIMEX/LEWEX for Spaceborne SAR Sensing of Ocean Wave Spectra

R. Keith Raney

Canada Centre for Remote Sensing
and RADARSAT
2464 Sheffield Road
Ottawa, Ontario K1A 0Y7

Paris W. Vachon

Data Acquisition Division
Canada Centre for Remote Sensing
2161 Sheffield Road
Ottawa, Ontario K1A 0Y7

Abstract

Based on analysis of SAR data and on theoretical predictions, observations are presented regarding improved performance from airborne and spaceborne SARs when observing ocean waves. The data set includes both open water waves and waves moving under a floating ice field. The theory summarized includes both known and new results.

Key Words: SAR, Ocean Wave, Spectrum

1 Introduction

LIMEX (Labrador Ice Margin Experiment) and LEWEX (Labrador Extreme Waves Experiment), occurring simultaneously in March, 1987, offered the opportunity of using airborne SAR (synthetic aperture radar) to image "known" surface gravity waves both in the open ocean and penetrating into floating ice (Vachon *et al.*, 1988; Raney *et al.*, 1989). Having two distinctly different reflecting media, ice and water, subjected to the same gross wave dynamics, presented unique experimental conditions for observing and logically separating key parametric aspects of the SAR imaging process.

Quantitative use of SAR systems from space for estimation of ocean wave directional spectra depends on two principal considerations: *radiometric response* leading towards calibration with respect to wave height; and *useful fidelity in the azimuth direction* (along track) of the SAR-derived spectral components. Using experimental airborne SAR data from LIMEX/LEWEX, together with analysis, progress has been made in both areas.

The role of *time dependence* on the azimuth response in SAR wave observation is central to wave spectral estimation. Two classes of motion need to be considered: radar sensitive *coherent* motion of the sea (orbital velocity), leading to velocity bunching (spatial) non-linearities and spectral cut-off (coherence time limitation); and *non-coherent* motion (phase velocity), the cause of "focus" dependence of SAR wave imagery and the basis for spectral ambiguity removal. We suggest *system considerations* concerning airborne and spaceborne SAR design intended to provide better estimates of ocean wave spectra.

2 Radiometric Response

If correctly calibrated, a SAR might be considered an "imaging scatterometer". A SAR system, including the image processor and spectral analysis stages, preserves calibratable estimates of mean reflectivity (Raney, 1983; Raney and Vachon, 1988a) if and only if the power norm is used in the image domain and the spectral transform is simply a Fourier transform. If this approach is taken, the mean reflectivity is proportional to the "DC digital number" in the transform domain. (It should be noted that "conventional" SAR processors use the *square root* of image power, and that "conventional" spectral analysis uses *the square of the Fourier transform magnitude*, which are not energy conserving, even in combination.) Having the mean reflectivity, the wave (image) contrast is quantitatively expressed in the directional spectral domain as variance as a function of (vector) wavenumber, when normalized by the DC digital number. Under the assumption that the azimuth component of the wave image spectrum is preserved, this approach has been verified using LIMEX/LEWEX data (Raney *et al.*, 1989).

The remaining radiometric issue concerns relating the reflectivity variance to useful physical properties such as wave height. This requires understanding and appropriate modelling of the four principal conditions governing wave reflectivity: hydrodynamic modulation; tilt modulation; velocity bunching; and coherence time limitations. Finding a practical relationship between a directional *image* spectrum, and the desired directional *wave height (or slope)* spectrum, at least in our view, requires additional progress.

The principal difference between waves on the ocean and the same waves that have penetrated a field of floating ice is that in the ice, the longer wavelengths survive, whereas the shorter waves (from capillaries to waves of length comparable to about one half of the average floe size) are quickly suppressed. From a SAR point of view, this means that in the open sea all four reflectivity conditions apply, whereas in the ice field the velocity bunching mechanism predominates. Experimentally, this means that it should be possible to estimate the relative importance of the various scattering mechanisms through comparison of related ocean and ice wave image properties. This is the subject of ongoing work.

3 Time Dependency

There are two aspects to the time dependency of the azimuth response of a SAR when observing a wave field: coherent and non-coherent. The key paramters introduced in this section are summarized in Table 1.

3.1 Coherent

It is well known that the SAR coherent (single-look) integration time is

$$T_L = \frac{R\lambda}{2\rho_A V_{\mathrm{TAN}}} \quad \text{(seconds)} \tag{1}$$

dependent on radar wavelength λ, single-look azimuth resolution ρ_A, and the component of vehicle velocity parallel to the earth's surface V_{TAN}. The integration time is proportional to

$$T_R = \frac{R}{V_{\mathrm{TAN}}} \quad \text{(seconds)} \tag{2}$$

which we denote as the (range dependent) radar *integration time constant*. This time constant is appropriate for system considerations that are time dependent in proportion to radar slant range. In particular, when the *scene coherence time* τ is less than T_L, reduced azimuth wave response results (Raney 1983), characterized as an "azimuth cut- off" (*e.g.* Monaldo and Lyzenga, 1986). This is often expressed in terms of the "R over V ratio". If the minimum detectable ocean wavelength is taken as two azimuth resolution cells, Monaldo and Lyzenga's minimum wavelength expression may be used to find

$$\tau \approx \frac{\lambda}{2\sqrt{H_S}} \quad \text{(seconds)} \tag{3}$$

showing τ to be proportional to radar wavelength and inversely proportional to the square root of significant wave height.

It is also well known that "velocity bunching" plays a major role in the SAR azimuth wave image formation process. Whereas most analyses of velocity bunching are expressed in terms of the "R over V" ratio, we feel that this is not appropriate. Velocity bunching, although arising from coherent properties of the Doppler domain, is a purely geometric effect in the sense that its impact is not proportional to integration time. We much prefer that the geometric image properties be expressed in terms of the *SAR dwell constant*

$$T_H = \frac{H}{V_{\mathrm{TAN}}} \quad \text{(seconds)} \tag{4}$$

where H is the sensor altitude above nadir. Variations in velocity bunching naturally occur as a function of incidence angle and wave aspect, and may be expressed with appropriate trigonometric weighting factors (Raney and Vachon, 1988a).

Whereas the distinction between the "SAR dwell constant" and the "integration time constant" may appear to be frivolous, we observe that they play different roles in the SAR wave imaging problem. Using LIMEX wave in ice data (Raney *et al.*, 1989), measurements show that wave image contrast increases with increasing incidence angle, consistent with velocity bunching predictions in the absence of coherence time limitations. When coherence time is included in the model (Raney, 1980), an additional coherence factor of

$$\frac{1}{\sqrt{1 + (T_L/\tau)^2}} \tag{5}$$

Table 1: SAR/Wave performance parameters.

SAR dwell constant	$T_H = H/V_{\mathrm{TAN}}$	seconds
One look integration time	$T_L = R\lambda/2\rho_A V_{\mathrm{TAN}}$	seconds
Integration time constant	$T_R = R/V_{\mathrm{TAN}}$	seconds
Coherence time	$\tau = \lambda/2\sqrt{H_S}$	seconds
Exposure time	$T_N = T_R \sin(\beta)$	seconds
Relative phase velocity	$dC_O = C_O/V_{\mathrm{TAN}} \cos(\phi)$	
Scale perturbation	$S = (1 - dC_O)$	
Coherence factor	$1/\sqrt{1 + (T_L/\tau)^2}$	

appears (see Eq. 1 Raney and Vachon, 1988a) in the amplitude weighting term, thus suppressing the onset of non-linearities as the integration time exceeds scene coherence time. When the T_L/τ term is large, it dominates the coherence factor which then is nearly proportional to $(R/V_{\mathrm{TAN}})^{-1}$. In this case, it tends to offset the H/V_{TAN} which is present due to velocity bunching, but not exactly. They are not inter-changeable. The lack of time coherence compensates to some extent for the geometric velocity bunching non-linearities. We conclude that in order to extend the linear dynamic range of a SAR observing waves, "some" scene coherence time limitation is beneficial from the point of view of SAR wave image fidelity.

An experiment could be designed to take advantage of certain geometric properties of velocity bunching. Since the spatial scale of the azimuth wavelength is dependent on the sign of the sensor velocity relative to the azimuth wavenumber, the nonlinearity resulting from velocity bunching is likewise affected. (This, as predicted, has been observed in LIMEX/LEWEX data.) By gathering wave data both "up-wave" and "down-wave", the responses could be compared, and the relative contribution of the non-linear component estimated. Image contrast resulting from the other mechanisms would be the remainder. The approach should be qualified with wave in ice data, then applied to the open water case. The first-order effects should not depend on coherence time limitations.

3.2 Non-coherent

It is known that the phase velocity of waves merits fundamental consideration in the SAR wave imaging problem (Raney and Vachon, 1988b). The principal ideas are addressed here under three general headings: scale distortion; multi-look spectral processing; and directional ambiguity removal. In each case, we now have experimental results from LIMEX/LEWEX data consistent with earlier predictions.

3.2.1 Scale Distortion

Consider a wave field moving at (phase) velocity C_O and direction ϕ relative to the along track vector of an observing scanning sensor with speed V_{TAN}. The resulting image will have a scale error along track in the amount

$$S = \left(1 - \frac{C_O}{V_{TAN}\cos(\phi)}\right) . \qquad (6)$$

The effect is very significant for most aircraft imaging geometries, but is negligible for most satellite imaging situations. Its application for scale correction, in either the image domain or the spectral wavenumber domain, is known.

The main message in scale distortion is that the scene reflectivity pattern moves during observation. For a SAR, the data required to form the total multi-look image may be gathered over an interval T_N of several seconds from each spatial position in the scene. The total *exposure time* T_N is an important parameter in this and related contexts, and is given by the product of the integration time constant and the angular extent $\sin(\beta)$ over which the SAR beam(s) gather data for looks.

3.2.2 Spectral Processing

For a moving scene, individual images of that scene taken at different times, will appear in different locations. Conventional SAR processing, which "always" combines looks as if they were motionless, creates motion blur in the composite image which subsequent processing cannot correct. The severity of the blur is proportional to T_N. As a result, directional spectra derived from such an image show decreasing contrast with increasing range, all else equal. (For some geometries, this is an additional cause of "azimuth cut-off", although from a non-coherent cause.) We denote the standard approach to such spectral calculations as "look-sum", since the looks are summed prior to Fourier transformation.

Some improvement in spectral contrast would result if the processor focus were adjusted. The "focus paradox" is resolved by noting that the relative position of imaged data is differentially shifted at the same time that focus is changed. For a simple azimuthal wave field, the apparent result is more satisfactory if the focus is perturbed. This is a false solution in general, however.

Better results are obtained in the following way (Raney and Vachon, 1988b; Raney *et al.*, 1989). Let each look have an integration time much smaller than the wave period, be imaged (power domain), and then Fourier transformed. (This requires either a modified SAR/processor combination, or access to the unprocessed SAR signal record.) The desired directional spectrum is formed by adding the magnitudes of the individual spectra. We denote this approach to spectral calculations as "spectral-sum". This is an energy conservative approach. It is impervious to non-coherent scene translation between looks and compensates for wavenumbers with both azimuth and range components. We strongly recommend it.

We demonstrate the contrast between these two methods of spectral calculation with the following example using wave in ice data gathered during LIMEX'87. Having formed the two different spectra from the same data set, the ratio of look-sum to spectral-sum is taken. The result is shown in Fig. 1. The unity value of the central (DC) pixel value illustrates conservation of energy for both methods. The spectral-sum method always shows higher spectral values, hence better wave contrasts, than does the look-sum approach.

3.2.3 Removal of Directional Ambiguity

Having acknowledged scene translation between looks, in the preceding section we described how to minimize the resulting dis-

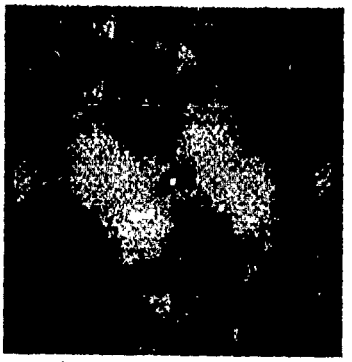


Figure 1: Ratio of look-sum to spectral-sum spectra for an example of LIMEX waves in ice data, 21 March, 1987. The central pixel (DC) value is unity. Grey levels are (proportionately) less than unity.

advantages. Here we comment on an advantage to be gained by using the same effect to eliminate the 180° ambiguity that otherwise plagues each directional component. (Details and results based on LIMEX data may be found in Vachon and Raney, 1989).

The directional ambiguity is removed by observing the direction of motion of each wave component in the set of looks. This may be done either explicitly (wave tracking or spatial correlation methods), or implicitly (Fourier analysis). (There are other methods which require two opposed directions of sensor travel. This is to be avoided, especially for satellites.)

The transform method is preferred. With N spectra available, one for each look, an N point Fourier transform (one sided) is performed over the set. The magnitude of the resulting spectrum is free of directional ambiguity. Confidence (and resolution) of the resulting spectrum is increased with larger N (more looks), and with *larger exposure time* T_N.

4 System Considerations

Several observations follow from the LIMEX/LEWEX experience. The general trend towards *smaller SAR dwell constant* (*e.g.* Beal, 1987) is supported. However, this trend should be treated carefully in light of spacecraft design considerations, acceptance of the coherence time limitation, and velocity bunching multi-mode imaging non-linearities. One could argue that the latter have been observed in SIR-B data due to the small SAR dwell constant! Also, we have seen evidence of non-linear behaviour for LIMEX ocean imagery in near range data with small dwell constant. There would seem to be a trade-off between small T_H with T_L constrained to be about the same size as, or slightly larger than, τ. Thus, the coherence factor may be used to some positive degree.

There is an alternative way to achieve a smaller SAR dwell constant without reducing the spacecraft altitude. One could choose an elliptical orbit, thus increasing the (periapsis) velocity of the spacecraft. (There is precedent for SAR designed for elliptical orbit, with Magellan as the prime example.) The SAR dwell constant that applies for selected elliptical orbits is given in Fig. 2. The key aspect is that the component V_{TAN} of spacecraft orbital velocity tangential to the (spherical) earth surface is the velocity which governs the image (Doppler) properties (Raney, 1986).

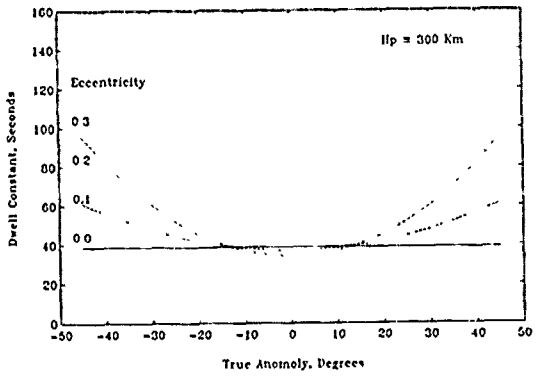


Figure 2: Representative dwell constants for elliptical orbits.

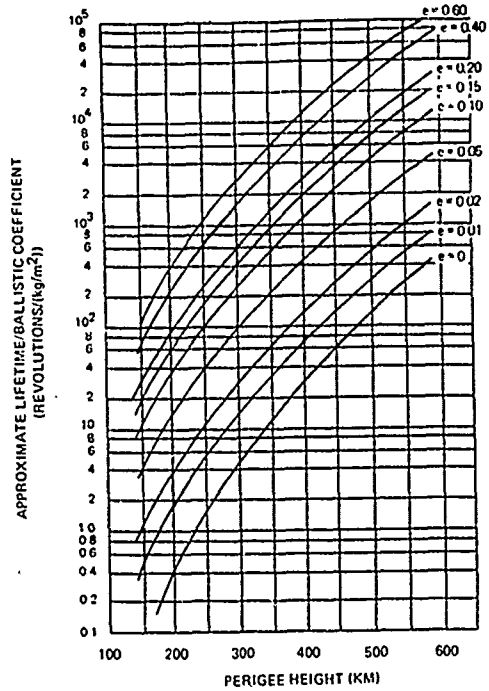


Figure 3: Drag coefficient as a function of elipticity and altitude.

It is apparent from Fig. 2 that there is only a modest impact on the dwell constant due to the small ellipticities suggested here. However, a substantial benefit may be found in the spacecraft orbital stability. Using data from Wertz (1978) as reproduced in Fig. 3, it is clear that with only a modest increase in orbit elipticity, two orders of magnitude reduction in drag coefficient would be realized. For example, rather than altitude boost every ten days or less, an elliptical orbit should require altitude boost only every year or so. This in turn reflects in much lower cost for spacecraft hardware, and should be given serious consideration as a modification to the "Spectrasat" concept (Beal, 1987).

The (multi-look) exposure time bears some reflection, for with T_N sufficiently long (considering the coherence time limited spatial resolution and the wave phase velocity of interest) the SAR by itself would be capable of producing unambiguous directional spectra. Additional instruments would not be required on the payload for this purpose. Exposure time is proportional to antenna beamwidth assuming only one antenna pattern, or more generally, to the angular sector used by the set of side-looking beams if more than one. If coherence time is indeed proportional to radar wavelength, the T_L/τ ratio would be independent of λ. For a longer wavelength, say P-band rather than C-band or even L-band, a simple antenna design (and lower power radar) would suffice.

Wavelength dependence of τ should be checked using JPL airborne radar data. Care should be taken to treat each data set for optimal spectral contrast using our spectral-sum approach. Otherwise, the increasing inter-look translation of the waves with increasing radar wavelength (T_L and T_N) would confuse and bias the conclusions against longer wavelengths.

When comparing airborne data sets to those obtained from or anticipated from spacecraft, the pertinent spatial and temporal scale factors (Table 1) should be carefully considered. In particular, for typical aircraft speeds, the relative phase velocity and the associated scale perturbation deserve attention.

References

1. Beal, R.C., "Spectrasat: A hybrid ROWS/SAR approach to monitor ocean waves from space", *Johns Hopkins APL Technical Digest*, Vol. 8, No. 1, pp 107-115, 1987.

2. Monaldo, F.M. and D.R. Lyzenga, "On the estimation of wave slope and height variance spectra from SAR imagery", *IEEE Trans. Geoscience Rem. Sens.*, Vol. GE-24, No. 4, pp 543-551, 1986.

3. Raney, R.K., "SAR response to partially coherent phenomena", *IEEE Trans. Antennas and Propagat.*, Vol. AP-28, No. 6, pp 777-787, 1980.

4. Raney, R.K., "Transfer functions for partially coherent SAR systems", *IEEE Trans. Aerospace Elect. Sys.*, Vol. AES-19, No. 5, pp 740-750, 1983.

5. Raney, R.K., "Doppler properties of radars in circular orbits", *Int. J. Rem. Sens.*, Vol. 7, No. 9, pp 1153-1162, 1986.

6. Raney, R.K. and P.W. Vachon, "SAR-seen multi-mode waves in ice: Evidence of imaging non-linearities", *Proc. IGARSS '88*, Edinburgh, Scotland, pp 141-144, 1988a.

7. Raney, R.K. and P.W. Vachon, "Synthetic aperture radar imaging of ocean waves from an airborne platform: Focus and tracking issues", *J. Geophys. Res.*, Vol. 93, No. C10, pp 12,475-12,486, 1988b.

8. Raney, R.K., P.W. Vachon, R.A. DeAbreau, and A.S. Bhogal, "Airborne SAR observations of ocean surface waves penetrating floating ice", *IEEE Trans. Geoscience Rem. Sens.*, (LIMEX/LEWEX Special Issue), Vol. 27, No. 5, Sept., 1989.

9. Vachon, P.W., R.B. Olsen, C.E. Livingstone, and N.G. Freeman, "Airborne SAR imagery of ocean surface waves obtained during LEWEX: Some initial results", *IEEE Trans. Geoscience Rem. Sens.*, Vol. 26, No. 5, pp 549-562, 1988.

10. Vachon, P.W. and R K. Raney, "Resolution of ocean wave propagation direction in single-pass airborne SAR imagery", *Proc. IGARSS '89*, Vancouver, Canada (this issue), 1989.

11. Wertz, J.R., Ed., Spacecraft Attitude Determination and Control, *D. Reidel Publishing*, Dordrecht, Holland, 1978.

DETECTABILITY OF OCEAN WAVES BY THE SYNTHETIC APERTURE RADAR ON ERS-1

Albert O. Scheffler, Clifford L. Rufenach, and Robert A. Shuchman

Radar Science Laboratory
Advanced Concepts Division
Environmental Research Institute of Michigan
Ann Arbor Michigan 48107 USA
(313) 994-1200
TELECOPIER (313) 994-0944 Ext. 2204
TELEX 4940991 ERIMARB

Johnny A. Johannessen
Nansen Remote Sensing Center
Bergen, Norway
47-5-297288
Telecopier 47-5-200050

Ocean surface waves are imaged by Synthetic Aperture Radar (SAR) through a combination of mechanisms, including surface slope, surface roughness, and surface motion. Over a limited range of ocean/radar conditions each of these mechanisms can be described by a linear modulation transfer function which allows the directional wave spectra to be inferred from the image spectra. However, it is generally understood that surface motion in most cases causes a non-linear transfer function and hence is the most important of these mechanisms.

The purpose of the present work is to determine when ocean surface waves can be correctly imaged or not imaged by the C-band SAR scheduled for launch on the first EUROPEAN EARTH SATELLITE (ERS-1) in 1990. More precisely, to determine under what conditions the modulation transfer function is linear and if non-linear under what conditions the waves can be correctly imaged or not imaged due to these surface motions. Indeed, as the non-linearity increases a distortion in the imaged spectral peak occurs resulting in a rotation of spectral peak toward the range direction. This distortion limits the ability to infer wave parameters from the imagery. However, if the non-linearities are not too large (slightly non-linear), in principle the intensity spectra can be corrected to obtain the directional wave spectra.

A two dimensional numerical simulation (LYZENGA, 1986) and a single parameter, c, suitable for characterizing the non-linearity (ALPERS, 1983) are used to investigate when the ERS-1 SAR can correctly image ocean waves, $c = 1200 * SWH * \cos \phi_p * \lambda_p^{-3/2}$ where ϕ_p is the peak wave direction relative to the flight direction (azimuth angle), λ_p is the peak wavelength, and SWH is the significant waveheight both in meters. In the case of wind seas, the simulation gives a linear range corresponds to approx. $c < 0.3$, whereas the highly non-linear case, where waves cannot be imaged, corresponds to approx. $c > 1.0$. Distorted waves in the imagery correspond to approx. $0.3 < c < 1.0$ and can be corrected for c near 0.3. Based on the above criteria and typical waveheights; say SWH = 2 m, the linear case corresponds to wavelengths longer than approximately 400 m for azimuth traveling waves or longer than 300 m for waves traveling 45 degrees to azimuth. The non-linear non-wave imaging case corresponds to wavelengths shorter than approximately 175 m for azimuth traveling waves or shorter than 150 m for waves traveling at 45 degrees. Ocean wavelengths between these two limits are called non-linear and for the wavelengths closest to the linear case (slightly non-linear) may be corrected to obtain the peak wavelength and direction. The above results are consistence with the airborne C-band SAR measurements taken during the NORCSEX experiment off the coast of Norway in March, 1988.

Alpers, W.R., "Monte Carlo Simulations for Studying the Relationship Between Ocean Wave and Synthetic Aperture Radar Image Spectra", J. Geophys. Res., 88 (C3), pp 1745-1759, 1983.

Lyzenga, D.R., "Numerical Simulation of Synthetic Aperture Radar Image Spectra for Ocean Waves", IEEE Trans. Geosci. Remote Sens., vol., GE-24 (6), pp 863-871, 1986.

Strategies to Assess the Capacity of on an On-Board Real-Time SIR-C Processor to Produce Ocean Wave Spectra

Frank Monaldo

The Johns Hopkins University Applied Physics Laboratory
Johns Hopkins Road, Laurel, Maryland 20707-6099
U.S.A.

Abstract

During the spaceborne imaging radar mission (SIR-C), scheduled for launch on a space shuttle in 1992, L-band, C-band, and X-band synthetic aperture radar (SAR) data will be acquired. Also on board the shuttle will be a processor designed to produce real-time estimates of two-dimensional wave image spectra from C-band, HH-polarization data. In an attempt to predict processor performance, we are pursuing several strategies: (1) We simulate the raw SAR signals of point targets and wave scenes and use a software simulation of the on-board processor to produce corresponding imagery and image spectra. (2) Raw, L-band, signal data from SIR-B is to be used to exercise the processor. (3) Computer simulations using SIR-B image spectra are used to determine appropriate averaging strategies for the on-board processor and to assess the precision with which the wave number and propagation direction of ocean surface waves can be determined. The results and progress of this work are discussed.

Keywords: synthetic aperture radar, real-time processing, ocean wave spectra.

1 Introduction

High resolution SAR imagery from the Seasat satellite and SIR-B SARs have demonstrated the potential to infer information about ocean surface waves The peaks found in two-dimensional SAR image spectra have shown a correspondence with independent measurements of the ocean (Monaldo and Lyzenga, 1988) Although there still exists some controversy (Alpers et al., 1986) as to the exact mechanisms by which a SAR is able to image waves and the optimum procedures for interpreting SAR wave imagery, there is a general consensus that the limiting factor in the usefulness of SAR image spectra is the apparent smearing caused by ocean surface motion. A SAR achieves high azimuth (along satellite ground track) resolution by using the Doppler information found in the return signal. Corruption of the Doppler signal by ocean surface motion can consequently reduce azimuth resolution. The extent of this degradation is proportional to R/V, the range-to-velocity ratio of the SAR platform.

The Seasat SAR, orbiting at an altitude of 800 km had an R/V = 120 s, about four times larger than the corresponding R/V of the SIR-B SAR orbiting at 230 km. Beal et al., (1986) showed that image spectra computed from SIR-B imagery exhibited considerably less azimuth resolution degradation than found in Seasat. These results argue for the use of a low-altitude as opposed to high-altitude orbiting SAR to retrieve the maximum amount of ocean wave information. One possible method to obtain ocean wave spectra on a global basis would be to launch a low-altitude, low-cost, single-purpose satellite dedicated to producing SAR wave image spectra. By limiting the SAR swath width to about 10 km and calculating image spectra on-board, it is possible to radically reduce the down-link data rate and minimize associated costs (MacArthur and Oden, 1987). A spaceborne processor capable of performing these functions is currently being fabricated at the Applied Physics Laboratory. The processor is to accompany the SIR-C SAR on board a space shuttle flight to provide a demonstration of the feasibility of such an approach. This paper discusses some of the efforts we have made to verify the processor design and predict its performance.

Figure 1 is a functional block diagram of the SIR-C processor. MacArthur and Oden (1987) give a more detailed description of processor design. Each output SAR image spectrum is comprised of the data accumulated in 1664, C-band radar pulses. The input video signal is filtered to 10 MHz, 1-bit quantized in I and Q, and ranged compressed by an array of 7, 64-bit correlators. After range correlation, each of the 1664 pulses has been sampled into 256 range cells of 30 m ground range resolution. Each two adjacent pulses are averaged together reducing the effective number of pulses to 832. The reduced number of pulses is dictated by hardware constraints. Azimuth compression is performed by using 256-point digital Fourier transforms, multipling by the appropriate azimuth chirp replicas, and performing the inverse Fourier transforms. Because the 256-point transforms are not sufficiently large to compress all 832 pulses at the same time, azimuth compression is performed in sections or sub-images and the entire image pieced together. Following azimuth compression, we discard the leading and trailing edges of the image which are contaminated by the fact that full azimuth compression is not possible at the edges. After this sequence, we are left with a complex array of 256 by 768 range and azimuth cells. Azimuth resolution is 10 m. The complex values are squared to yield image intensity and averaged by three along azimuth to reduce speckle noise. The final SAR image is an ar-

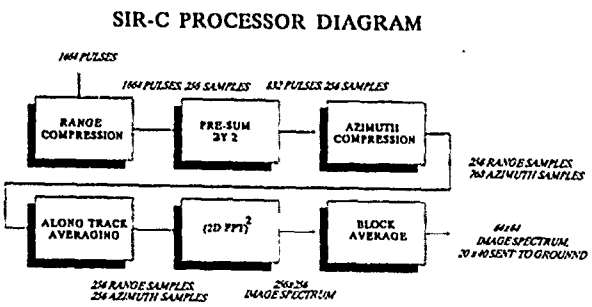


Figure 1 Schematic block diagram of the on-board SIR-C processor

ray of 256 by 256 values, each cell representing a 30 m by 30 m area on the surface.

A two-dimensional Fourier transform of the image is then computed and squared producing an image intensity-variance spectrum. One-half of this spectrum is redundant and can now be discarded. To reduce both sampling variability and the down-link data rate, the spectrum is averaged in blocks of 4 by 4 elements. The size of spectral array is thus reduced to 32 by 64. A sub-sample of 20 by 40 elements is then to be transmitted to the ground. At the rate of one spectrum per second the required down-link data rate is 6.4×10^3 bps (bits per second), down by a factor of 100 from the input signal date rate of 8.5×10^5 bps.

2 Test Strategies

We have pursued three strategies to test the capacity of and predict the performance of the on-board processor: (1) to simulate point targets and wave imagery to verify basic processor design, (2) to use actual SIR-B raw signal data to exercise the processor hardware, and (3) to simulate processor output spectra to verify spectral smoothing strategies. Significant progress on the first and third strategies has been achieved. The second strategy awaits the completion of the hardware fabrication.

2.1 Simulated Signals

The use of simulated point targets and wave image signals has been limited to the testing of the azimuth compression process and calculation of the two-dimensional spectrum. Since the Applied Physics Laboratory has been involved in a number of spaceborne altimeters with significant experience in the designing and fabrication of instrumentation to perform range compression, it is processing occurring after range compression that we are now most closely focusing on In addition, the azimuth compression is to be performed on sub-images and the resulting image constructed from these smaller images. There is concern that the edges of the sub-images might not be appropriately combined.

We tested the azimuth compression strategy by simulating the hardware functions on a multi-purpose computer, specifically an HP-9816. All the computations were made at 8 bit precision to mimic the calculations of the on-board processor Any consequent round off errors would therefore also appear in the software simulation.

The anticipated signals from a series of point targets in azimuth, after range compression, were input into the software simulation of the processor. Figure 2 is a representation of the image produced by the simulation. The point targets are accurately reproduced, even across sub image boundaries

Both the azimuth compression and Fourier transform processes were tested using simulated signal of wave images. A SAR image spectrum from SIR-B was inverse Fourier transformed and re-sampled to produce the corresponding SAR image. Using an inverse azimuth compression procedure, the uncompressed SAR signal was simulated. Again this artificial signal is used as input for the software simulation. Both the original SAR spectrum and the one computed from the simulation are shown in Figure 3. The result again verifies the processor design and the fact that sub-images are appropriately combined with no residual effects on the output spectrum.

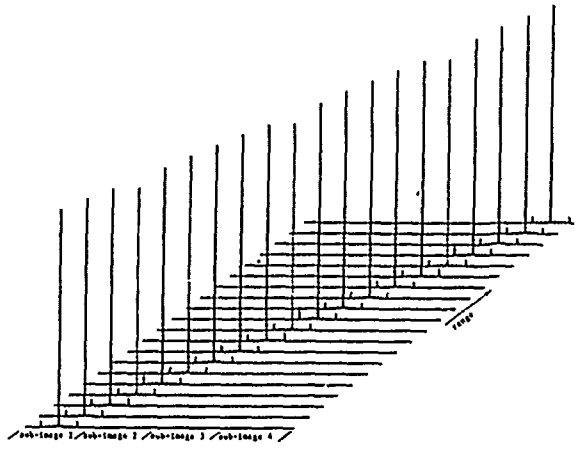


Figure 2 The result of azimuth compression of twenty point-targets by a software simulation of the on-board processor. Horizontal is the azimuth direction. Successive lines represent successive range lines. Twenty range lines are shown above.

2.2 Real Signal

Although SIR-B acquired only L-band SAR data and the on-board processor is designed to work with the C-band channel from the SIR-C SAR, much can be learned from the SIR-B signal. Our current plans are to take raw SIR-B signal from the particularly dynamic situation of Hurricane Josephine and use the actual processor to produce the corresponding image spectra. Although the resulting image resolution will not be as high when a processor designed for C-band is applied to L-band signals, the ocean waves are still sufficiently long to be imaged. Image spectra from his area have shown a particular sensitivity to the focus used to perform the azimuth compression and they should therefore provide a severe test for the processor. The implementation of this test awaits the fabrication of the hardware.

2.3 Performance

A spectrum of a SAR image is, at best, a single realization of the ensemble-mean spectrum (Monaldo, 1989). Each individual spectral value represents a random variable having a χ^2-distribution with two degrees of freedom. The mean of this distribution is the value of the corresponding ensemble-mean spectral value. The standard-deviation of a χ^2-distribution having only two degrees of freedom is equal to the mean. Any single spectrum produced by the on-board processor is therefore a noisy representation of the ensemble-mean spectrum. This type of noise is often referred to as sampling variability. Each 256 by 256 point spectrum computed by the on-board processor is block averaged to reduce both sampling variability and the down-link data rate

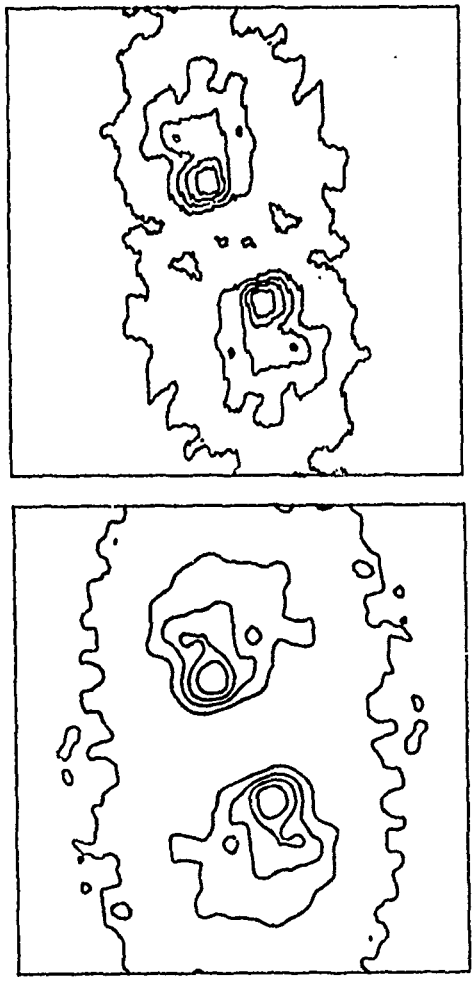


Figure 3 The top contour plot is a representation of a SAR 256 by 256 element image spectrum from SIR-B. The imagery, from which the spectrum was calculated, was acquired from data take 91.5 off the coast of Chile. The center of the box is zero wave number. The edges of the box represent a wave number of $2\pi/(60m)$. This spectrum was used to simulate an artificial SAR signal which before azimuth compression. The lower plot is the 64 by 64 spectrum computed by the software simulation of the on-board processor from this artificial signal. The spectrum has been smoothed to about 100 degrees of freedom.

Every group of 4 by 4 elements is averaged together producing a 64 by 64 array. The resultant averaged spectrum now has 32 degrees of freedom and consequently less sampling variability.

The effect of sampling variability on the precision with which dominant ocean wave number and propagation direction can be determined was estimated with a computer simulation. We start with a smoothed, 256 x 256 element SIR-B SAR image spectrum. This spectrum is asserted *a priori* to be the ensemble-mean spectrum. Noise, having a χ^2-distribution with two degrees of freedom is applied to the spectrum. The resulting spectrum is a simulated realization of the ensemble-mean spectrum. Every group of 4 by 4 elements of this spectrum is averaged, in a manner analogous to the averaging done by the on-board processor. This 64 by 64 point spectrum is a simulated output spectrum from the processor having 32 degrees of freedom. The dominant wave number and propagation direction, calculated using a center-of-mass algorithm for the spectrum is calculated. This procedure was repeated 50 times.

The root-mean-square variability on the wave number estimate was found to be 1.6×10^{-3} rads/m. For a 200 m wavelength wave system, this variability is equivalent to an precision of 10 m in wavelength. Propagation angle was estimated with a root-mean-square variability of 1.3°. Both values are comparable to the precision attainable from SIR- B image spectra processed on the ground.

3 Conclusions

Using a software simulation of the on-board SIR-C processor, we have verified the azimuth compression and spectral analysis design. We are currently awaiting hardware fabrication to exercise the processor with actual SIR-B SAR signal. Further, the result of computer simulations suggest that we may expect that the precisions with which wave number and propagation direction can be determined are 1.6×10^{-3} and 1.3^0, respectively.

4 References

Alpers, W.R., C. Brüning, and K. Richter, "Comparison of simulated and measured synthetic aperture radar image spectra derived from ocean wave spectra during the Shuttle Imaging Radar-B mission", IEEE Trans. Geosci. and Remote Sensing, Vol. 88, pp. 559-566, 1986.

Beal, R.C., F.M. Monaldo, D.G. Tilley, D.E. Irvine, E.J. Walsh, F.C. Jackson, D.W. Hancock III, D.E. Hines, R.N. Swift, F.I. Gonzalez, D.R. Lyzenga, and L.F. Zambresky, "A comparison of SIR-B directional wave spectra with aircraft scanning radar spectra", Science, Vol. 232, pp. 1531-1535, 1986.

MacArthur, J.L. and S.F. Oden, "Real time global wave spectra from SIR- C: systems designs", 1987 IEEE Int. Geosci. and Remote Sensing Symposium Digest, Vol. II, pp. 1105-1108, 1987.

Monaldo, F.M. and D.R. Lyzenga, "Comparison of Shuttle Imaging Radar-B ocean wave imagery with linear model predictions", J. Geophys. Res., Vol. 93, pp. 15,374-15,388, 1988.

Monaldo, F.M., "The effect of sampling variability on the estimation of wave number and propagation direction from SAR wave spectra", IEEE Int. Geosci. and Remote Sensing Symposium Digest, this issue, 1989.

THE INFLUENCE OF PLATFORM ATTITUDE DYNAMICS ON THE DOPPLER FREQUENCY SYNTHESIS OF SIR-B OCEAN DATA NEAR HURRICANE JOSEPHINE

David G. Tilley, Joyce L. Kerr and Cynthia L. Vogt

The Johns Hopkins University/Applied Physics Laboratory
Johns Hopkins Road, Laurel, Maryland 20707-6099
U.S.A.

ABSTRACT

The space shuttle Challenger was quickly maneuvered into a SAR data acquisition mode to capture a SIR-B scene of Hurricane Josephine which presented itself as an unscheduled target of opportunity during NASA Mission 41-G during October of 1984. Steeply sloped seas approaching 10 meters in height were observed several hundred kilometers in advance of the hurricane track. Mission attitude data indicate that the shuttle pitch, roll, and yaw were unstable during the data take. The statistical clutterlock employed by our ground SAR processor yields a Doppler frequency estimate that tracks the shuttle yaw variation but is ambiguous with a single discontinuity equal to one multiple of the radar pulse repetition frequency. The SIR-B signal data have been compressed under several Doppler parameterizations to investigate the ambiguity of the clutterlock. Resultant ocean images are analyzed for their contrast and resolution properties.

INTRODUCTION

A critical stage in synthetic aperture radar (SAR) image processing is the coherent Doppler correlation of the radar signal $S(r,v,t)$; measured in terms of the downrange distance, r, scene velocity v, and illumination time t, with a reference signal producing a two dimensional mapping of cross sectional reflectivity $\sigma(r,a)$ in terms of the downrange coordinate and the along track coordinate $a = \text{Dopp}^{-1}(v,t)$. The azimuth (i.e., along track) compression of the radar data is parameterized by the median Doppler frequency shift observed for a discrete number of radar pulses defining a synthetic array of antenna elements along an orbit in space. A statistical clutterlock procedure [Li et al., 1985] is often used to compute a Doppler centroid from azimuth frequency spectra of complex SAR data after range compression to exclude the backscatter power contributed by targets illuminated by a partial aperture. The attitude of the SAR platform influences the median Doppler shift,

$$f_c = \frac{2}{\lambda} \vec{v} \cdot \vec{r} \qquad (1)$$

for the surface scattering elements illuminated by the pulsed radar with wavelength λ. The target velocity distribution is influenced by the trajectory of the spaceborne SAR platform in relation to a rotating earth and by wave motion at the ocean surface. The latter may be considered a random influence that broadens and shifts the Doppler spectrum. The SAR orbital geometry is usually known well enough so that downrange velocity of the shuttle at its point of closest approach to a given surface picture element (pixel) along track can be used to deterministically define a nominal antenna pattern at beam center. A matched filter function can be correlated with the frequency spectrum of the SAR signal to yield an estimate of the Doppler centroid prior to range compression. However, when the Doppler matched filter is based on a nominal antenna pattern without accurate data on the attitude of the SAR platform, an arbitrary choice of the central pulse corresponding to beam center in the orbital array of discrete illumination elements results in a mainlobe versus sidelobe ambiguity in the definition of the synthetic aperture antenna. When the azimuth pulse ambiguity is the result of an unknown shuttle pitch and yaw, the physical antenna may be squinted fore or aft of the spacecraft rather than side-looking. Hence, the correct center Doppler frequency employed in defining the SAR reference signal could be different from the estimated Doppler centroid by one or more multiples of the pulse repetition frequency. If the shuttle attitude is known the pulse ambiguity can be resolved and the Doppler centroid estimate should track the squint variation to align the mainlobe of the synthetic aperture with the mainlobe of the physical aperture. The radar look angle θ, pitch angle, ϕ and yaw angle γ are usually specified in a frame of reference centered at the spacecraft [Li and Johnson, 1983] relative to a plane containing the earth's angular velocity vector and the spacecraft velocity vector. The earth's rotational velocity at the radar target, v_t, can be placed in a parallel plane by making a flat earth approximation at the location of the spacecraft. The directional cosine of the spacecraft yaw aligns these vectors in their respective planes while the directional sine of the earth look angle and spacecraft pitch angle project the velocity vectors onto the downrange coordinate of the radar. Hence, the Doppler centroid,

$$f_c = \frac{2}{\lambda} (v_s \cos\gamma - v_t) \sin(\theta - \phi), \qquad (2)$$

can be estimated from the earth's rotational velocity and the ground track velocity, v_s, of the space shuttle when its orbital attitude and trajectory are known.

On October 12, 1984 the space shuttle imaging

radar (SIR-B) aboard Challenger recorded 90 seconds of SAR signal data covering a 600 kilometer strip of The North Atlantic Ocean at 30°N and 72°W in the near vicinity of Hurricane Josephine. This SIR-B data set provides an excellent example of the pulse ambiguity along track complicated by an extremely dynamic sea state. These data were initially processed through the NASA ground SAR processor at The Jet Propulsion Laboratory (JPL) in Pasadena, California, assuming that the shuttle had stabilized in a right-looking mode at a look angle of 24.6 degrees. Subsequent analyses [Gonzalez et al., 1987] of the ocean imagery revealed a discontinuity in wave contrast during the data take. It was later discovered that image contrast and resolution could be improved [Tilley, 1988] by using a Doppler centroid approximately equal to the pulse repetition frequency, 1363 Hz, during azimuth compression The unusually large value, 1333 Hz, of the Doppler centroid was estimated by a statistical clutterlock procedure used by a ground SAR processor [McDonough et al., 1986] operated by The Johns Hopkins University Applied Physics Laboratory.

Subsequent examination of PATH (i.e., postflight attitude and trajectory history) tapes for NASA Mission 41C archived at JPL indicate that the space shuttle Challenger was in an unstable attitude during the early portion of the data take. The yaw, pitch, roll and altitude data from the PATH tape are plotted in Figure 1. Note that the shuttle was descending from an altitude of 228 km at a rate of 14 m/s. The yaw, pitch and roll of the shuttle, however, all migrate through an inflection point at about 16.31.00 hrs GMT The yaw and roll each experience a second inflection point of opposite sign at 16.31.39 hrs GMT (i e., 39 seconds later). It appears that the shuttle was being maneuvered into an unscheduled right looking SAR mode at a yaw of 270° and that the data take may have begun before the shuttle had stabilized in that mode. Hence, in this unusual case, SAR image processing requires frequent recalculation of the Doppler centroid so that azimuth compression can be accomplished with a coherent reference function that tracks changes in the attitude of the physical antenna.

DATA PROCESSING PROCEDURES

Although the Doppler frequency at beam center can in theory be determined from the spacecraft ephemeris and attitude data, these data are generally not known with sufficient accuracy to provide an azimuth filter parameter which results in high-quality imagery. The Doppler centroid is, therefore, estimated from the raw data with a "clutterlock" algorithm. In this method, 4096 pulses of uncompressed raw data are Fourier transformed to complex data and converted to baseband by discarding the negative frequencies in the spectrum and downshifting the spectrum from the video offset frequency. The complex data are then reduced by 1) averaging groups of eight pulses to give 512 range lines, and 2) selecting 100 samples across the swath at intervals of 25 samples. At each of the 100 range distances, the azimuth frequency spectrum is correlated with the radar's antenna pattern. The peak of the correlator output represents an estimate of the Doppler center frequency for that range distance. The straight line which is fitted through a plot of the 100 peak amplitudes against slant range distance gives the desired estimate of the Doppler centroid,

$$f_c = a_1 + a_2 \left[1 - \left(\frac{h}{r}\right)^2\right]^{1/2} \qquad (3)$$

where h is the altitude of the SAR platform and the coefficients a_1 and a_2 are computed by linear regression of the squared error.

A VAX 11-780 computer was used to host an FPS-164 array processor that processed input files containing 22 Mbytes of complex data to output approximately 2 Mbytes of single look image data. The Doppler centroid was computed for 3 second batches of signal data at 6 second intervals along track, while the Doppler rate was fixed at a value suggested by an earlier experiment [Tilley, 1988]. Each of the 14 batches of SAR input data were compressed in azimuth with a single look process employing their unique Doppler parameterization. Figure 2 depicts the variation in the Doppler characteristics employed during the 90 seconds of the SAR data take. Note that the Doppler rate parameter remained constant at about -1844 Hertz/sec throughout while the Doppler centroid varied from 268 Hertz to 2931 Hertz, more than twice the pulse repetition frequency. The resultant image data amounted to about 20 kilometers along track for each batch in a ground range perspective with 5.5 m pixels.

The image segments were reduced to 11 meter square pixel representations by averaging 2x2 arrays of the smaller pixels and segmented into 4 databases of 512 x 512 pixels to represent each batch. Figure 3 depicts the single look image data near the middle of the data take with comparison to the four look imagery at a similar, if not identical, location.

The 4 databases representing each 3 second batch of SAR ocean imagery were subjected to fast Fourier transform routines after squaring pixel values to form 16-bit radar cross sections. Pre-processing was limited to the subtraction of a mean cross section computed over the 512 x 512 pixel databases. Post-processing of the resultant SAR image power spectra was limited to a power normalization [Irvine and Tilley, 1988] that equated the power spectral variance with the variance-to-mean squared ratio of the image database. The 4 power spectra representing each 3 second SAR segment were averaged to yield a single image spectrum for each of the 14 batches of APL processed ocean segments. Figure 4 compares the Fourier spectra for the single look APL process with the Fourier spectra computed in a similar manner from the 4-look JPL processed imagery.

DISCUSSION

The ideal SAR platform moves over the earth at a constant altitude while maintaining the physical antenna in a plane parallel to its direction of travel. When these conditions are met, the platform velocity does not have a component of velocity in the downrange direction that contributes to the Doppler frequency shift (i.e., equation 1) of the pulsed radar signal. When the space shuttle yaw angle squints the radar antenna either forward or backward of the right-looking mode, the horizontal shuttle velocity (e.g., v_s = 6700 m/sec) will have a component in the downrange direction, particularly when the difference in the radar look and pitch angles is large. Comparing the yaw angles in Figure 1 with the clutterlocked Doppler frequencies in Figure 2, it appears that the horizontal shuttle velocity should not be neglected in SAR data compression. Furthermore, for this data take over Hurricane Josephine the shuttle platform was descending at a constant rate of 14 m/sec. When this vertical velocity is projected in the side-looking direction of the SAR at a look angle of 24.6 degrees, the rate of descent amounts to approximately 108 Hertz of the Doppler shift.

The APL processor compensates for range migration [McDonough et al., 1986] of the Doppler centroid with empirical coefficients, a_1 and a_2. Using the flat earth approximation (i.e., $\cos(\theta-\phi) = h/r$), Equation (3) can be interpreted in terms of the horizontal

component of shuttle velocity, v_s, and the rotational velocity, v_t, of a target at the earth's surface. Including the 108 Hertz shift owing to the vertical shuttle velocity, dh/dt, with the predictions of equation 2, the APL coefficients can be expressed:

$$a_1 = \frac{2}{\lambda}\left[\frac{dh}{dt}\cos\theta + v_s \cos\gamma \sin(\theta-\phi)\right] \quad (4)$$

and

$$a_2 = -\frac{2}{\lambda} v_t \quad (5)$$

As the earth rotates through a circle equal to the circumference of the earth (40,212 km) in one day (86,400 s) at a latitude, α, the rotational speed of a surface target is

$$v_t = 465 \cos(\alpha) \cos(\beta) \text{ m/sec} \quad (6)$$

relative to the ground track of the SAR moving in a direction, β, east of a northern heading. For the Hurricane Josephine data take $\beta = 127^o$ and latitude varied from 36.1^oN to 32.5^oN. Table 1 lists significant parameters that can be used to relate the Doppler centroids estimated empirically and used in this study with those computed using equations 3-6.

In comparing the contrast properties of the SAR ocean imagery generated by the JPL and APL scene correlators, major differences in both the coherent and non-coherent statistical procedures for computing cross sectional intensities exist and should be considered. The JPL 4-look process involves a short coherent integration period (e.g., 0.6 seconds) to achieve 24 meter resolution, approximately. Incoherent temporal averaging follows as 4 sequential looks are overlayed spatially and averaged. The APL single look process involves a longer coherent integration period (e.g., 3 seconds) to achieve 6 meter resolution, approximately. Incoherent spatial averaging follows as 2x2 arrays of pixels in range and azimuth are averaged to yield 11 meter pixels. The SAR variance-to-mean squared ratios for the image intensity were greater than 1.0 for the 11 meter APL pixels and about a factor of 4 less for the 12.5 meter JPL pixels. In both cases the incoherent statistics involve 4 samples. However, the single look process attempts to obtain a coherent spatial resolution better than the pixel size achieved after incoherent processing whereas the 4-look process attempts to improve the temporal coherence of the cross section measurement.

Table 1
Doppler Centroid Estimates

Time (min:sec)	Latitude (deg)	Yaw (deg)	Pitch (deg)	Eq.3 (Hz)	Fig.2 (Hz)
31:09	36.1	271.5	-33.8	2805	2864
31:19	35.6	270.5	-34.4	1978	1563
31:29	35.1	269.2	-34.6	882	575
31:39	34.6	268.6	-34.6	379	268
31:49	34.0	268.7	-34.7	477	452
31:59	33.5	268.8	-34.7	573	667
32:09	33.0	268.9	-34.7	668	884
32:19	32.5	269.1	-34.8	850	1081

CONCLUSION

Knowledge of SAR platform attitude can be applied to compute approximate values of the Doppler centroid based upon considerations of orbital geometry and scene velocity. When such approximations are used together with empirical estimates, computed by a statistical clutterlock algorithm, ambiguities equal to one or more multiples of the pulse repetition frequency can be resolved in the selection of an appropriate Doppler centroid. For the SIR-B data take HJ117.40 near Hurricane Josephine, the center Doppler frequency of the signal data tracks the shuttle trajectory and attitude variations as archived on NASA Mission 41G PATH tapes. Unusually large values of the Doppler centroid are required in excess of the pulse repetition frequency early in the data take before the shuttle had stabilized in a right-looking mode. The APL clutterlock procedure results in a smooth evolution of the ocean wave spectrum as it varies along the SAR track.

Residual differences in the statistical and deterministic Doppler centroid estimates are the result of ensemble surface wave motions. These effects should be included in a deterministic model of the Doppler frequency variation with wave phase in future research activities.

Comparisons of one-look and four-look imagery to assess differences in the Doppler clutterlock estimation procedures and subsequent SAR azimuth compressions are difficult. Hence, future investigations are needed to optimize Doppler parameters in a four-look azimuth compression. The single look process achieves a better wavenumber response for the hydrodynamic wind-wave modulation that is important in the imaging of high sea states. The four-look process achieves better suppression of image speckle which may be helpful in imaging temporally coherent scenes via the velocity bunching and tilt modulation mechanisms.

ACKNOWLEDGMENT

SAR signal data and PATH tape data were provided by B.M. Holt at The Jet Propulsion Laboratory in Pasadena, California. The authors have benefitted from many helpful discussions with R.C. Beal, R.N. McDonough and B.E. Raff at the Applied Physics Laboratory in Laurel, Maryland. The NASA Office of Space Science and Applications has funded this research.

REFERENCES

1. Li, F.K., D.N. Held, J.C. Curlander and C. Wu, "Doppler parameter estimation for spaceborne synthetic aperture radar", IEEE Trans. Geosci. Remote Sensing, GE-23, No. 1, 47-56, 1985.

2. Li, F.K. and W.T.K. Johnson, "Ambiguities in Spaceborne Synthetic Aperture Radar Systems", IEEE Trans. Aerospace Elec. Sys., AES-19, No. 3, 389-397, 1983.

3. Gonzalez, F.I., B.M. Holt, and D.G. Tilley, "The age and source of ocean swell observed in Hurricane Josephine", The Johns Hopkins APL Tech. Dig., 8, No. 1, 94-99, 1987.

4. Tilley, D.G., "The influence of processor focus on speckle correlation statistics for a Shuttle Imaging Radar scene of Hurricane Josephine", Proc. of the 32nd SPIE Symposium, Statistical Optics, 976, 20-30, San Diego, CA., 18-19 August 1988.

5. McDonough, R.N., B.E. Raff and J.L. Kerr, "Image formation from spaceborne synthetic aperture radar signals", The Johns Hopkins APL Tech. Dig., 6, No. 4, 300-312, 1986.

6. Irvine, D.E. and D.G. Tilley, "Ocean wave directional spectra over the Agulhas Current from the SIR-B synthetic aperture radar", J. Geophys. Res., 93, No. C12, 15,389-15,401, 1988.

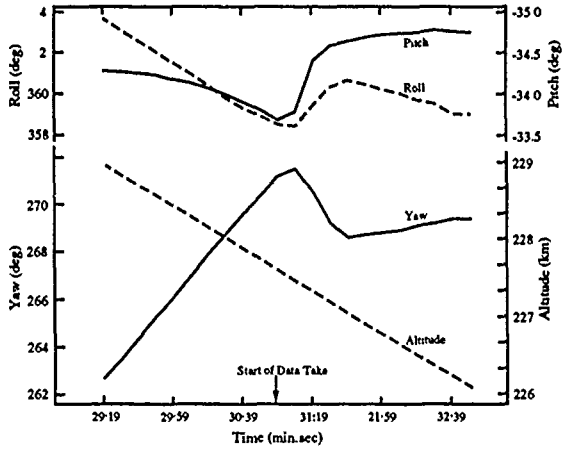


Figure 1 Shuttle ephemeris vs time (yaw, pitch, roll, and altitude vs min:sec).

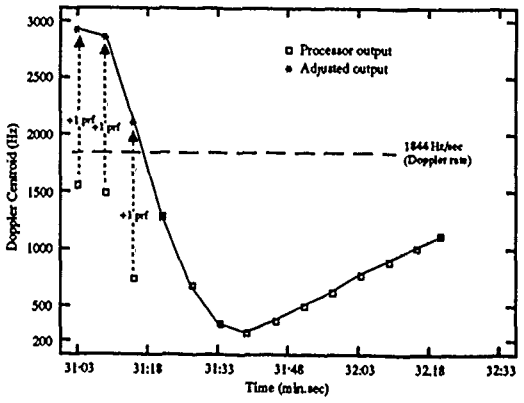


Figure 2 Doppler processing parameters vs time

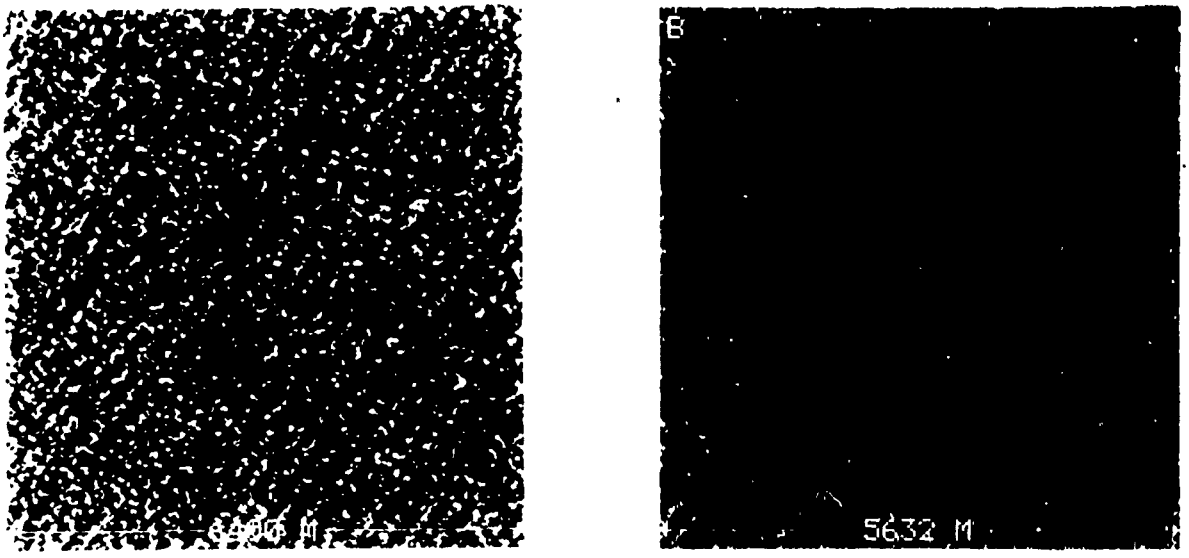


Figure 3 Four-look (a) and single look (b) SAR images have variance-to-mean squared ratios of 0.84 and 2.20, respectively.

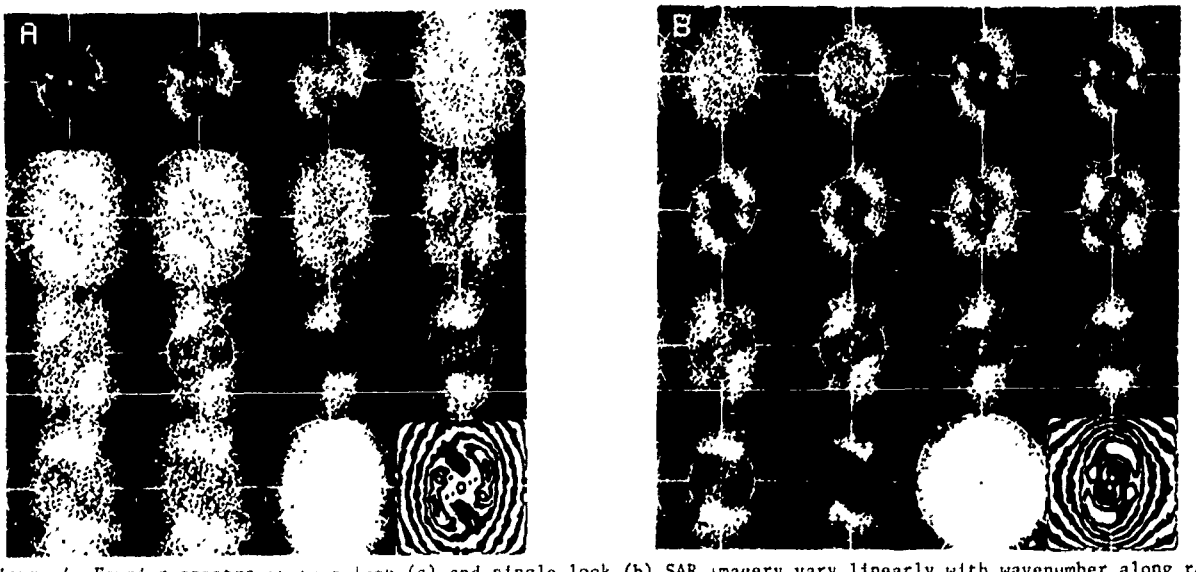


Figure 4 Fourier spectra of four-look (a) and single look (b) SAR imagery vary linearly with wavenumber along radii from the center out to limiting values of $2\pi/100$ and $2\pi/88$ rad per meter at 14 locations along track.

RADAR SURFACE SIGNATURES BASED ON THE TWO-DIMENSIONAL TIDAL CIRCULATION OF PHELPS BANK, NANTUCKET SHOALS, MA.

S. R. Chubb[1], G. R. Valenzuela[2], and D. A. Greenberg[3]

[1]Sachs/Freeman Assoc., Inc.
Landover, Md. 20785-5396
[2]Naval Research Laboratory
Washington, D. C. 20375-5000
[3]Bedford Institute of Oceanography
Dartmouth, Nova Scotia, Canada B2Y 4A2

ABSTRACT

Numerical results for the M_2 tidal currents over Phelps Bank (Greenberg et al., 1989), with 1/8' x 1/8' resolution, are used to predict the radar surface signatures (cross sections σ) of Phelps Bank using the Alpers and Hennings (1984) first order relaxation model. The numerical predictions for σ incorporate all two dimensional contributions from the tidal currents and include both the residual and first harmonic terms. Not unexpectedly, strong correlation between spatially resolved, contour plots of σ, over a wide range of directions and time intervals within the tidal cycle, with variations in the bottom topography of the Phelps Bank is obtained even when Bragg wave scattering alone is included. Additional contributions to the radar surface signatures should come from other scattering mechanisms (i.e., specular, wave-breaking, etc.) and from the shoaling of long ocean waves over the bottom features modulating the energy of the Bragg resonant waves. The predictions for σ are contrasted with previous measurements in Phelps Bank by NRL (Valenzuela et al., 1985), and a discussion is given of the importance of including the off-diagonal elements from the current stress tensor in the evaluation of σ.

Keywords: SAR imaging, Bragg scattering, current-wave interaction, numerical modeling

I Introduction

It has been well-established that in regions of shallow water, the bottom topography of the ocean floor often can be observed through Synthetic Aperture Radar (SAR) and Real Aperture Radar (RAR) imaging of the ocean surface. This phenomenon has been identified in SAR images (Fu and Holt,1982; Lodge, 1983a,b; Kenyon, 1983) derived from data associated with the Seasat satellite mission in 1978 (Born et al., 1979) and in RAR imagery (de Loor, 1976, 1978; de Loor and Brunsveld van Hulten, 1978; de Loor, 1981) of the North Sea. Extensive measurements over Phelps Bank, Nantucket Shoals performed by the Naval Research Laboratory in 1982 (Valenzuela et al., 1985) provided further evidence of the phenomenon.

Although the phenomenon is well-documented by these various experiments, a complete, theoretical formulation of the processes involved has not yet been established. A major reason for this fact is the indirect nature of the imaging process: microwave backscatter into SAR's and RAR's is primarily from the ocean surface since the wavelength of the microwave radiation (on the order of tens of centimeters) is considerably larger than the penetration depth (typically 1 centimeter) of the associated electromagnetic fields below the surface, and the water depth is of the order of tens of meters. Complicating the analysis further, especially in rough oceans, is the introduction of non-linear hydrodynamic phenomena such as wave-breaking and wave-cresting.

On the other hand, recently, Alpers and Hennings (1984; henceforth, AH) have pointed out that some of the correlation between features associated with the ocean bottom and images of the surface observed both in de Loor's RAR measurements and in Seasat SAR data could be accounted for directly within the context of the two-scale scattering model (Keller and Wright, 1975; Hughes, 1978; Wright, 1978) in which an implied coupling between short, surface (Bragg) resonant waves and the bottom topography emerges in radar cross sections (σ) as a natural consequence of the single wave (relaxation time approximation) scattering of short gravity-capillary waves through the large-scale motions of the ocean surface.

In particular, it has been observed that within this two-scale, Bragg-scattering model, the ratio between the deviation of the cross-section ($\delta\sigma$) at a given point in space from its mean value (σ_o) to the value of σ_o can be directly related to the gradient of the tidal current distribution in the limit in which Bragg wave scattering provides the only source of radar return. In the past various simplifying assumptions have been made, 1) associated with the dimensionality of tidal flow (assumed to be one-dimensional), 2) the flow be laminar free and very nearly uniform, 3) to neglect the "local" scattering terms relative to the advective and relaxation terms in the radiation balance equation, and 4) in many cases it is valid to neglect even the advective terms.

These approximations allowed AH to make estimates of $\delta\sigma/\sigma_o$, based solely on a) an estimate of a single component of the tidal velocity evaluated at one point in space, b) a knowledge of the depth profile (bathymetry) and its variation along the direction of tidal flow, and c) a knowledge of the "look" angle of the radar relative to the tidal flow. Even with these assumptions, however, AH have been able to correlate, somewhat quantitatively, various calculated features in the surface signatures with variations in the bottom topography in reasonable agreement with findings derived from SAR images of the North Sea.

The goal of the present work has been to investigate more quantitatively the implications of applying a full-fledged, two-scale scattering model, in which the underlying assumptions about dimensionality, laminar free flow and near uniformity of flow of the currents

made by AH are no longer valid to the case of Phelps Bank, Nantucket Shoals. We have done this using direct, numerical calculations of $\delta\sigma/\sigma_o$, based on a realistic, calculated description of the tides, in which 1) the tidal flow is a two-dimensional solution of the Navier-Stokes (NS) equation (Greenberg et al., 1989), and, hence, 2) there is no restriction associated with the form of flow (other than the imposition of open boundary conditions). Within this context, the coupling of bathymetry to radar return enters indirectly, as a boundary condition associated with the depth profile of the region that is imposed on the solution of the NS equation, and $\delta\sigma/\sigma_o$ is evaluated numerically, using the gradient of the simulated barotropic (depth-averaged) currents. Thus, we model the surface currents using depth-averaged solutions of the NS equations. Therefore, we make none of the simplifying assumptions associated with approximating the tidal flow as one-dimensional, laminar free and nearly uniform. We do assume, consistent with the two-scale model, that Bragg scattering provides the main contribution to the radar return. However, we do neglect effects from the shoaling of long ocean waves over the bottom features modulating the energy of the Bragg resonant waves and the presence of non-linear effects relating to the surface wave field and contributing to the radar return. Also, for bottom features of large spatial scale, the advection scattering effects are insignificant compared with relaxation effects Velocity bunching effects are neglected in this work.

II Results and Discussion

We have evaluated $\delta\sigma/\sigma_o$ by using an expression that in essence was derived by AH. In particular, we have used

$$\delta\sigma/\sigma_o = -\frac{(\alpha+1-\gamma)}{\mu}\sum_{i,j}\epsilon_i\,\epsilon_j\,\frac{\partial U^i}{\partial x^j}$$

$$= -\frac{(\alpha+1-\gamma)}{\mu}\left(\sin^2\phi\,\frac{\partial U^x}{\partial x}\right.$$

$$\left. + \sin\phi\cos\phi\,\left(\frac{\partial U^y}{\partial x}+\frac{\partial U^x}{\partial y}\right) + \cos^2\phi\,\frac{\partial U^y}{\partial y}\right) \quad (1).$$

Here, in the first line of Eq 1, ϵ_j denotes the j^{th} component (x or y) of a unit vector that points along the projection of the wave-vector of the incident signal onto the plane of the surface, U^i is the i^{th} component of the tidal current, and the x[y]-direction is directed eastward [northward]. In the second line, ϕ is the "look" angle (measured clockwise, relative to north). The relaxation constant is μ, the straining constant is

$$\gamma = 1/2\ (1 + 3\ k^2\ T)/(1 + k^2\ T) \quad (2),$$

where $T=T/(g\rho)$. Here, g is the gravitational constant, T is the surface tension, ρ is the density of water, and we have adopted the assumption of "Phillips equilibrium" for the short-wave energy spectrum, E_o, of the ocean at the surface in which

$$E_o = \omega^2 k^{-(\alpha+1)} \quad (3),$$

where ω is the radian frequency of the Bragg wave. Throughout, k is the Bragg wave-vector = 2 k_o, where k_o is the projection of the radar wave vector onto the surface plane. There are minor differences between our Eq. 1 and Eq 38 of AH that result because of different conventions associated with our respective definitions of ϕ and in the assumed functional forms of E_o.

AH present this relationship as an intermediate step in the development of their paper but never fully exploit it because of the various approximations that they employ. In fact, the sole approximation associated with the relationship (Eq. 1) beyond the assumption that Bragg scattering provides the dominant contribution to the backscatter is that advective contributions from the radiation balance equation can be neglected. Of significance for the present work is the fact that the relationship can be applied to arbitrary current distributions.

We have numerically evaluated the values of $\delta\sigma/\sigma_o$ for a variety of different "look" angles and times during the tidal cycle using the depth-averaged currents, derived from the barotropic model of Greenberg (Greenberg, 1983), including periodic and residual contributions, as applied to the Phelps Bank, Nantucket Shoals (Greenberg et al., 1989), and Eq. 1. In Fig. 1, we present representative, spatially resolved contour maps for two "look" directions (east and north) of this quantity, evaluated for one case (t=0).

The relationship of this plot to actual images of Phelps Bank is complicated by 1) the possibility of temporal and spatial variations in the value of the (usually assumed constant) prefactor $-(\alpha + 1-\gamma)/\mu$ found in Eq. 1, 2) the importance of the effects (cited above) associated with non-linear dependencies of radar return on surface wave height spectrum as well as the importance of advection terms (which introduce a coupling to the wind, for example) within the model, 3) the need for precise, temporally and spatially resolved M_2 tidal information associated with the area, which is necessary for correlating the physical time evolution of the tidal cycle with the result of our calculation, and 4) the coarseness of the 1/8′ x 1/8′ mesh in relation to the SAR resolution and finite size effects associated with the calculation.

On the other hand, various features of the plot do seem to be relevant to our understanding of SAR and RAR imagery of the region. In particular, there is in fact a strong correlation between the region of large positive return (marked by arrows) for eastward look angles with the steep fluctuation in bathymetry (see Fig. 2) along the bank. Additional analysis (Chubb et al., in preparation, 1989) reveals that this feature persists in time throughout a large portion of the tidal cycle, indicating that when realistic currents are used, even within a first order relaxation/Bragg model, a strong correlation exists between the surface signatures and bottom topography of Phelps Bank for some instances in the tidal cycle.

As expected, there is a rather strong dependence of the surface signatures on look angle. As a consequence, though there is evidence of large backscatter from the same region of the bank in the case of a northward look angle, the magnitude of the surface signature is significantly reduced. Furthermore, the two plots in Fig. 1 are quite different, demonstrating the sensitivity of surface signature to look angle. This is a feature of $\delta\sigma/\sigma_o$ that would not be readily identifiable in a model that neglects the two-dimensional variation in the current density associated with use of Eq. 1. Additional plots, in other directions (not shown) further confirm the sensitivity of $\delta\sigma/\sigma_o$ to variation in look angle.

Existing SAR imagery (Valenzuela et al., 1985) does reveal strong return from a region of the bank near the location marked by arrows in Fig. 1. The associated data was derived with a "look" angle of 103°, which corresponds to roughly to an east-southeastward direction. This direction is quite close to the eastward look angle that was used in the portion of Fig. 1 in which the large return is found. On the other hand, it should be noted that there is some sensitivity with respect to variations in time in our calculated intensities from this region when the look angle is set to 103°. At the current time, we are investigating the

correlation of this feature with existing imagery more fully as well as including the effect of long wave refraction by the bottom features.

III Conclusion

We have applied the first order relaxation/Bragg model for evaluating surface signatures using realistic current distributions to assess the two-dimensional effect of current field. These currents are solutions of the two dimensional NS equation associated with a barotropic model of the M_2 tidal cycle. Our calculations of surface signatures thus involve a modelling of the tidal cycle that is considerably more complete than in previous investigations, where tidal currents have been treated within a one dimensional framework, and the flow is nearly uniform and laminar free. Our results show 1) that in fact, even in the limit of this first order/relaxation approximation, there is strong correlation between bottom topography and surface signature, and 2) that this correlation is strongly dependent on look angle in relation to current direction. We have found reasonable agreement between variations in surface signature and available imagery data. We have also identified various factors which complicate the relationship between imagery and calculations of surface signature. These include: the possibility of temporal and spatial variations in various quantities that are usually assumed to be constant, associated, for example, with variations in environmental parameters; the importance of non-linear dependencies of radar return on wave height spectrum, as well as the importance of advection terms within the model. Comparisons between our calculations with existing data reveal the need for explicit, two-dimensional information of the current field as a function of space and time.

References

Alpers, W., and I. Hennings, "A theory of the imaging mechanism of underwater bottom topography by real and synthetic aperture radar", J. Geophys. Res., 89, 10529-10546, 1984.

Born, G. H., J. A. Dunne, and D. B. Lame, "Seasat mission overview", Science, 204, 1405-1406, 1979.

de Loor, G. P., Radar methods, in Remote Sensing for Environmental Science, edited by E. Schanda, 147-186, Springer-Verlag, New York, 1976.

de Loor, G. P., "The observation of tidal patterns, currents and bathymetry with SLAR imagery of the sea", IEEE Trans. Geosci. Electron., OE-6, 124-129, 1981.

de Loor, G. P., and H. W. Brunsveld van Hulten, "Microwave measurements over the North Sea", Boundary Layer Meteorol., 13, 113-131, 1978.

Fu, L.-L., and B. Holt, "Seasat views oceans and sea ice with synthetic-aperture radar", Jet Propulsion Laboratory Publ. #81-120, 1982.

Greenberg, D. A., "Modeling the mean barotropic circulation in the Bay of Fundy and Gulf of Maine", J. Phys. Ocean., 13, 886-904, 1983.

Greenberg, D. A., G. R. Valenzuela, and D. T. Chen, "Determining the currents over Phelps Bank', in Radar Scattering from Modulated Ocean Waves by Non-Uniform Currents, edited by G. Komen and W. Oost, Reidel Publ., Dordrecht, Neth., 1989.

Hughes, B. A., "The effect of surface waves on surface wind waves", J. Geophys. Res., 83, 455-465, 1978.

Keller, W. C., and J. W. Wright, "Microwave scattering and the straining of wind-generated waves", Radio Sci., 10, 193-147, 1975.

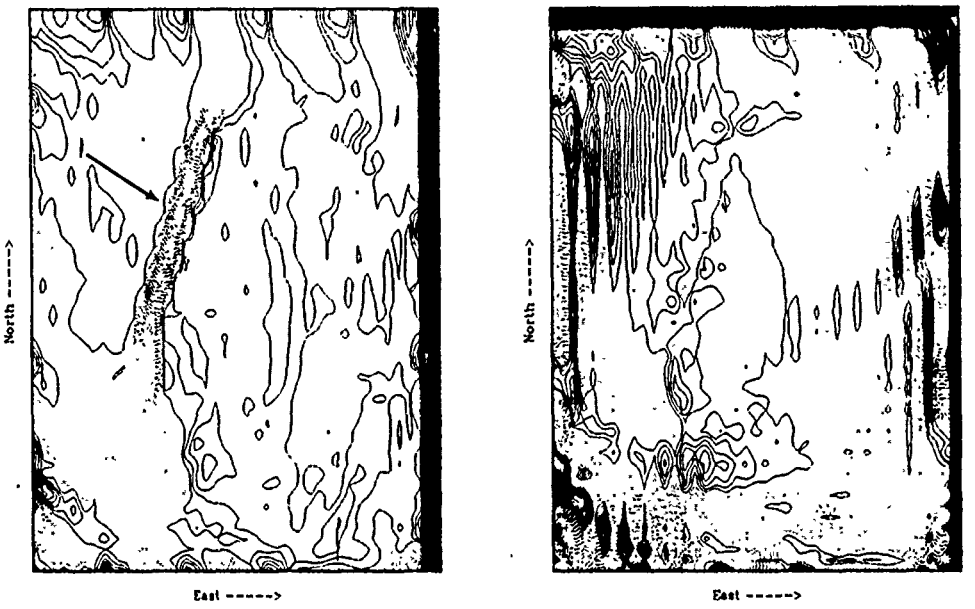
Kenyon, N. H., "Tidal current bedforms investigated by Seasat", in Satellite Microwave Remote Sensing, edited by T. D. Allan, chap. 16, Ellis Horwood Limited, Chichester, England, 1983.

Lodge, D. W. S., "Expressions of bathymetry on Seasat synthetic aperture radar images" in Satellite Microwave Remote Sensing, edited by T. D. Allan, Chap. 15, Ellis Horwood Limited, Chichester, England, 1983a.

Lodge, D. W. S., "Surface expressions of bathymetry on Seasat synthetic aperture radar images", Int. J. Remote Sensing, 4, 639-653, 1983b.

Valenzuela, G. R., W. J. Plant, D. L. Schuler, D. T. Chen, and W. C. Keller, "Microwave probing of shallow water bottom topography in the Nantucket shoals", J. Geophys. Res., 90, 4931-4942, 1985.

Wright, J. W., "Detection of ocean waves by microwave radar; The short gravity-capillary waves", Boundary layer meteorol., 13, 101-102, 1978.


North ------>
East ----->
North ------>
East ----->
North ------>
East ----->

COMPARISONS BETWEEN MEASURED AND THEORETICAL RESULTS FOR C- AND K_U-BAND BACKSCATTER FROM NATURALLY-OCCURRING INTERNAL WAVE CURRENT PATTERNS

T. W. Dawson & B. A. Hughes

Defence Research Establishment Pacific FMO, Victoria, B.C., Canada, V0S 1B0

Abstract

This paper presents a comparison between predicted and measured values of radar backscatter and of internal wave surface currents, based on data obtained during the SCATTMOD internal wave experiment. Radar backscatter and surface truth measurements were obtained during 6 days in August, 1985, and cover 9 sets of tide-generated internal waves in Georgia Strait.

The radar portion of this data consists of approximately 75 sets each of C- and K_u-band fanbeam airborne scatterometer signals, each processed to 25 incidence angles and an along-track resolution of either 12.5m or 6.25m. Aircraft navigation data were also recorded. Simultaneous surface measurements, including wave slope, wave height, current, wind, and ship position, were obtained from CFAV Endeavour Current meters were located both fore and aft to allow internal wave phase velocity estimates to be computed.

A first analysis of the data is based on linear first-order Bragg inversion theory, to provide estimates of surface current from the radar modulation. Correlations between predicted and measured current profiles through the internal wave fields range from negligible to approximately 0.9, with K_u-band predictions generally providing the better correlations. Current amplitudes (but not phases) predicted from C-band data are generally the best, with an error factor ranging from 1 to 15. Current amplitudes predicted from K_u-band data are typically an additional factor of 3 in error. However, the strength of the internal wave currents, and variations in the Bragg surface wave vector due to aircraft attitude variations provide a significant source of error for this inversion technique.

Results from a second analysis, based on a forward modelling of the radar backscatter which accounts for both nonlinearities in the surface wave interaction with the current field, and variations in the Bragg wave vector along the flight path, are also discussed.

Keywords : *internal waves, scatterometer, microwave remote sensing, modelling.*

1 Introduction

In the formation of images of internal wave patterns by radar remote sensing techniques there are two primary sets of interactions that operate. The first occurs between the internal wave current and the surface wind wave field thereby setting up patterns on the ocean surface that are sometimes easily discernible by eye. This interaction has been previously modelled (Hughes, 1978; Thompson *et al.*, 1988) and has been subjected to various experimental tests (Hughes and Grant, 1978; Lewis *et al.*, 1974). The second interaction occurs between the surface waves and the radar field (with some influence still by the internal wave current directly on the radar waves) and it too has been subjected to considerable experimental and theoretical examination (Kasilingham and Shemdin, 1988, Alpers, 1985).

It has been shown that linearized approximations provide reasonably accurate modelling for gravity waves and internal wave current fields that are not too 'steep' (Hughes, 1978). It has also been shown that synthetic aperture radar imaging of internal waves can be modelled within a factor of 2 or 3 using linear interaction dynamics and 1^{st}-order Bragg scattering for L-band wavelengths ($\approx$ 30 cm), and that greater discrepancies appear for shorter wavelengths (Hughes and Hughes, 1987). Recent experiments suggest that more accurate radar backscatter modelling, compared to 1^{st}-order Bragg, improves the correspondence between theory and experiment (Caponi *et al.*, 1988, Shuchman *et al.*, 1988), and that inclusion of non-linear hydrodynamical interaction terms in the modelling also improves the correspondence (Caponi *et al.*, 1988).

The present experiment (SCATTMOD I) was performed in southern Georgia Strait on the West Coast of Canada during the time period 5 Aug - 10 Aug, 1985. Strong natural internal waves are generated in this region during the summer by tidal flow through the connecting inter-island passages, and result in very well-developed surface modulations in the wind-wave field. Remote imagery was obtained with a fan-beam scatterometer mounted in a CV-580 aircraft flown by the Canada Centre for Remote Sensing (CCRS). Two wavelengths were used, C-band ($\approx$ 6 cm) and K_u-band ($\approx$ 2 cm) and these were subsequently processed using synthetic aperture techniques into highly resolved narrow strip imagery (one pixel wide). Simultaneous (but not necessarily coincident) surface measurements were made by equipment mounted on the research ship CFAV Endeavour.

The analysis and results described herein represent an attempt to use measured surface currents to predict the scatterometer modulations, and also to do the reverse problem, that is, to start with the the scatterometer modulations and to predict the currents.

Two questions can be asked of the data set : (i) At these short wavelengths, how accurate is the linear 1^{st}-order Bragg modelling in predicting either the currents or the scatterometer modulations given one or the other ?, and (ii) how accurate can the modelling be made ?. The first question has been answered by using forward and reverse modelling; the second question has not been fully answered, although non-linear fluid dynamical interactions have been included in the forward modelling.

The data are noisy, sparse and subject to strong assumptions in the modelling. The scatterometer traces, although showing surprisingly large modulations ($\approx$ 10 dB) especially at the shorter K_u-band wavelengths, are inherently noisy. This has been dealt with to a certain degree by averaging over several adjacent incidence angle bands, but more noise remains than is desirable. The current measurements are also 'noisy' because the internal waves themselves contain irregularities, and because instrument noise is not completely removed. Some residual instrument motion also occurs within the internal waves due to ship and sensor-mount responses to the current and wind-wave fields. The aircraft and ship measurements were taken along lines that were as nearly the same as possible, but they are, of course, separate, and one was done at low speed and the other at high speed compared to the internal waves. This disparateness has required assuming that propagation and evolution features of the internal wave and wind-wave fields are negligible over time scales of the order of 1/2 to 1 hour, and further that the internal waves are straight-crested and propagate strictly perpendicular to their crests at a common phase speed Both of these are only approximations to reality.

The ambient wind-wave field is assumed to be 'saturated' and modelled by a simple power law description in wavenumber space. Previous measurements (Hughes and Hughes, 1987) indicate that this is also a significant assumption and that it should be allowed for in the interpretation

of the results.

Although the data set is imperfect because of the above reasons, it does represent a unique and valuable set of measurements Very few remote sensing data sets exist with supporting surface truth measurements of almost any kind, particularly at these short radar wavelengths. It is hoped that the imperfections of the data will be outweighed by the significance and originality of the obtained results for this largely untested area of oceanographic remote sensing.

The following sections of this report provide, in order, a description of the data, the basic theory used in the modelling, the data processing, results in the form of regressions, and discussion and conclusions. Errors do exist in estimations of the parameters of the experimental conditions, (e.g. aircraft speed, internal wave propagation speed and direction, wind speed) and these are considered as well.

2 Data Set

The equipment on the ship was essentially the same as that described elsewhere (Hughes *et al*, 1984) A boom-mounted instrument platform ahead of the bow of the ship carried a DREP-constructed laser wave-slope meter, a ('bow') current meter at a nominal 1-m depth, a waveheight wire, thermistors, accelerometers and an anemometer. A second anemometer was mounted on the ship mast, as was a trisponder unit associated with three shore-based remote units at known locations. A second boom positioned off the starboard side at the stern of the ship carried a second ('stern') current meter Data from these instruments were recorded on computer tapes, along with ship's heading information.

During the experiment, the ship made one or more passes through sets of naturally-occurring internal waves under the direction of a helicopter-borne observer There were a total of nine sets of internal waves for which data from simultaneous aircraft overflights are available, some of which were measured during more than one ship pass These data are described in detail elsewhere (Dawson and Hughes, 1989).

Also available during the ship measurements are several hundred time-encoded photographs of the experiment taken by the helicopter observer, at typical altitudes of several hundred metres. The internal waves are clearly visible in many of these photographs, and this helped to identify internal wave crests in the ship data files (see Figure 1). The rest of the data were obtained from the CV-580 aircraft This data set has been discussed elsewhere (van Halsema *et al.*, 1986) It consists of approximately 75 sets of C- and K_u- band along-track fan-beam scatterometer data, along with pertinent aircraft track-related data, obtained during repeated overflights of the internal waves, with a track approximately perpendicular to the internal wave crests.

The data and parameters are summarized elsewhere (Dawson and Hughes, 1989). The data represent a mixture of polarizations and resolutions (nominal 6-m or 12-m pixels), in both C- and K_u- band. The data for these flights was processed for DREP by CCRS from the aircraft-recorded data. The processed files each consist of several hundred records consisting of a time stamp, a time offset from midnight (UT), a calibration signal for each band, ground speed, aircraft altitude, and instantaneous track heading, together with 25 values each of incidence angle, radar look direction (measured relative to the track angle), and scattering coefficients for each of the C- and K_u- signals. Integration of the track data gives an aircraft position relative to some arbitrary origin, which is made precise using a nadir photograph of the ship. Incidence angle and look direction give a relative direction from the aircraft to the center of each pixel in the record, and allow computation of a representative Bragg wave vector for the pixel. These pixels were processed from the aircraft data using a constant footprint or simultaneously constant incidence angle width mode.

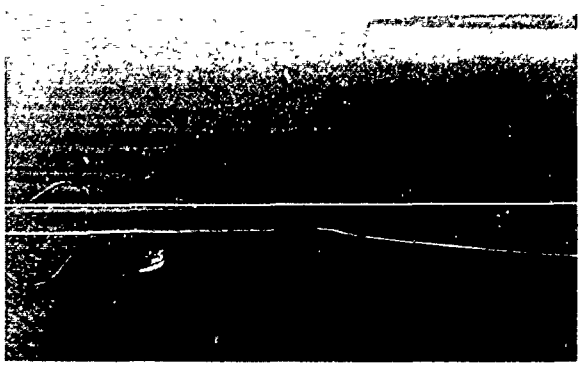


Figure 1: Photograph of CFAV Endeavour commencing the internal wave measurement run for Set #16, taken from the helicopter at 10:29:21 (PDT).

3 Theory

The basic model is defined with respect to a stationary land-based Cartesian reference frame (x,y), where x is East and y is true North. Phenomenologically, a spectrum of surface water waves with wavevector field $\vec{k}(\vec{x},t)$ and ambient action spectrum $\psi_0(\vec{k})$ evolved under the influence of a water current field $\vec{u}(\vec{x},t)$ and wind field $\vec{w}$. The resulting surface wave field was then viewed by the airborne scatterometer, while simultaneous ship-borne measurements of relevant surface parameters were taken. The two sets of data were linked by a nadir-looking photograph of the ship at some point during each overflight. Typical ship measurement runs were of the order of one hour, and the aircraft measurement runs of the order of two minutes.

Parameters measured on the ship which are relevant to this article include time, position $\vec{s}$ with respect to a set of land-based trisponders at known locations, vector water current $\vec{u}$ relative to the ship at bow and stern current meters, and wind velocity with respect to the ship. These data were processed to get estimates of the current with respect to the above land-based coordinate frame along the ship's track, i.e. $\vec{u}[\vec{s}(t),t]$.

For modelling purposes, it was necessary to extrapolate these current and wind measurements to other positions and times, and some assumptions were therefore necessary. The currents of interest during the experiment consisted of tidally-generated internal wave packets, propagating through tidal drift currents. It was assumed that the drift component was a uniform translation $\vec{u}_0$, and that the internal waves were sufficiently long-crested that crest curvature could be ignored, and of sufficiently long wavelength that dispersion could be neglected. Thus the surface current model consists of a two-dimensional longitudinally propagating non-dispersive internal wave packet, plus a uniform ambient drift. A second Cartesian coordinate system (X,Y) was introduced, defined to be at rest with respect to the mean water, and oriented so that the internal wave propagation is along the X-axis at phase speed C. The internal wave current was then *assumed* to be adequately modelled by the functional form

$$\vec{U}(\vec{X},t) = \widehat{X}\Phi(X - Ct), \tag{1}$$

so that the total water current in the (x,y)-frame was

$$\vec{u}(\vec{x},t) = \vec{u}_0 + \widehat{X}\Phi(X - Ct). \tag{2}$$

Coordinates in the two reference frames were related by $\vec{x} = \vec{X} + \vec{u}_0 t$, provided that the origins coincide at $t = 0$ To make things definite, the time origin for each aircraft overflight was taken to be that of the nadir photograph, and the common coordinate origin at this time was taken to be the location of the bow current meter.

The raw data obtained during the experiment were the north and east components of the water velocity relative to the ship, $\vec{u}_r(t) \equiv \vec{u}(t) - \dot{\vec{s}}(t)$ These data were contaminated by internal-wave coupled noise (Dawson and Hughes, 1989), which was removed by fitting separate cubic polynomials to each of the easting and northing measurements, and using these to compute an instantaneous ship velocity estimate. This approximation was used in the above expression to estimate values of $\vec{u}(t)$. The mean drift $\vec{u}_0$ was estimated as the average of the water current with respect to land during each ship measurement period. Correlation methods using the time differences between the two current meter signals were used (Hughes *et al.*, 1984) to estimate a phase propagation speed C for the internal wave part of the current field, and the direction of propagation $\widehat{X}$ was estimated at the time of the experiment and corrected using the nadir photographs provided by CCRS. The internal wave current envelope Φ and the spatial coordinate relative to the moving internal wave, χ, were then estimated with respect to land from the bow current meter signals according to

$$\chi = \widehat{X} \cdot [\vec{s}(t) - \vec{u}_0 t] - Ct, \qquad \Phi(\chi) = \widehat{X} \cdot (\vec{u}[\vec{s}(t),t] - \vec{u}_0). \tag{3}$$

These data were then used to construct a weighted least-squares cubic spline approximation to Φ which was used in solving the differential equations.

The surface-internal wave interaction equations require the wind field $\vec{W}(\vec{X},t) \equiv \vec{w}(\vec{x},t) - \vec{v}_0$ relative to the still ambient water, where $\vec{w}(\vec{x},t)$ is the wind in the land-based (x,y)-frame. The measurements were of the wind relative to the ship at the ship's location, i e. the data file contains samples of $\vec{w}_s(t) \equiv \vec{w}[\vec{s}(t),t] - \vec{s}(t)$. For lack of a better approximation, the wind velocity was assumed to be constant throughout each aircraft run, and obtained from the ship data as a time average over each aircraft overflight period, i.e. $\vec{W} = < \vec{w}_s(t) > + < \vec{s}(t) > - \vec{v}_0$.

During each overflight, the aircraft data available included time, aircraft heading and speed, beam incidence angle and look direction, as well as the associated backscatter coefficients, referenced to the nadir position. Measurements from the nadir photograph, together with the known dimensions of the ship, were used to compute a position vector offset $\vec{\Delta}_0$ of the bow current meter position with respect to the aircraft nadir point at time t_0. From the aircraft velocity information $\vec{v}_p(t)$, the aircraft position $\vec{p}(t)$ was computed by integration.

For each aircraft data sample, the aircraft track angle θ_a, incidence angle θ_i, look direction θ_ℓ and radar wavelength λ_r information allowed computation of a surface wave Bragg vector

$$\vec{k}_B[\vec{p}(t),t] = 4\pi \sin(\theta_i)\{\hat{x}\cos(\theta_\ell + \theta_a) + \hat{y}\sin(\theta_\ell + \theta_a)\}/\lambda_r. \qquad (4)$$

The equations governing the interaction between surface and internal waves are those of Hughes (1978) These are written in the (X,Y)-frame, in which the ambient water is at rest. In Cartesian tensor form (with indices running from 1 to 2, and implied summations over repeated pairs), the evolution of a surface wave vector $\vec{K}$ and associated action spectrum ψ in time and space is governed by

$$\mathcal{L}K_i = -K_j \partial U_j/\partial X_i, \quad \mathcal{L}\psi = \beta\psi(1 - \psi/\psi_0). \qquad (5)$$

Here the differential operator $\mathcal{L}$ is defined by

$$\mathcal{L} \equiv \partial/\partial t + (U_j + \partial\omega/\partial K_j)\,\partial/\partial X_j, \qquad (6)$$

and $\psi_0(\vec{K})$ denotes the value in the absence of any current gradients. Computations in this report are based on an ambient action spectral form

$$\psi_0(\vec{K}) \equiv A_0|\cos[(\theta - \theta_0)/2]|^s K^a[\omega(\vec{K})]^b, \qquad (7)$$

where A_0, s, a, b are constants, and θ and θ_0 respectively denote the polar angles of $\vec{K}$ and $\vec{W}$. The surface wave intrinsic frequency $\omega(\vec{K})$, incorporating both gravity and surface tension effects is

$$\omega(\vec{K}) = \sqrt{g|\vec{K}| + T|\vec{K}|^3}. \qquad (8)$$

If a height spectrum containing a $K^{-\nu}$ behavior is to be used, then $a = -\nu - 1$ and $b = 1$. This uses a factor ω^2/K (which is just the constant g for long surface waves, but which may deviate from g for the short Bragg waves) to convert the height spectrum to an energy spectrum, which is then divided by ω to convert it to an action spectrum For purposes of this report, the 'saturated' ideal values $s = \nu = 4$ were used for lack of measured estimates of the true spectral shape during the experiment. The wind speed enters via (Plant, 1982, omitting wind angle dependence)

$$\beta(\vec{K}) = 0.04|\vec{K}|^2 u_*^2/\omega(\vec{K}), \qquad (9)$$

where u_*^2 is given by (Amorocho and deVries, 1980)

$$u_*^2 = |\vec{W}|^2\left\{0.00104 + 0.0015/\left[1 + \exp[-(|\vec{W}| - 12.5)/1.56]\right]\right\}. \qquad (10)$$

The scattering mechanism was assumed to be first-order Bragg. Thus it was assumed that the backscattered signal was proportional (Wright, 1968) to the integral of the total mean-square height of surface waves at the local Bragg wavenumber, over the instantaneous beam footprint. For a narrow band incident beam, this quantity is approximately

$$I(\vec{x},t) \propto \{\mathcal{H}[\vec{k}_B(\vec{x},t)] + \mathcal{H}[-\vec{k}_B(\vec{x},t)]\}\delta\vec{k}_B,$$

where $\mathcal{H}$ is the local height spectrum in the surface waves, and $\delta\vec{k}_B$ is a measure of the beam bandwidth. The height spectrum can be related to the action spectrum ψ and the intrinsic frequency ω of the surface waves by $\mathcal{H} = k\psi/\omega$ and since ω depends only on $|\vec{K}|$, the scatterometer image can be modelled as

$$I \propto \{\psi(\vec{k}_B) + \psi(-\vec{k}_B)\}\{|\vec{k}_B|/\omega(\vec{k}_B)\}\delta\vec{k}_B. \qquad (11)$$

If the environmental parameters are available in a numerically tractable form, the above equations can be solved either to predict values of $\psi(\vec{x},t))$, thereby simulating a scatterometer image, or inverted in some manner to predict the internal wave current distribution from the scatterometer signal, as will be discussed below.

4 Forward Modelling

This section is concerned with attempts to model the radar image from the surface measurements. Using the method of characteristics, Equations (5) were converted to a set of coupled ordinary differential equations for the values of ψ and $\vec{K}$ along a trajectory $\vec{X}(t)$,

$$\dot{X}_j = U_j + \partial\omega/\partial K_j, \quad \dot{K}_i = -K_j \partial U_j/\partial X_i, \quad \dot{\psi} = \beta\psi(1 - \psi/\psi_0). \qquad (12)$$

However, in this reference frame, the current is given by Equation (1) and has only an X_1 component. The first pair of Equations (12) thus reduce to $\dot{K}_2 = 0$ and

$$\dot{X}_1 = \partial\omega/\partial K_1 + \Phi(X_1 - ct), \quad \dot{X}_2 = \partial\omega/\partial K_2, \quad \dot{K}_1 = -K_1 \partial\Phi(X_1 - ct)/\partial X_1.$$

Therefore K_2 is constant during the integration. It can also be verified from these equations that the quantity $J(t) \equiv (\Phi - C)K_1 + \omega$ is a constant during the integration, and this can serve as an accuracy control function during numerical solution of the characteristic equations. Alternatively, the equation $J(t) = J(t_-)$ (where t_- is an integration starting time) can serve as a solution for K_1 as a function of X_1, although this involves solving a sixth degree polynomial at each point, owing to the functional form Equation (8). Thus in principle, only the equation for X_1 needs to be integrated numerically. The other variable can then be expressed formally as integrals. The solution (Hughes, 1978) for ψ along the characteristic, once $\vec{X}(t), \vec{K}(t)$ are known, can be expressed as

$$\{\psi(t)\}^{-1} = \{\psi_0(t)\}^{-1} - \int_{t_-}^{t} (d\{\psi_0(\eta)\}^{-1}/d\eta) \exp\left(\int_\eta^t \beta(\tau)d\tau\right) d\eta.$$

This form shows that the actual initial conditions are 'forgotten' in a manner depending on $e^{-|\varpi|}$, where $\varpi \equiv \int_{t_-}^{t} \beta(r)dr$.

To compute a synthetic scatterometer trace, the following procedure was used. For each aircraft sample at time t_j, the instantaneous plane position $\vec{p}(t_j)$ was used as the 'initial' position, and an initial Bragg wave vector $\vec{k}_{Bj}$ at this position was computed using Equation (4) These were then expressed in the (X,Y)-frame to get initial vectors $\vec{X}_j$, and $\vec{K}_{Bj}$ at which $\psi(\vec{X}_j, \vec{K}_{Bj}, t_j)$ is desired. The differential equations for $\vec{X}(t), \vec{K}(t)$, augmented by the additional equation $\dot{\varpi} = \beta$ were then integrated from t_j backwards to some earlier time t_- determined by either the trajectory being permanently outside of the support of Φ, or by ϖ having exceeded some specified upper bound, indicating that the actual starting values for the forward problem to determine ψ should be relatively unimportant. The differential equation for ψ was then integrated forward in time from t_- to t_j, using the initial condition $\psi(t_-) = \psi_0[\vec{K}(t_-)]$. The resulting value of $\psi(t_j)$ then gives the expected Bragg contribution to Equation (11). The procedure was then repeated for the anti-Bragg wave vector, and the results expressed as a synthetic image according to Equation (11), the bandwidth term being incorporated into the overall (indeterminate) multiplicative constant.

5 Reverse Modelling

This section is concerned with attempts to estimate the internal wave current from the scatterometer image. The procedure is analogous to that of Hughes and Hughes (1987). Specifically, Equations (5), together with the functional form $\psi_0 = \psi_0(\vec{K})$ can be used to derive the (exact) result

$$(\mathcal{L} + \beta)\mathcal{F} = \beta\{(1 - \psi/\psi_0) + \mathcal{F}\} + \Gamma,$$

where $\mathcal{F} \equiv \log(\psi/\psi_0), \Gamma \equiv (K_j/\psi_0)(\partial\psi_0/\partial K_i)(\partial U_j/\partial X_i)$ If the currents are sufficiently weak, then $\psi \approx \psi_0$, and the current term in Equation (6) can be neglected relative to the group velocity, provided that the surface wave is not directed along the internal wave crests. Moreover the wave vector remains constant in this approximation (Hughes, 1978), and the above equations immediately simplify to

$$\{\partial/\partial t + \beta + (\partial\omega/\partial K_j)\,\partial/\partial X_j\}\,\mathcal{F} = K_X \psi_0^{-1}(\partial\psi_0/\partial K_X)\,\partial\Phi(X - Ct)/\partial X.$$

It is then assumed that the modulation pattern is locked onto the internal wave, so that $\mathcal{F} = \mathcal{F}(X - Ct)$. With $\chi \equiv X - Ct$, this finally reduces to an ordinary differential equation

$$\alpha\mathcal{F}'(\chi) + \beta\mathcal{F}(\chi) = \gamma\Phi'(\chi), \qquad (13)$$

where $\alpha \equiv \partial\omega/\partial K_X - C$, $\gamma \equiv (K_X/\psi_0)(\partial\psi_0/\partial K_X)$ are constants depending on the particular wavevector $\vec{K}$ under consideration.

If Equation (11) is linearized about a mean modulation I_0, then it is readily shown that

$$\mathcal{I} \equiv \log(I/I_0) \approx I/I_0 - 1 = r_+\mathcal{F}_+ + r_-\mathcal{F}_-, \qquad (14)$$

where $r_\pm \equiv \psi_{0\pm}/(\psi_{0+} + \psi_{0-})$, the subscript '+' referring to Bragg and '−' to anti-Bragg contributions Quantities dependent only on $|\vec{K}_B|$ cancel in this expression The logarithm of the scatterometer signal with the mean removed provides a direct estimate of $\mathcal{I}$, from which it is desired to reproduce an estimate of $\Phi(\chi)$.

The inversion algorithm is most easily derived with the help of Fourier transforms. Transformation of Equation (13) written for Bragg and anti-Bragg components leads to the "solutions"

$$\hat{\mathcal{F}}_\pm(\xi) = -i\xi\gamma_\pm\hat{\Phi}(\xi)/(\beta_\pm - i\xi\alpha_\pm), \qquad (15)$$

and application of Equation (14) then leads to an expression

$$\hat{\mathcal{I}}(\xi) = -i\xi\{r_+\gamma_+/(\beta_+ - i\xi\alpha_+) + r_-\gamma_-/(\beta_- - i\xi\alpha_-)\}\hat{\Phi}(\xi). \qquad (16)$$

relating the Fourier-transformed scatterometer data to the current field. Thus the current may be thought of as filtered version of the scatterometer signal $\hat{\Phi}(\xi) = \hat{\mathcal{I}}(\xi)\hat{O}(\xi)$ This is easily implemented (Dawson and Hughes, 1989) in a numerical scheme.

To estimate the current represented by each scatterometer trace the following procedure was used The data were processed (by CCRS) to a set of nominal incidence angles. In several cases, the traces contained a direct return from the ship, which was removed Then, for each incidence angle, an average Bragg-surface wave vector was computed for the entire aircraft pass. The scatterometer data were converted from dB to a zero-mean natural logarithm form, and a fitted cubic polynomial subtracted to remove any long-term trends due to radar and aircraft variations. These data were then used directly in Equation (14) as $\mathcal{I}$, and then processed (filtered) according to Equation (16) to obtain current estimate. The procedure was repeated separately for each nominal incidence angle, resulting in 25 current estimates for each of the C and K_u data sets for each overflight.

6 Results and Discussion

An example of the results of the modelling compared to the measured signals is shown in Figure 2. The bottom panel depicts the internal wave current as a function of distance (m) through the internal wave field, with the amplitude (ms^{-1}) given on the left-hand scale. The current is also depicted in a greyscale plot in a two-dimensional coordinate system (fixed to the internal waves) depicted on the right hand and bottom scales (m). The ship track (marked with x's at 60-s intervals) and the plane track (with ticks at 1-s intervals) are also shown in this two-dimensional frame. The times on the ship track are PDT; those on the aircraft track are in seconds relative to t_0. As explained earlier, the coordinate origin is the nadir position in the pertinent nadir photograph. The top panel shows the predicted scatterometer modulation from the forward modelling for C-band, while the panel below it shows the corresponding measured scatterometer data. The greyscale background shows the modulations from the 25 incidence angles used in each run, with the steepest incidence angles at the top, and the shallowest at the bottom of each panel. The superimposed curves represent variations in the average over all incidence angles, with amplitudes shown on the leftmost scale. Both panels are in a $\log_e$ domain, with the along-track mean removed from each incidence angle line separately. The fourth panel from the bottom contains the predicted K_u− modulations from the forward modelling, with the associated measured data below it, in a similar manner to the C-band data. The reverse modelling results, i.e. predicted current from the scatterometer signals, are presented in the fifth panel from the bottom for C-band, and in the second panel from the bottom for K_u-band Again, the greyscale contains variations predicted from each incidence angle separately, and the average is shown in the superimposed curve. All signals have been bandpass filtered, using a zero-phase, fourth-order Butterworth filter, with the passband for the data of Figure 2 corresponding to wavelengths between 96m and 288m, chosen to bracket most of the significant internal wave energy, as determined from a power spectral analysis. It can be seen that there is a high degree of correlation between the measured and the predicted traces for both forward and reverse modelling, i.e. between the topmost two panels for the C-band scatterometer, between the third and fourth panels from the bottom for the K_u-band scatterometer, and between the first, second and fifth panels from the bottom for the currents The apparent phase shift between current and scatterometer traces arises from the filtering operation described by Equation (16).

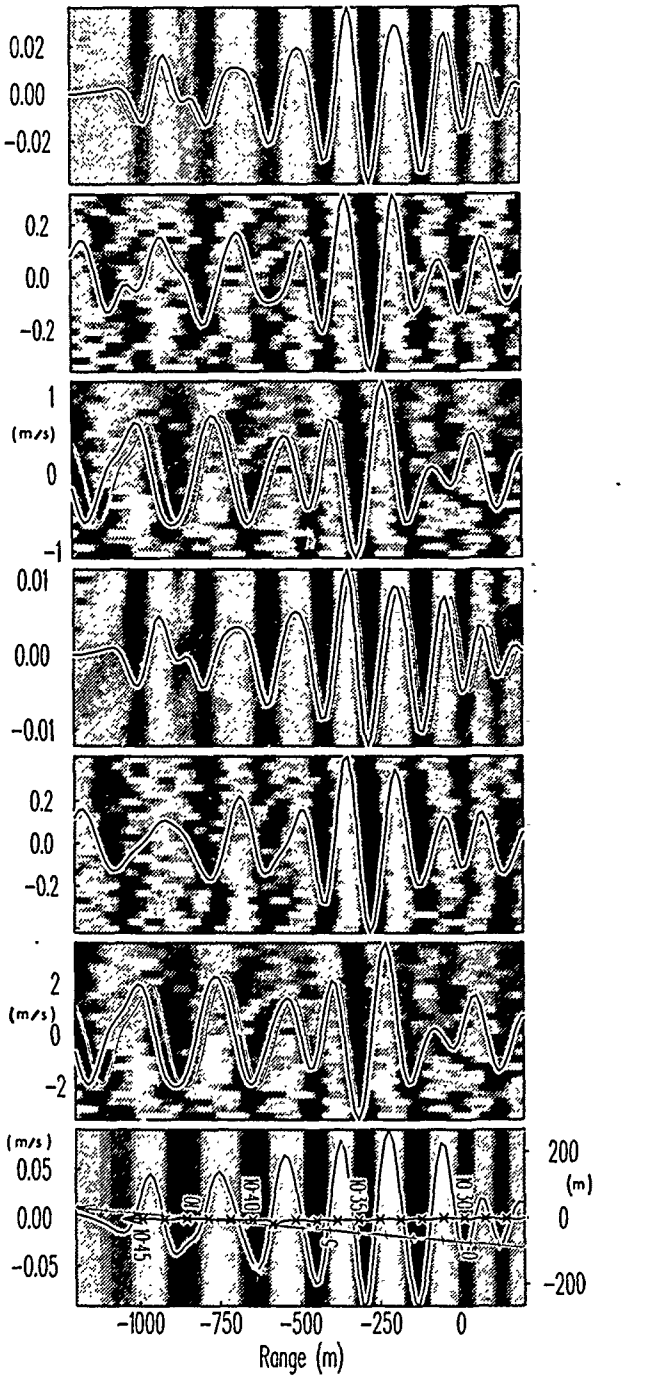


Figure 2 Measured and predicted results from Data Set #16. From top to bottom, the panels contain predicted C-band modulations, measured C-band modulations, predicted current from C-band, predicted K_u-band modulations, measured K_u-band modulations, predicted current from K_u-band, and measured current. Refer to the text for further details.

Table I: Correlation values and RMS ratios. The reverse modelling numbers are in parentheses, preceded by the forward modelling numbers

Data	C-band		K_u-band	
Set	Correlation	RMS Ratio	Correlation	RMS Ratio
16	0.82 (0.80)	8.6 (11.4)	0.88 (0.80)	32.5 (34.3)
17	0.78 (0.78)	13.7 (13.9)	0.77 (0.78)	47.4 (42.9)
18	0.82 (0.77)	11.6 (15.7)	0.82 (0.77)	45.6 (47.4)
20	0.70 (0.55)	7.2 (9.7)	0.76 (0.56)	38.5 (30.0)
28	0.70 (0.68)	1.3 (1.6)	0.73 (0.69)	8.8 (5.0)
29	< 0.70	—	0.71 (0.62)	10.9 (8.0)
47	< 0.70	—	0.72 (0.68)	7.5 (7.5)
55	0.71 (0.57)	1.4 (1.3)	0.80 (0.57)	3.2 (3.8)
56	0.78 (0.70)	1.8 (1.6)	0.82 (0.68)	4.5 (4.5)
63	0.77 (0.72)	6.6 (10.2)	0.75 (0.72)	26.8 (32.5)
64	0.78 (0.71)	8.1 (13.2)	0.74 (0.71)	31.6 (40.2)

A photograph taken from the helicopter at 10:29:21 PDT is shown in Figure 1. The ship (length 67m) is seen commencing a pass through the internal wave set, in accordance with the layout shown in the bottom panel of Figure 2.

Numerical correlation studies were performed on the data from all runs as follows. Each data set was averaged over groups of 5 adjacent incidence angles down to 5 'bins', as well as to a single overall average, forming a sixth bin. Each predicted signal bin (both bands, forward and reverse modelling) was allowed a range of moderate horizontal stretch/shrink factors and shifts to account for any uncertainties in the registration and velocity values. From the entire data collection, 8 C-band sets show correlations ≥ 0.70 with shifts of less than 100 m and stretch factors of less than 10%. The K_u-band data shows correlations ≥ 0.70 for 11 sets (including all of the preceding 8) under the same limits. These correlations refer to the average over all incidence angles. Similar values are obtained from the bin-averages. The curves shown in Figure 2 have correlations of 0.82 at a 4% stretch and shift of 6 m for the C-band forward modelling, and 0.88 for the K_u-band prediction at a shift of 12 m and 2% stretch. These were the highest correlations obtained from the data collection. For the reverse modelling, correlations between predicted and measured current traces for these examples were 0.80 for both the C and K_u cases, both at a shift of 6 m and 4% stretch factor. In all four panels of Figure 2 containing predicted data, the heavy superimposed curve is the result of these correlation-optimized stretch and shift factors. The original unstretched, unshifted curves are visible as the faint traces on these same panels.

Comparisons between the amplitudes of the predicted and measured data have been obtained as ratios of their root-mean-square values. Standard regression coefficients assume errors in the ordinate only and provide a slope value that is simply the rms ratio multiplied by the correlation coefficient. A regression slope-value can be obtained as well with errors in the abscissa only, and the geometric mean of this value and the ordinate-error value is simply the rms ratio. Since errors can exist in both dimensions, the rms ratio seems to be a reasonable estimator. For the data in Figure 2, the ratios are 8.6 and 32.5 for the forward modelling in C- and K_u-band respectively, the scatterometer modulations being underpredicted in both cases. For the reverse modelling, the ratios are 11.4 and 34.3 respectively, with the currents being overpredicted

Values of correlation and rms ratio are given in Table I for each of the 'best' data sets, i.e. those satisfying the above mentioned criteria on correlation, stretch and shift. If the modelling were perfect, these values would all be unity. Deviations in the rms-ratios from unity probably indicate inadequacies in the theory. It can be seen that the ratios range from 1.3 to 15.7 for C-band, and from 3.2 to 47.4 for K_u-band (There is perhaps a slight trend with incidence angle, not shown in Table I, such that steeper angles possess modestly reduced upper values. This might be an indication (again) that longer Bragg waves are modelled better than shorter ones.) In all cases but one (#17), the forward (non-linear) modelling provides higher correlation values than the reverse (linear) modelling. The rms ratios are also better (smaller) for the majority of cases. In those data sets for which the forward modelling was better, the overall reduction in rms ratio is 35% for C-band and 15% for K_u-band.

Numerical experiments have been carried out on data set #16 in an attempt to determine the sensitivity of the rms ratios to the various model parameters. In particular, β was given an angular dependence similar to the underlying ambient wave spectrum, i.e. $\cos^4[(\theta-\theta_0)/2]$. Also the logarithmic slope of the ambient energy spectrum in wavenumber (magnitude) space (i.e. 1+'a' in Equation (7) was varied from -3 to -6. For Data Set #16 the β-variation produces a factor of about 3 improvement in the rms-ratio values for both bands and for both forward and reverse modelling (values in Table I changing from 8.6(11.4) to 3.1(3.8) for C-band, and from 32.5(34.3) to 10.3(11.0) for K_u-band.). Changing 1+'a' to -6 from the 'standard' value of 4 produces very nearly a factor 1.5 improvement in the rms-ratios, again for both bands and both models. There is some evidence (Hughes and Hughes, 1987) that for these short waves, in this location, the slope is steeper than -4 and is more nearly -6 or -7, and so perhaps a factor of 1.5 reduction is permissible for at least some of the data sets. Variations in phase speed and direction of the internal waves were also investigated, the former being changed by $\pm 20\%$ and the latter by $\pm 15°$, each producing changes of approximately 20% in the predicted results. Finally, the wind-speed was varied by $\pm 25\%$, with factor of 2 changes resulting (due to the resultant changes in β). The effects of deviations in the angular portion of the spectrum, i.e. 'b' in Equation (7), have not been examined.

All of these are quite plausible inaccuracies in the model parameters (or variations within a given data set), and so the rms ratios can be expected to show some significant deviations solely due to these effects. The variation in β with angle is very poorly understood. In previous modelling attempts, various angular dependencies have been used, and the two given here tend to represent the extremes (i.e. independent of angle, or the same as the energy spectrum angular dependence). The evidence so far seems to be that allowing an angular dependence similar to that of the energy spectrum provides an improved fit between 1st-order Bragg scattering theory and measurements. As yet, not enough tests have been applied to make this a firm conclusion, although it does support a similar observation reported elsewhere (Hughes and Hughes, 1987).

The rms-ratio values reported here are all larger than unity. This is somewhat different from previously reported (Hughes and Hughes, 1987). values from this same area (but taken under different circumstances). The reason for this difference is not understood as yet, although it could be due to different current strengths for the present data collection. It could also be due to different amounts of interaction between the surface wave components of disparate length scales producing excessive modulations in the real field compared to the predicted field. One data set in the above reference (JOWIP 12-3) showed rms ratios larger than unity by a significant amount, and it also exhibited strong deviations in the surface wave spectrum from the ambient. Perhaps the present collection is exhibiting the same features. Specular scattering has also been observed in other data sets for the steeper incidence angles (Kwoh *et al.*, 1988), and this, if occurring here, could lead to underestimations in the predictions of the scatterometer modulations because it is not included in the present model.

Other reported comparisons between predicted and measured modulations show the same trends as the present data, i.e. underpredictions in the SAR modelling at X-band (Caponi *et al.*, 1988; Shuchman *et al.*, 1988; Gasparovic *et al.*, 1988).

7 Conclusions

Using 1st-order Bragg scattering theory only, forward (non-linear hydrodynamic) modelling of scatterometer traces from measured currents and reverse (linear hydrodynamic) modelling of currents from measured scatterometer data both provide many predictions that correlate well with the associated measurements. Eight data sets in C-band and eleven sets at K_u-band show correlations of at least 0.70. The corresponding amplitude predictions are typically a factor of 3-10 in error for C-band, and even further in error for K_u-band data, with the scatterometer modulations being consistently underpredicted, and the currents overpredicted Some of the discrepancy may be due to inaccuracies in the measured parameters or to unsuitability in the functional forms chosen for the ambient wind wave spectrum and the relaxation parameter.

The 1st-order Bragg theory is certainly inadequate for C- and K_u-band modulation estimates. It also appears that more accurate measurements of the underlying parameters are necessary in order to resolve the amount of inadequacy. Inaccuracy in the parameters can easily account for factors of 3 or more in the resulting amplitude predictions.

Forward (non-linear) modelling provides a better correspondence with measurements than the reverse (linear) modelling.

References

[1] W. Alpers. Theory of radar imaging of internal waves. *Nature*, 314:245–246, 1985.

[2] J. Amorocho and J.J. deVries. A new evaluation of the wind stress coefficient over water surfaces. *J. Geophys. Res.*, 85(C1):433–422, 1980.

[3] E.A. Caponi, D.R. Crawford, H.C. Yuen, and P.G. Saffman. Modulation of radar backscatter from the ocean by a variable surface current. *J. Geophys. Res.*, 93(C10):12249–12263, 1988.

[4] T.W. Dawson and B.A. Hughes. *SCATTMOD I Summary Report. Data Description and Comparisons Between Measured and Theoretical Results for C- and K_u-band Backscatter From Naturally-Occurring Internal Wave Current Patterns.* Technical Memorandum (in preparation), Defence Research Establishment Pacific, Forces Mail Office, Victoria, B.C., Canada, V0S 1B0, 1989.

[5] R.F. Gasparovic, J.R. Apel, and E.S. Kasischke. An overview of the SAR internal wave signature experiment. *J. Geophys. Res.*, 93(C10):12304–12316, 1988.

[6] B.A. Hughes. The effect of internal waves on surface wind waves. 2. theoretical analysis. *J. Geophys. Res.*, 83(C1):455–465, 1978.

[7] B.A. Hughes and H.L. Grant. The effect of internal waves on surface wind waves. 1. experimental measurements. *J. Geophys. Res.*, 83(C1):443–454, 1978.

[8] B.A. Hughes, S.J. Hughes, and T.W. Dawson. *Objectives, Procedures and Description of DREP Data for the Georgia Strait Experiment (July/August 1983).* Technical Memorandum 84-7, Defence Research Establishment Pacific, Forces Mail Office, Victoria, B.C., Canada, V0S 1B0, 1984.

[9] S.J. Hughes and B.A. Hughes. *Estimation of Internal Wave Surface Currents From SAR Imagery.* Technical Memorandum 87-1, Defence Research Establishment Pacific, Forces Mail Office, Victoria, B.C., Canada, V0S 1B0, 1987.

[10] D.P. Kasilingham and O.H. Shemdin. Theory for synthetic aperture radar imaging of the ocean surface : with application to the tower ocean wave and radar dependence experiment on focus, resolution and wave height spectra. *J. Geophys. Res.*, 93(C11):13837–13848, 1988.

[11] D.S.W. Kwoh, B.M. Lake, and H. Rungaldier. Microwave scattering from internal wave modulated surface waves: a shipboard real aperture coherent radar study in the Georgia Strait experiment. *J. Geophys. Res.*, 93(C10):12235–12248, 1988.

[12] J.E. Lewis, B.M. Lake, and D.R.S. Ko. On the interaction of internal waves and surface gravity waves. *J. Fluid Mech.*, 63(4):773–800, 1974.

[13] W.J. Plant. A relationship between wind stress and wave slope. *J. Geophys. Res.*, 87(C3):1961–1967, 1982.

[14] R.A. Shuchman, D.R. Lyzenga, B.M. Lake, B.A. Hughes, R.F. Gasparovic, and E.S. Kasischke. Comparison of joint Canada-U.S. ocean wave investigation project synthetic perature radar data with internal wave observations and modelling results. *J. Geophys. Res.*, 93(C10):12283–12291, 1988.

[15] D.R. Thompson, B.L. Gotwols, and R.E. Sterner II. A comparison of measured surface spectral modulations with predictions from a wave current interaction model. *J. Geophys. Res.*, 93(C10):12339–12343, 1988.

[16] D. van Halsema, A.L. Gray, S.J. Hughes, and B.A. Hughes. *C-* and *K_u*-band scatterometer results from the SCATTMOD internal wave experiment. In *Proceedings of IGARSS'86 Symposium, Zürich 8–11 Sept., 1986*, pages 311–317, ESA Publications Division, August 1986.

[17] John W. Wright. A new model for sea clutter. *IEEE Trans. Antennas and Propagation*, AP-16(2):217–223, 1968.

SURFACE CURRENT AND WAVE MODULATION MEASUREMENTS OVER BOTTOM TOPOGRAPHY - A COMPARISON WITH THEORY

by Peter Smith and Donald Johnson

Naval Ocean Research and Development Activity
Stennis Space Center, MS 39529

ABSTRACT

The wave-current interaction model of Hughes (1978) is applied to the case of flow over bottom topography. The surface current is prescribed using data from four surface-following current meters positioned over a bottom feature on Nantucket Shoals. The ambient wave spectrum is prescribed using data collected by a drifting spar buoy. The current profiles across the submerged sill exhibits minima in both components of current. The dominating wave growth mechanism is the refraction of waves near the minimum in the along-isobath component of flow. The spikey nature of the wave modulation dictates that finite amplitude together with wave-wave interaction terms should be included in the action transport equation.

KEY WORDS: SAR, BOTTOM TOPOGRAPHY, WAVE-CURRENT INTERACTION

EXPERIMENT

In September, 1987 we deployed a series of four current meters across a major bottom topographic feature on the Nantucket Shoals. An east-west bottom profile as determined by a shipboard acoustic depth-sounder appears in figure 1. The bottom rises rapidly from 20 m depth to 7 m depth at the top of the sill. The slope is about 5° at the steepest point. The current meters measure the surface current speed and direction and are described in Johnson (1987). A spar buoy equipped with an array of capacitance wires (Smith and Johnson, 1988) was deployed up-stream of the array and drifted through the array and across the ridge as pictured in figure 2. At the time of this measurement the current was ebbing and exhibited a strong along-trough component (denoted by v) of 1 m/s and a weaker across-sill component (denoted by u) of 20-30 cm/s. The entire across-sill distance under study was 400 m. The arrows in figure 2 represent the speed and direction of the surface current at the respective location and show the speed to decrease going from deep water to shallow water. The spar buoy collected wave height data during its drift, and that time series is displayed in figure 3 together with an expanded view of the bottom topography. The amplitude of the waves exhibit a sudden increase at the crest of the ridge.

ANALYSIS

We assume that the bottom topography is two dimensional and all derivatives with respect to the along-trough direction (denoted by the y direction) are assumed to be zero. We also assume that outside

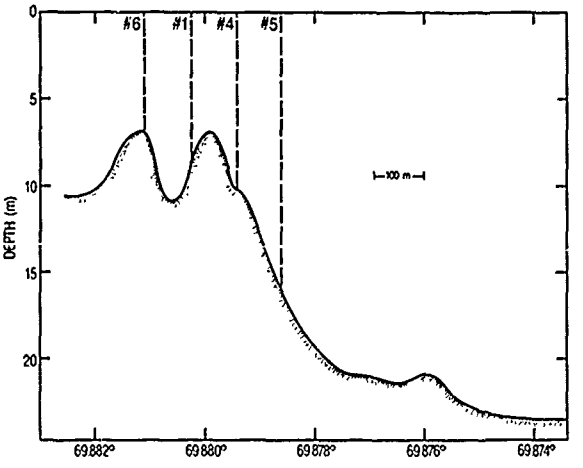


FIGURE 1

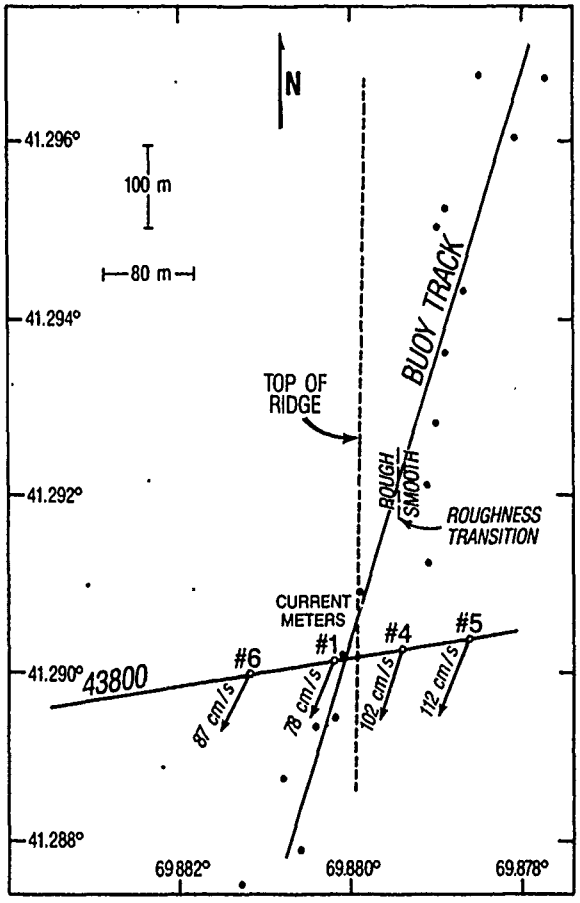


FIGURE 2

of the region spanned by the current meters the current strain is nearly zero. This defines a 240 m "interaction region" which modulates the ambient wave field. In order to interpolate the current between the data points we fit a cubic spline to the data and force the derivatives of both u and v to be zero outside of the interaction region. A plot of the resulting u and v profiles are shown at the bottom of figure 5. The u component exhibits a minimum near the crest of the sill and the v component exhibits a minimum near the trough of the sand wave.

The wave and current interaction theory of Hughes (1978) is applied to the ambient wave height spectra as measured by the spar buoy and the current velocity contours as inferred from the current meter data. The relevant equations are:

$$\frac{dx}{dt} = \frac{\partial \omega_o}{\partial k_x} + u \qquad (1)$$

$$\frac{dy}{dt} = \frac{\partial \omega_o}{\partial k_y} + v \qquad (2)$$

$$\frac{dk_x}{dt} = -k_x \frac{\partial u}{\partial x} - k_y \frac{\partial v}{\partial x} \qquad (3)$$

$$\frac{dk_y}{dt} = 0 \qquad (4)$$

and $$\frac{d\Psi}{dt} = \beta(k_x, k_y)[\Psi - \Psi^2/\Psi_o]. \qquad (5)$$

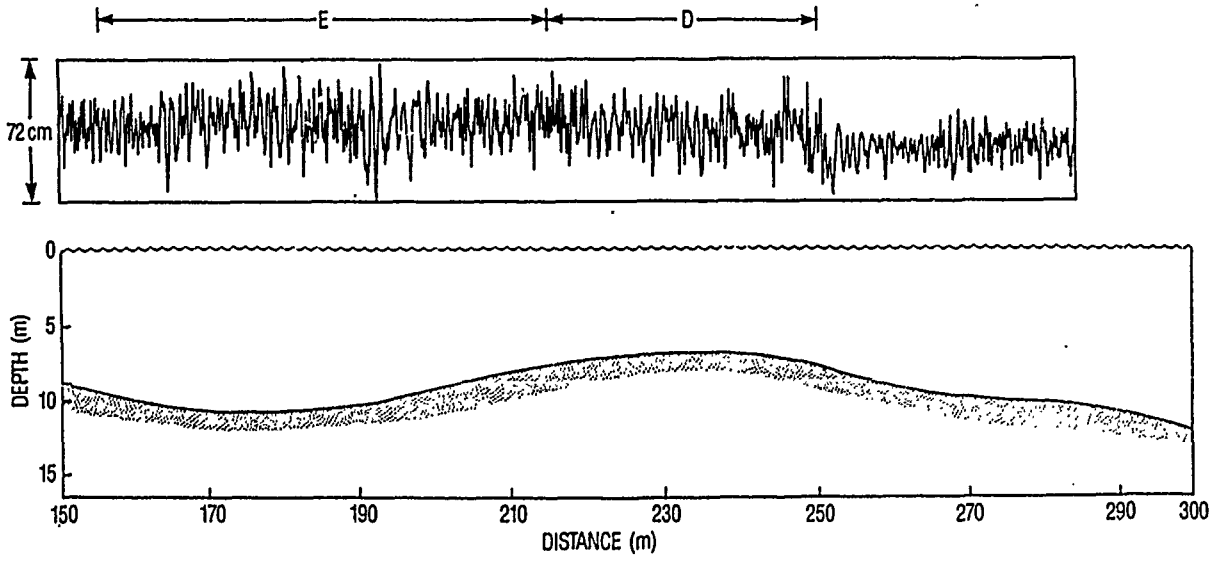


FIGURE 3

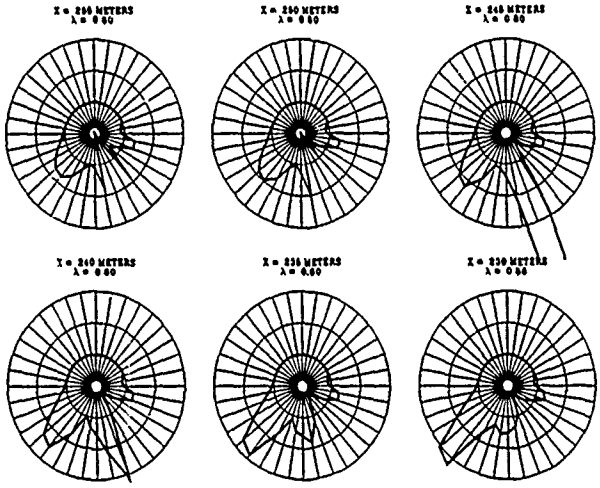


FIGURE 4

Ψ is the wave action spectral density, Ψ_0 is the equilibrium spectral density evaluated at the point, k is the wave number, and ω is the radian frequency. The ray equations (equations 1-3) are integrated backwards from observation points within the interaction region to determine the ray trajectories. Observation points are spaced 5 m apart. The form of these equations are identical to those used by Hughes (1978) and by Thompson et al. (1988) to describe the modulation by internal waves except that the term, $k_y \frac{\partial v}{\partial x}$, is added to equation 3 to include the effect of lateral shear. It will be seen below that the inclusion of this term makes an important difference in the results. Equation 5 leads to an integral equation for Ψ following Hughes (1978). This integral equation is integrated forward into the interaction region along the rays to provide the directional modulation functions, H(θ) at each observation point. ($\Psi = H\Psi_0$) Six such H(θ)'s are shown in figure 4 for the case of λ = 80 cm. The directional spectrum, Ψ, at each observation point is then integrated in θ, and the results for each wavelength, λ, are plotted in figure 5.

Figure 5 shows that the wave energy is amplified near the minimum in v at all wavelengths. The principal effect of this minimum is to refract wave trajectories. Some trajectories enter the interaction region and get trapped by this minimum and give a large contribution to the energy due to the length of time the wave spends in the environment of the shear. For 60cm<λ<90cm the minimum in u produces an amplification of the wave amplitude, however the effect is not as great as is

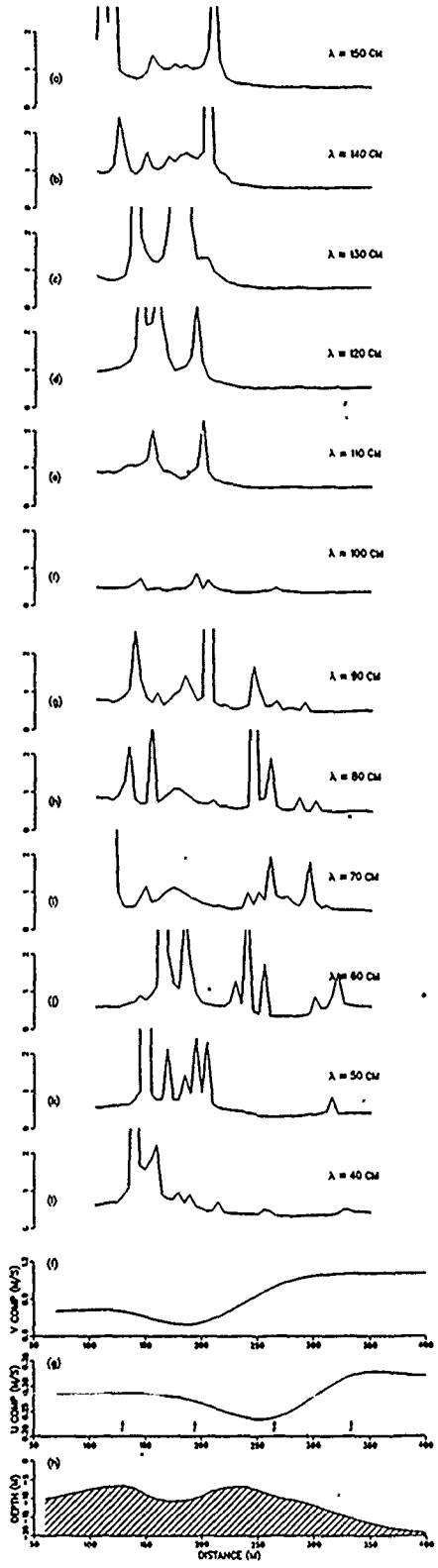


FIGURE 5

the refraction effect. The two regimes, refraction and surface divergence, transition in the neighborhood of x= 210m. These two regimes are labeled E and D, respectively, in figure 3. Power spectra for the respective regions, calculated from the spar buoy data, are shown overlayed in figure 6. The spectra are similar within the 95% condidence limits in spite of the fact that the theory (figure 5) predicts that the "E" spectrum should be dominant for λ>1m and λ<50 cm. A possible explanation for this disparity is discussed below.

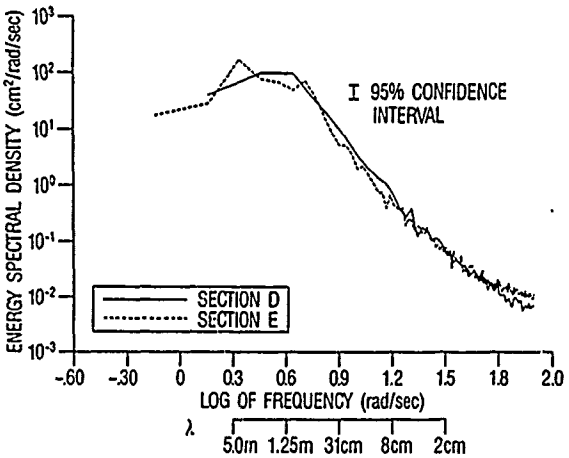


FIGURE 6

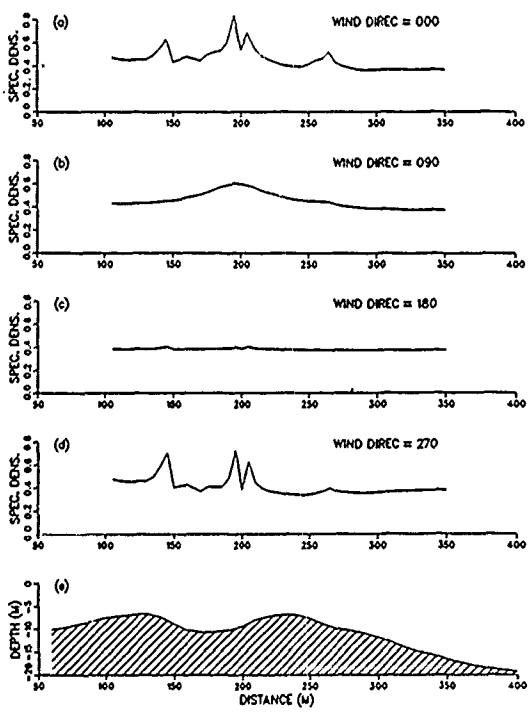


FIGURE 7

The effect of wind direction on the interaction was determined by changing the directional form factor for the ambient wave spectrum. The results for λ=1m and for four points of the compass are shown in figure 7. Modulation of the waves is significant for winds blowing from the west and north, however easterlies and southerlies produce very little modulation.

CONCLUSIONS

Application of the theory of Hughes to the case of wave modulation by bottom topography results in a very spikey signal both in space and in direction as a result of the resonance-like behavior of the blocking phenomenon. The predicted amplitudes are so much greater than the ambient spectral levels that the small perturbation assumption inherent in the form of equation 4 are likely violated. The incorporation of a wave-wave interaction term would redistribute the wave energy both directionally and spectrally. The similarity of the two spectra of figure 6 suggests that such a redistribution has occurred.

The specification of the surface current profile is marginal using the four current measurements as described here if the current does exhibit minima as we have asserted. In the future, more moorings spaced at closer intervals should be employed.

Wave refraction appears to be dominant, according to the model, due to the strong lateral shear in the along-trough component of flow.

ACKNOWLEDGEMENTS

Contribution No. PR 89:031:321. This document has been reviewed and is approved for public release.

REFERENCES

Hughes, B.A., 1978, "The effect of internal waves on surface wind waves 2. theoretical analysis," J. Geophys. Res., Vol 83, pp 455-465.

Johnson, D.R., 1987, "A surface current meter," Continental Shelf Research, Vol 7, pp 975-986.

Smith, P. and D. Johnson, 1988, "Some characteristics of short ocean waves as microwave scatterers," Proc., IGARSS '88 Symposium, pp 395-396.

Thompson, D.R., B.L. Gotwols, and R.E. Sterner II, 1988, "A comparison of measured surface wave spectral modulations with predictions from a wave-current interaction model," J. Geophys. Res., Vol 93, pp 12,339-12-343.

SAR OCEAN SURFACE FEATURE ANALYSIS IN CONCEPTION BAY, NEWFOUNDLAND

Keith A. Thomson, Scott A. Akenhead[1] and Brian G. Sanderson[2]
Seaconsult Limited
Suite 301, 187 Gower Street
St. John's, NF
A1C 1R2
Telephone: (709) 722-7023
Facsimile: (709) 722-7353

SAR imagery was collected in Conception Bay, Newfoundland for the purpose of examining the tidal and wind-driven surface and thermal features in order to enhance a study of the oceanographic influences on the distribution and movement of Northern Cod (Gadus morhua) in the bay.

Two sets of SAR overflights were conducted by the Canadian Centre for Remote Sensing. The first, on September 4, 1988, coincided with the maximum currents of a flood tide and high northwesterly winds. The second, on September 5, 1988, coincided with the maximum currents of an ebb tide and moderate southeasterly winds. The X and C-band SAR imagery revealed very different surface structures on each day. On September 4, the effects of the high wind dominated that of the flood tide. The turbulence created by the wind on the lee side of the islands was very apparent, as were several surface wind-mixing fronts. The surface structures shown in the SAR data on September 5 were due to the effects of the ebb tide, which were sufficient to dominate the effects of the moderate wind. The circulation and thermal patterns included signatures of coastal upwelling events, closed eddies, hammerhead circulation features and tidal jets.

An extensive ground-truth data collection program was carried out coincident with the SAR overflights. These data include: Eulerian current profiles, Lagrangian surface currents and surface temperatures obtained with an acoustic doppler current profiler and drogues in a sub-area of the bay; Eulerian currents at three depths, temperature profiles and tidal heights at each of two moorings at either end of the bay; surface wind speed and direction from a meteorological station established on an island in the bay; and several surface signatures of local ferries as they crossed a tidal jet.

This paper will present the SAR imagery and discuss the observed surface features in light of the ground-truth data. The results will be further discussed with respect to the study relating the local oceanographic phenomena to the distribution and movements of inshore cod.

---

[1]Department of Fisheries and Oceans, Box 5667, St. John's, NF, A1C 5X1
[2]Physics Department, Memorial University of Newfoundland, St. John's, NF, A1C 5S7

AIRBORNE SAR CHARACTERISTICS OF ARCTIC ICE SHELVES AND MULTIYEAR LANDFAST SEA ICE, AND THE DETECTION OF MASSIVE ICE CALVINGS AND ICE ISLANDS

Martin O. Jeffries
William M. Sackinger

Geophysical Institute
University of Alaska Fairbanks
Fairbanks, AK 99775-0800

ABSTRACT

Mid-winter, airborne, X-band SAR image characteristics of the pack ice, multiyear landfast sea ice (MLSI) and ice shelves of northern Ellesmere Island, and the near-real-time detection of a massive calving from the Milne Ice Shelf, N.W.T., Canada are described. Also, the utility of SAR for the detection and characterization of ice islands (tabular icebergs) in the Arctic Ocean is demonstrated. A distinctive SAR characteristic of the ice shelves, ice islands and the MLSI is the ribbed texture resulting from their undulating surfaces. The SAR detects the ice shelf and ice island undulations regardless of their orientation relative to the illumination direction, but the smaller-scale MLSI undulations show only when they are oriented normal to the illumination direction. Strong returns from the inner parts of the northern Ellesmere Island fiords are believed to be from multiyear lake ice and further evidence of year-round, freshwater-seawater stratification related to inverted dams of floating ice shelves and MLSI.

Key words: Airborne SAR; landfast sea ice; ice shelves; ice islands.

INTRODUCTION

In February 1988 extensive airborne SAR data was obtained of the inter-island channels and the Arctic Ocean margin of the Canadian Arctic Archipelago, for the Canadian Arctic Marine Ice Atlas. The data includes the first non-proprietary SAR imagery of the fast ice fringe off the north coast of Ellesmere Island, and of the ice islands (tabular icebergs) that calved in 1982-83 from the Ward Hunt Ice Shelf (Fig. 1). With the possible exception of thick, old sea ice in the fiords of northern Greenland, the fast ice fringe of northern Ellesmere Island is unique because it includes the most prominent known ice shelves (40-100 m thick) and the most extensive areas of multiyear landfast sea ice (MLSI: 3-10 m thick and 2-25 years old) in the Arctic (cf. Jeffries, 1987).

In addition to their uniqueness, the ice shelves and the MLSI are significant because they are the source of tabular icebergs in the Arctic Ocean (commonly called ice islands), and thick sea ice floes (TSIFs), which periodically calve from the ice shelves and MLSI respectively. Ice islands and TSIFs are the largest ice features in the Arctic Ocean, and once they have calved they can drift around the ocean, often for decades, embedded in the pack ice. Their drift takes them into the coastal waters of the Beaufort and Chukchi Seas, and occasionally into the inter-island channels of the Canadian Arctic Archipelago, where they are a hazard to offshore petroleum development and shipping.

The object of this paper is to describe and explain some of the SAR characteristics of the ice shelves and the MLSI, and to demonstrate the utility of SAR for the detection of ice calvings and ice islands.

IMAGE ACQUISITION AND FIELD SURVEY

The 1:300,000 scale SAR data was obtained from an altitude of 9,450 m using the STAR-2 airborne SAR (X-band, HH polarization) operated by Intera Technologies Ltd. On 19 February 1988 a group of ice islands was imaged at approximately 79.98° N, 106.64° W, about 80 km NNW of Ellef Ringnes Island (Fig. 1). On 22 February the fast ice and pack ice along the north coast of Ellesmere Island (Fig. 1) was imaged.

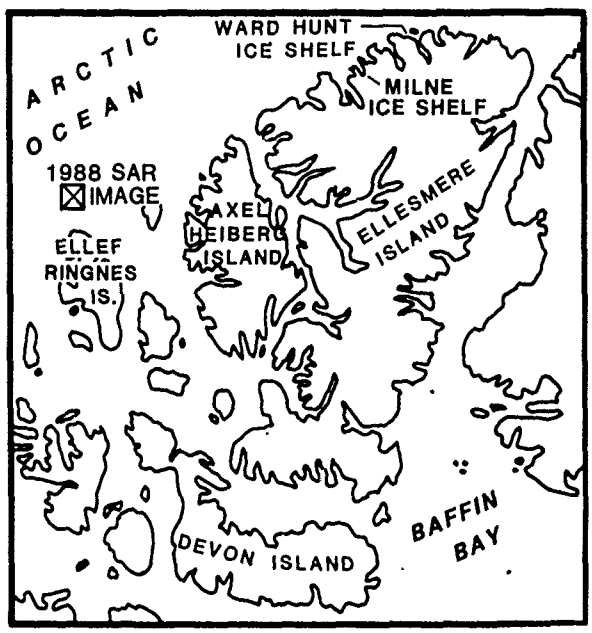


Figure 1. Map of the Queen Elizabeth Islands, Canadian High Arctic, showing the location of the Milne and Ward Hunt Ice Shelves on the north coast of Ellesmere Island, and the ice islands imaged by SAR in February, 1988, NNW of Ellef Ringnes Island.

The interpretation of the SAR imagery is supplemented by observations and measurements made each spring from 1982-86 on northern Ellesmere Island, and in fall 1985 and spring 1986-87 at the ice islands noted above. Fieldwork included observations of ice types and morphology, and analysis of ice physical-structural properties.

REGIONAL ICE CONDITIONS, FEBRUARY 1988

A representative section of the SAR data is reproduced to illustrate the regional pack and fast ice conditions at the time of the SAR data acquisition (Fig. 2). The grey tones of the pack ice returns are quite variable, from extremely bright to very dark (Figure 2). There are many rounded multiyear floes with relatively weak returns or mid-grey tones and they are generally tightly packed with 10/10 concentration. The bright tones represent strong returns from ice rubble in pressure ridges around the floes, and their abundance is

Figure 2. Annotated SAR image of the Arctic Ocean pack ice and fast ice along the north coast of Ellesmere Island. 1: Milne Ice Shelf (see also Fig. 3); 2: Ayles Ice Shelf; 3: Ward Hunt Ice Shelf; 3a: East Ward Hunt Ice Shelf (the source of the ice islands in Fig. 4); YB; Yelverton Bay; MF: Milne Fiord; AF: Ayles Fiord; MI: M'Clintock Inlet. The distance along the bottom edge of the image is 170 km. The radar illumination direction is from the top to the bottom edge.

evidence of the intense, wide-scale ice ridging that occurs in this region due to the almost continuous ice motion towards the coast. There are, however, numerous, extensive refrozen leads and polynyas, with very dark tones (weak returns), which suggest a recent, major reversal of the normal ice motion. The presence of ice on the leads and polynyas is indicated by the numerous narrow, strong returns running through the dark tones, particularly in the large polynya in the centre of the image (Fig. 2). The strong returns are probably backscatter from small ice deformation features.

The fast ice includes the Milne, Ayles and Ward Hunt Ice Shelves, plus MLSI, notably in Yelverton Bay, Ayles Fiord, and M'Clintock Inlet (Fig.2). The MLSI has a fairly dark tone because it is relatively saline and, therefore, a high-loss material; however, the MLSI also includes strong returns from pressure ridges (e.g. Yelverton Bay) and from embedded icebergs (e.g. Ayles Fiord). The ice shelf returns are generally stronger than the MLSI returns and, even at this scale, a distinctive ribbed texture is visible. At the surface most of the shelf ice is a type of freshwater ice, with a negligible salinity and, therefore, a lower-loss material than the MLSI. In the innermost parts of the fiords, at the far range of the SAR, there are some especially bright, textureless returns, particularly in Milne Fiord, Ayles Fiord and M'Clintock Inlet (Fig. 2).

DETECTION OF AN ICE CALVING AT MILNE ICE SHELF

The Milne Ice Shelf (Fig. 3) is the second largest ice shelf (area 290 km^2) remaining in this region. In the mid-1960s an ice island broke off the north-west front of the ice shelf and was subsequently replaced by the growth of MLSI in an area known as the Milne Re-entrant (Jeffries, 1986a; Jeffries et al., 1987). In addition to the re-entrant there was already a narrow zone of MLSI along the north-east front of the ice shelf and, together, they formed a continuous fringe of MLSI (Jeffries et al., 1987: Fig 2b). In the SAR image the MLSI is the relatively dark-toned border with a ribbed texture, but clearly it is no longer continuous (Fig. 3). The north-east fringe remains attached to the ice shelf, but it is now separated from the re-entrant by a roughly S-shaped, dark-toned feature. This, in turn, separates the re-entrant from the ice shelf. The SAR image shows that the Milne Re-entrant has been detached from the Milne Ice Shelf and displaced about 1.7 km to the south-west. The new TSIF has dimensions of 3.6 x 7.2 km.

The S-shaped feature is a recently refrozen lead. The evidence for this is the narrow, castellated, strong return running through the dark tone. The castellations strongly resemble finger-rafting which occurs in young sea ice up to 0.3 m thick The strong return probably is due to backscatter from the margins of the elevated ice

Figure 3. Annotated SAR image of the Milne Ice Shelf. MR: Milne Re-entrant; I: Inner unit; C: Central unit, O: Outer unit. Arrows indicate strong returns from linear moraines. The distance along the bottom edge is 20 km. The radar illumination direction is from the top to the bottom edge.

thrusts or fingers at the surface of the young ice. The finger rafting and the inferred ice thickness are also evidence that the calving occurred only a short time before the SAR flight. It has already been noted that the pack ice in this region is almost always moving towards the north coast of Ellesmere Island, particularly in mid-winter, and that the leads and polynyas (Fig. 2) are evidence of a major reversal of ice motion. Reversals of this kind are due to intervals of persistent offshore geostrophic winds, which can cause not only extensive leads in the pack ice, but which may also contribute to calvings from the fast ice (Ahlnås and Sackinger, 1988). A search through the February 1988, High Arctic weather maps revealed a four day period of offshore winds of up to 10 m s^{-1} between 12 and 15 February, i.e. only 7-10 days before the SAR overflight (Jeffries and Sackinger, unpublished MS[a]). Thus, the calving of the Milne Re-entrant was fortuitously detected in near-real-time.

MILNE ICE SHELF SAR CHARACTERISTICS

On the basis of ground observations and air photographic records of surface topographic variations, the ice shelf has been divided into three units; an inner, a central and an outer unit (Jeffries, 1986b). The inner unit in the SAR image is the bright area at the base of Fig. 3. Above this are the central and outer units, with the ribbed texture mentioned in the context of Fig. 2. As noted in the previous section the MLSI also has a ribbed texture, which is only visible at the larger scale (Fig. 3).

The ribbed texture is related to the undulating topography of linear hummocks and depressions (commonly called "rolls") which are characteristic features of the ice shelves and the MLSI. On the central and outer units of the Milne Ice Shelf the depressions are as much as 7.5 m deep and the crest-to-crest spacing of the hummocks varies between 135 m and 450 m (Jeffries, 1986b). On the MLSI the distance between the hummocks is shorter (60-100 m) and the depressions are shallower (max. 1m) than on the ice shelf. The morphological variations are clearly visible in the SAR image (Fig. 3). Both the ice shelf and the MLSI undulations are oriented normal to the radar illumination direction.

Cutting across the general trend of the rolls on the outer unit are two features that form an inverted T-shape (Fig. 3). These, too, are depressions, and believed to be the traces of former, deep fractures that have since rehealed (Jeffries, 1986b; Jeffries et al., 1987). Although the undulations on the outer unit are not completely straight in places, they are essentially linear features. In contrast, the rolls on the central unit have a more pronounced curvi-linear appearance (Fig. 3). This is believed to be derived from the flow pattern of the glaciers which, although they are presently quiescent, once actively fed the central unit of this ice shelf (Jeffries, 1986b; 1988). Many moraines remain on the surface of the central unit as additional evidence of former glacier activity (Jeffries, 1986b) and some of these, too, are evident as bright returns in the SAR image (Fig. 3).

The inner unit also has an undulating topography, with depressions of up to 1 m depth and a wavelength of 60-100 m (Jeffries, 1986b). However, they are not visible in the SAR image and, instead, as noted in a previous section, there is a uniformly bright tone.

ICE ISLAND SAR CHARACTERISTICS

Unless there has been another calving since February 1988, the loss of the Milne Re-entrant is the largest calving to have occurred since 1982-83, when a number of ice islands broke off the East Ward Hunt Ice Shelf (Jeffries and Serson, 1983). These ice islands are shown in Fig. 4. There are twelve ice islands in the SAR image and their dimensions range from 0.15 x 0.25 km to 5.7 x 8.7 km. The latter is Hobson's Choice (#1: Fig. 4), the largest known ice island in the Arctic Ocean today. Some general ice island characteristics are (Jeffries and Sackinger, unpublished MS[b]): 1) not unexpectedly, they have a ribbed texture; 2) they tend to be long and thin, more angular and sometimes larger than the surrounding sea ice floes; 3) the edges are often bright and well defined.

A number of the ice islands in Fig. 4 comprise shelf ice only, with the distinctive ribbed texture. However, as Fig. 3 shows, MLSI can become attached to the fronts of ice shelves and, consequently, when ice islands calve, they can have shelf ice and MLSI components. This includes ice islands 1, 2 & 3 (Fig.4), but there also can be a third component. The shelf ice component of Hobson's Choice is easily identified by its ribbed texture, but it, in turn, is flanked by two roughly triangular areas of dark, almost textureless tone. The dark component on the right side of the shelf ice is the MLSI that was attached to the front of the East Ward Hunt Ice Shelf at the time of the calving and which has remained attached since then. Along the outer edge of the MLSI is a narrow strip of consolidated multiyear pack ice (MYPI) that has become attached only since the calving. A far greater amount of MYPI is attached to the left side of the shelf ice. The MLSI and the MYPI are considered to be integral parts of the ice island and, with the addition of MYPI, the area of Hobson's Choice has increased by 30% to 34 km^2 and the mass by 6% to 7.4 x 10^{11} kg (Jeffries and Sackinger, unpublished MS [b]).

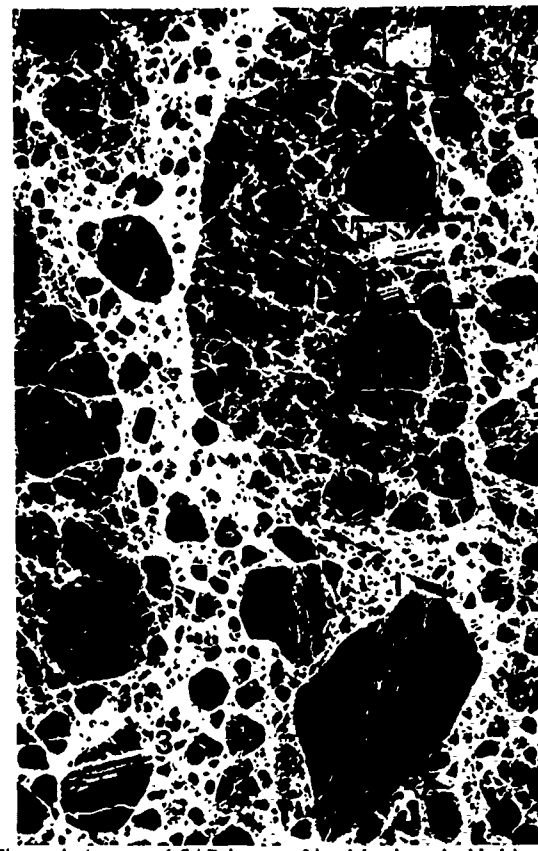


Figure 4. Annotated SAR image of ice islands embedded in pack ice Number 1 (Hobson's Choice) and Number 2 are the largest ice islands presently known in the Arctic Ocean; Number 3 undoubtedly calved in 1982-83 from the East ward Hunt Ice Shelf, but has not previously been observed; the square brackets enclose a cluster of six ice islands, and three individual ice islands are enclosed by boxes The radar illumination direction is from the top to the bottom edge.

There are sometimes vertical ice cliffs at the edges of ice islands, but these are often over-ridden by ridges of ice rubble exceeding the height of the ice island surface. In either case, there will be strong returns from the flat, angular facets, and the margins of ice islands will be well-defined in a SAR image. Where depressions intersect the edges of the ice islands there are often deep incisions due to the catastrophic drainage of the lakes in summer (M. Schmidt, personal communication).

DISCUSSION OF SAR RETURNS

Topography and radar illumination direction

The potential of SAR for the detection of ice islands was first demonstrated by SEASAT, which imaged ice island T-3 in 1978 in

the eastern Beaufort Sea (Fu and Holt, 1982: pp. 126-127). One of the notable features of the T-3 SAR return was a ribbed texture similar to that described in this paper. The ribbed texture of the SAR returns from the ice shelves and the shelf ice component of ice islands has two different forms. On the Milne Ice Shelf, and others (Fig 2), and on those ice islands oriented approximately perpendicular to the illumination direction (e.g. # 3, Fig. 4; also T-3 in Fu and Holt, 1982), the ribbed texture has a distinctive and relatively bright appearance of light grey and mid-grey tones. This can be attributed to the variation of reflectivity with the incidence angle, since there will be high backscatter from the relatively steep hummock slopes directly facing the radar, and reduced backscatter due to low grazing angles of the slopes facing away from the radar. Some of the latter slopes also might be in radar shadow, which, together with possible specular reflection off flat lake ice surfaces in the depressions, will cause a weak return.

On those ice islands oriented nearly parallel with the illumination direction the ribbed texture has a more subdued appearance and the shelf ice has a basic mid-grey tone, with lineaments defined by discontinuous narrow light and dark grey tones (e.g. # 1 and #2: Fig. 4)). In this case the incidence angle varies much less and, consequently, the radar returns are less variable. Also at this orientation, specular reflection from the flat lake ice surfaces in some depressions should result in weak returns. However, where lakes have drained off the edge, leaving an irregular topography, there will be discontinuous weak and strong returns, thus offering the possibility of distinguishing drained lakes from undrained lakes.

The variation of reflectivity with the incidence angle on the undulating surfaces applies also to the MLSI. It is recalled that the MLSI at the front of the Milne Ice Shelf, oriented normal to the illumination direction, has a ribbed return (Fig. 3). This is not the case with the MLSI attached to Hobson's Choice and Ice Island #2, both of which are oriented almost parallel to the illumination direction. Thus, although the ice island/ice shelf undulations are detected regardless of incidence angle and illumination direction, probably because of their relatively high relief, the SAR is a less sensitive indicator with regard to the MLSI undulations, probably because of their relatively low relief and their higher salinity.

SAR evidence of under-ice freshwater-seawater stratification

The inner unit of Milne Ice Shelf and the innermost parts of other fiords have bright, textureless returns (Figs. 2 and 3). Airborne radio-echo sounding in 1980 revealed that the inner unit of the Milne Ice Shelf is less than 10 m thick (Narod et al., in press). Subsequent ice coring has shown it to be as little as 3 m thick, underlain by freshwater in late-winter before the ablation season, and composed of multiyear lake ice with many tubular bubbles and a candled surface (Jeffries, 1986b; 1988; unpublished data). Weeks et al. (1978) have discussed lake ice SAR returns in some detail. In the case of the inner ice shelf unit, the bright tone is most likely due to some combination of scattering from the rough, candled, upper surface and the bubbles, and a strong return resulting from the dielectric constant change at the bottom ice-water interface.

It is believed that the central and outer units of the Milne Ice Shelf act together as a floating, inverted dam causing year-round, freshwater retention and freshwater-seawater stratification below the inner unit (Jeffries, 1988). Hence, the freshwater is found below the inner unit ice cover and provides water for the growth of multiyear lake ice. Fiord stratification is found at other locations along this coast; for example, behind the Ward Hunt Ice Shelf, Disraeli Fiord also has a 3 m thick multiyear lake ice cover (Keys, 1978). Although stratification has not been found by direct oceanographic studies in the innermost parts of Ayles Fiord and M'Clintock Inlet, the strong returns from these areas (Fig 3) are believed also to be evidence of lake ice growth associated with year-round freshwater-seawater stratification. It is interesting to note that the bright return from the ice in M'Clintock Inlet reaches further up the east side of the inlet than it does up the west side (Fig. 2). It may be speculated that this is evidence of the Coriolis force which would tend to move outflowing freshwater over to the right or the east side of the inlet

CONCLUSION

Arctic ice shelves, MLSI and ice islands are interesting targets for SAR, and it, in turn, is a valuable instrument for studies of these unusual and important ice features. The ice shelf/ice island undulations are perhaps the most characteristic feature of the SAR images and the SAR will depict them regardless of their orientation relative to the radar. On the other hand, SAR appears to be able to depict the MLSI undulations only when they are oriented normal to the illumination direction. However, this was not an impediment to the near-real-time detection in February 1988 of the calving of a massive piece of MLSI from the front of the Milne Ice Shelf. This event indicates that SAR is ideal for regular, periodic monitoring of the ice shelves and the MLSI, both for the detection of ice losses and for the assessment of ice regrowth. The SAR data has shown that ice islands can significantly increase their area and mass as they drift, a phenomenon of interest to those involved in arctic navigation and offshore petroleum development. Finally, SAR data of this High Arctic region apparently contains indirect evidence of the effects of the thick ice of the outer fiords on the oceanography and ice cover of the inner fiords.

ACKNOWLEDGEMENTS

This work was supported by the U.S. Dept. of Energy (Morgantown Energy Technology Centre) and the Geophysical Institute, University of Alaska Fairbanks. Logistical support for fieldwork was provided by the Polar Continental Shelf Project. Special thanks are extended to Canarctic Shipping Company Limited, and Pietro de Bastiani in particular, for making available the SAR data.

REFERENCES

1. Ahlnås, K. and Sackinger, W. M., "Offshore winds and pack ice movement episodes off Ellesmere Island". In, Sackinger, W. M. and Jeffries, M. O. (eds.), Port and Ocean Engineering under Arctic Conditions, Vol. III, Geophysical Institute, University of Alaska Fairbanks, 1988.
2. Fu, L. L. and B. Holt, "Seasat views oceans and sea ice with synthetic aperture radar", JPL Publication 81-120, Jet Propulsion Laboratory, Pasadena, California, 1982.
3. Jeffries, M. O., "Ice island calvings and ice shelf changes, Milne Ice Shelf and Ayles Ice Shelf, Ellesmere Island, N.W.T.", Arctic, 39, 15-19, 1986a.
4. Jeffries, M. O., "Glaciers and the morphology and structure of Milne Ice Shelf, Ellesmere Island, N.W.T., Canada", Arctic and Alpine Research, 18, 397-405, 1986b.
5. Jeffries, M. O., "The growth, structure and disintegration of arctic ice shelves". Polar Record, 23, 631-649, 1987.
6. Jeffries, M. O., "Glaciers and the morphology and structure of Milne Ice Shelf, Ellesmere Island, N.W.T., Canada: A reply", Arctic and Alpine Research, 20, 369-371, 1988.
7. Jeffries, M. O. and H. Serson, Recent changes at the front of Ward Hunt Ice Shelf, Arctic 36, 289-290, 1983.
8. Jeffries, M. O. and Sackinger, W. M., "Near-real-time airborne synthetic aperture radar detection of a thick sea ice floe calving from the Milne Ice Shelf, Canadian High Arctic", submitted to Arctic and Alpine Research, unpublished MS (a).
9. Jeffries, M. O. and Sackinger, W. M., "Detection and characterization of ice islands with airborne synthetic radar", submitted to Journal of Geophysical Research, unpublished MS (b).
10. Jeffries, M. O., Sackinger, W. M. and Serson, H. V., "Remote sensing of sea ice growth and melt-pool evolution, Milne Ice Shelf, Ellesmere Island, Canada", Annals of Glaciology, 9, 145-150, 1987.
11. Keys, J. E., "Water regime of Disraeli Fiord, Ellesmere Island", Defence Research Establishment Ottawa, Report 792, 1978.
12. Narod, B. B., G. K. C. Clarke and B. T. Prager, "Airborne radar sounding of glaciers and ice shelves, northern Ellesmere Island, Canada", Canadian Journal of Earth Sciences, in press.
13. Weeks, W. F., A. G. Fountain, M. L. Bryan and C. Elachi, "Differences in radar returns from ice-covered North Slope lakes", Journal of Geophysical Research, 83, 4069-4073, 1978.

STUDIES OF ICE SHEET HYDROLOGY USING SAR

R. A. Bindschadler and P. L. Vornberger*

Code 671, NASA/Goddard Space Flight Center,
Greenbelt, Maryland 20771 USA

*ST Systems Corporation, 4400 Forbes Boulevard,
Lanham, Maryland 20706 USA

ABSTRACT

Analysis of SAR data collected of the Greenland ice sheet in summer and winter suggest the potential use of SAR to monitor the temporal hydrology of ice sheets. Comparison of each SAR data set with summer TM imagery show an areal positive correlation with summer SAR data and a negative correlation with winter SAR data. We hypothesize that the summer SAR data are most sensitive to the variable concentrations of free water in the surface snow while the winter SAR data indicate variations in snow grain size.

INTRODUCTION

The remoteness and harshness of the polar regions make remote sensing particularly useful as a technique for conducting scientific investigations in polar areas. For studies of ice sheets, they are so large that suitable sensors must be mounted on space platforms to enable complete continental and synoptic coverage. The application of space-borne remote sensing data of ice sheets is still in its infancy but investigations using Landsat imagery, radar altimetry, and passive microwave brightness temperatures have all shown exciting potential. This paper focuses on another instrument, the synthetic aperture radar (SAR) and demonstrates how the sensitivity of SAR to snowpack wetness and grain size could prove valuable in the monitoring of the ice sheet hydrology.

ICE SHEET HYDROLOGY

Ice sheets exist because over an extended period of time snow accumulation exceeded loss. The balance of mass gain versus mass loss determines whether the volume of an ice sheet is increasing or decreasing. Because changes in ice volume directly translate into changes in sea level, it is important to understand the effect of climate on the mass balance of ice sheets.

With the exception of the majority of the Antarctic ice sheet, ice sheets experience substantial melt of surface snow during summer. As this snow is melted, the water produced either flows along the local surface gradient, percolates below the surface, or a combination of these two. The component of water that percolates will refreeze if the surrounding snow is at a sub-freezing temperature. The refrozen water usually forms a broad, thin, horizontal ice lens or a narrow, vertical ice column. The latent heat released by the freezing will warm the surrounding snow until its temperature reaches the freezing point. Any additional percolating meltwater will pass through to deeper layers. A complete description of the effects of surface melt on snow stratigraphy is found in Benson (1962).

As fresh snow ages, it recrystallizes into larger, more rounded grains. Buried by subsequent snowfall, this metamorphism of the snow eventually forms incompressible glacier ice. The presence of free water is known to greatly accelerate this process (Male, 1980). Thus, larger-grained snow is a signature of wetter conditions, even when the catalystic water has been removed by refreezing or by water flow.

SENSITIVITY OF SAR

SAR imagery represents the backscatter power of a transmitted radar beam as a function of position. The backscatter is a combination of geometric (or surface) scattering and volumetric (or sub-surface) scattering. At the oblique antenna angles used for SAR, volumetric scattering usually is the larger of the two (Ulaby and others, 1981). This is particularly true over most of the ice sheets where surface slopes are 0.01 or less. While previous examination of SAR imagery has shown that it can detect many of the same surface features seen on more familiar Landsat visible imagery (Hall and Ormsby,1983; Bindschadler and others, 1987), this approach ignores some of the greater sensitivities of SAR. Two of the more dominant effects which we address here are absorption by free water and volume scattering by snow particles.

Energy in the microwave region is strongly attenuated by free water. Ulaby and others (1981) showed that for snow the attenuation is very large for even a small volume fraction of free water. Thus, wet areas should appear darker than dry areas in SAR imagery.

Because the size of snow grains is much less than microwave wavelengths, the scattering follows the Rayleigh criterion:

$$\sigma = \frac{64\,\pi^5\,r^6}{\lambda^4} \qquad (1)$$

where σ is the backscatter crossection, r is the particle radius, and λ is the wavelength (Ulaby and others, 1981, Eqn. 5.80). The high power of the particle size term indicates that the backscatter intensity will be sensitive to small variations in the

grain size. Regions of slightly larger grains will appear much brighter than the surrounding regions.

Finally, the penetration of the radar into the sub-surface could permit the detection of ice lenses. Being rather smooth and flat, ice lenses will tend to act as specular reflectors. Their size is generally less than the pixel resolution of current SARs so areas populated by ice lenses should appear slightly darker than areas without ice lenses.

Given these characteristics of SAR, it is easy to see that SAR could be useful for detecting and monitoring the hydrology of the ice sheets. Changes in their hydrology indicate changes in their mass budget, and are important as they relate to sea level change. In the sections that follow, we give two examples which illustrate the sensitivity of SAR to snow grain size and snow wetness.

SAR ACQUIRED IN WINTER

As a first example, we use SAR data collected of the Greenland ice sheet in April 1987 by the Intera Corporation STAR-1 system (X-band, 3.3 cm wavelength). For the purposes of our analysis, these data required numerous enhancements and co-registration with a Landsat Thematic Mapper (TM) scene (Vornberger and Bindschadler, submitted). The TM scene (id# 51236-14083) was acquired on July 20, 1987. Figure 1 shows identical areas of both SAR and

Figure 1. SAR (top) and TM (bottom) image subscenes of identical areas of the Greenland ice sheet (65.7 ° N, 49.2 ° W). Size of each image is 7.4 km (horizontal) x 7.2 km (vertical).

TM (band 4; 760- 900 nm). The presence of lakes and stream channels in the SAR image confirm that some of the image content is sub-surface structure and variation because in April, the surface of the ice sheet is covered by a homogeneous snow layer and no surface lakes or streams exist.

The comparison of the regional pattern of bright and dark areas in the SAR and TM shows a strong negative correlation; where SAR is bright, TM is dark, and vice versa. We hypothesize that the reason for this result is due to grain size variations in the different regions within the image area. Equation 1 identified the expected increase of SAR intensity with increasing grain size. Work by Dozier and Marks (1987) has shown that for TM band 4 there is a decrease in reflectance with increasing grain size. This is the effect seen in Figure 1.

We suspect that the variation in grain size is a result of topographically-driven concentrations of surface meltwater which subsequently accelerates the snow metamorphism creating larger snow grains. These areas of meltwater accumulation should correspond with topographic lows. The extreme cases are the surface lakes which collect water faster than water can drain along the surface or through the snowpack.

SAR ACQUIRED IN LATE SUMMER

When the SAR data are acquired in summer, the character of the backscatter is quite different. The example used here comes from the Seasat satellite SAR (L-band, 23.5 cm wavelength) where an image of western Greenland was acquired on October 9, 1978 (see Bindschadler and others, 1987, Figure 7a). Although not at the same geographic location as Figure 1, the subscene chosen (Figure 2) is in a similar hydrologic location: near the summer snowline where numerous lakes form. As with Figure 1, a summer TM scene is used for comparison (id# 51261-14033 acquired August 14, 1987). Figure 2 shows a positive correlation between SAR and TM: where SAR is bright so is TM and where SAR is dark so is TM.

We hypothesize that the cause of the positive correlation is the presence of free water which reduces the pixel intensities of both SAR and TM the greater the amount of water. As mentioned above, water produces larger snow grains in the wetter area but for SAR the brightening effect of larger snow grains is overwhelmed by the water absorption effect, while for TM both water and grain size act in the same direction. The effect of snow-grain size on SAR intensity will mollify the darkening due to water absorption but the latter definitely will dominate over the expected ranges of free water content and grain size.

DISCUSSION

We have given two examples of SAR data over ice sheets and compared each with summer TM imagery. Each area was in a region where melting occurs during the summer and some snow is left at the end of the summer melt season. The primary difference between these two comparisons is that one uses SAR data acquired during winter and the other uses SAR data acquired during late summer. Given the sensitivities of SAR to free water content and snow-grain size we feel this difference is significant. We predict that for a fixed location on an ice sheet where snow melts but survives the summer, the pattern of bright and dark areas in SAR imagery will reverse in summer and winter. The fact that the pattern will reverse in sign rather than shift in intensity removes any reliance in stably-calibrated SARs and should make

Figure 2. SAR (top) and TM (bottom) image subscenes of identical areas of the Greenland ice sheet (64.4° N, 48.8° W). Size of each image is 7.4 km (horizontal) x 7.2 km (vertical).

detection simple. Where stably-calibrated SARs could be used effectively is in temporal comparisons of backscatter from ice sheet areas as a means to quantify the free water fraction. Further data need to be collected at the same ice sheet area at different seasons to confirm this reversal behavior. In situ wetness and grain-size data would also prove very useful to examine whether subsurface ice lenses can be detected.

The importance of this study rests in the need to monitor temporal variations in global climate. The extent of the ice sheet zone we have investigated (a combination of the percolation and soaked zones as defined by Benson, [1962]) is a signature of the interaction of the atmosphere with the ice sheet through the processes of snow accumulation and melting. Variations in the size or character of this zone will indicate variations in the polar climate.

ACKNOWLEDGEMENTS

This research was supported by the NASA Oceans Program.

REFERENCES

Benson, C.S., 1962. Stratigraphic studies of the snow and firn of the Greenland ice sheet, US Snow and Ice and Permafrost Research Establishment, US Army Corps of Engineers, Research Report 70, 98pp.

Bindschadler, R.A., K.C. Jezek, and J. Crawford, 1987. Glaciological investigations using the synthetic aperture radar imaging system, Annals of Glaciology, Vol. 9, p. 11-19.

Bindschadler, R.A. and P.L. Vornberger, submitted. Multi-spectral analysis of ice sheets using co-registered SAR and TM imagery, Photogrammetric Engineering and Remote Sensing.

Dozier J. and D. Marks, 1987. Snow mapping and classification from Landsat Thematic Mapper data, Annals of Glaciology, Vol. 9, p. 97-103.

Hall, D.K. and J.P. Ormsby, 1983. Use of Seasat synthetic aperture radar and Landsat multispectral scanner subsystem data for Alaskan glacier studies, Journal of Geophysical Research, Vol. 88 (C3), p. 1597-1607.

Male, D.H., 1980. The seasonal snowcover, in Dynamics of Snow and Ice Masses, S.C. Colbeck (ed.), Academic Press, Inc. New York, p. 305-396.

Ulaby, F.T., R.K. Moore, and A.K. Fung, 1981. Microwave Remote Sensing, Active and Passive, Addison-Wesley Publishing Company, Reading, Massachusetts, 3 volumes.

KEYWORDS: ice sheet, SAR, hydrology

A MATCHED-FILTER TECHNIQUE FOR REMOVING HYPERBOLIC EFFECTS DUE TO POINT SCATTERERS SIMULATION AND APPLICATION ON ANTARCTIC RADAR DATA

G. Raju and R. K. Moore

Radar Systems and Remote Sensing Laboratory
University of Kansas Center for Research, Inc.
2291 Irving Hill Road
Lawrence KS 66045-2969, USA
913/864-4835
FAX: 913/864-7789, TELEX: 706352, OMNET: KANSAS.U.RSL

ABSTRACT

A modern coherent ice-probing radar (the University of Kansas coherent Antarctic radar depth sounder or CARDS) for probing the ice sheets of Antarctica and Greenland was successfully operated on Downstream B, a dynamic glacier in West Antarctica, in the austral summer of 1987-88. The results clearly showed strong bottom echoes and several layered structures. Quantitative results on the bottom roughness are described in a companion paper.

One of the inherent problems prevalent in wide-angle radar observations is the undesirable hyperbolic echo profiles caused by isolated scatterers. A similar problem in seismic shooting is solved by migration methods. Although this effect has been considerably reduced in our radar by choosing a higher frequency and unfocused processing, it is still observed in the data. A matched-filter technique has been developed for removing these effects. The method involves generating a number of reference functions appropriate for the actual data and cross-correlating with the data. The correlation maximum determines the location of the isolated scatterer/s. This process is done in the frequency domain using FFTs applied on both the data and the reference function.

We have simulated performance for several cases of point scatterers. Samples of actual data have been used to prove its practical importance and effectiveness.

INTRODUCTION

Radio echo (radar) sounding of the ice sheets in Antarctica and Greenland has provided much information about the ice thickness and properties of the bottom over the years. Recordings usually show the signal strength as intensity mapped on an "image" with abscissa the horizontal distance and ordinate the depth to the target (bottom or internal reflector/scatterer). Radars for this application must use low frequencies if the signal is to penetrate through the ice; 60 MHz has often been used.

Use of a low frequency means that the antennas for airborne or sled-mounted systems necessarily have wide beams. Consequently the radars receive signals from targets over a wide range of angles. For the larger angles the range to a point scatterer in or beneath the ice is considerably larger than the depth. When the radar traverses over the target, the signal range varies from a maximum when the edge of the beam strikes it through a minimum when the radar is over the target to a maximum as the beam last strikes the target. The result on the image is that each strong point target appears as a hyperbola. Figure 1 shows a simulation of the signals from a point target at several points along a traverse, indicating the effect of the range variation.

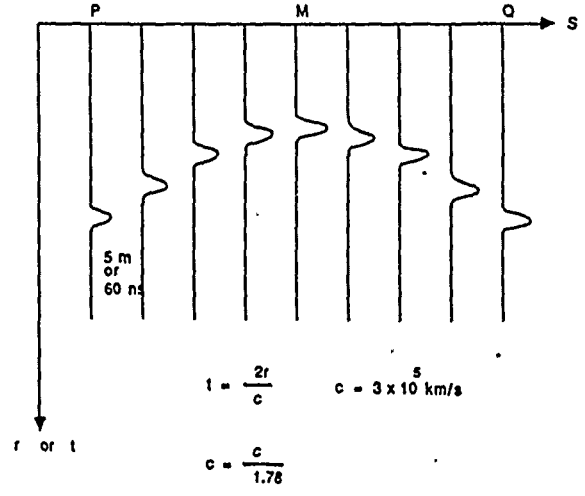


Fig. 1 Hyperbolic profile in recorded data due to a single scatterer

These hyperbolas frequently overlap, resulting in images that are difficult to interpret. Hyperbolic returns from shallower strong targets at the edges of the beam may obscure weaker returns from targets near the beam center. Thus, we need a method that can replace each hyperbola with a point at the location of the target. In this paper we present a correlation (or matched-filter) technique appropriate to the problem.

The University of Kansas developed a modern-technology 150-MHz radar as a replacement for the old-technology radars used previously. This radar was tested in the Antarctic in 1986 and 1987, and used for glaciology in 1988. We applied the correlation technique to samples of the 1987 measurements and found that it significantly improved the appearance of the images.

THE GEOMETRY

Figure 2 shows the geometry needed to develop expressions for the hyperbolic return signals. The radar travels along the ice surface on a path directly above a target whose depth is the minimum range R_o. The radar is shown at point O, a distance x along the track from the position of the target. The slant range R is also the displacement of the target in the vertical direction ON THE IMAGE. If y is the vertical displacement ON THE IMAGE, we can write

$$R^2 - x^2 = R_o^2 \quad \text{or} \quad y^2 - x^2 = R_o^2 \qquad (1)$$

as the equation of the image hyperbola. When this is combined with an appropriate finite pulse shape, we obtain the kind of result shown in Fig. 1. The figure shows the amplitudes for sample pulses; on the image these amplitudes would appear as intensities.

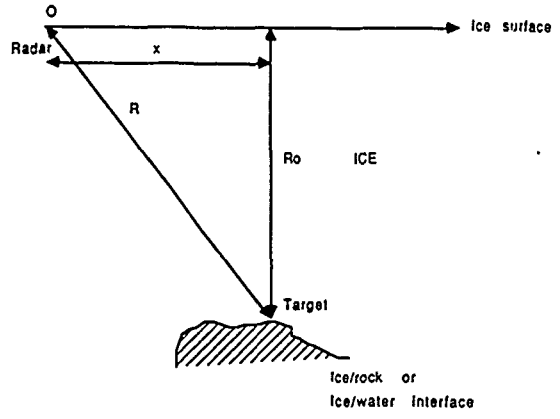


Fig. 2 Geometry for simulation

THE CORRELATION OR MATCHED-FILTER PROCESS

Although one can think of the process used here as a matched filter, a correlation approach is easier to understand. Consider the situation shown in Fig. 1. If we have a target that produces such a hyperbolic return and a reference function of the same shape with the same minimum range, we may translate the reference in the x direction. In correlation, we take the product of the signal function and the reference function and integrate. Clearly, if the reference function has its peak at the same x coordinate as the target, the functions overlap and the integral of the product is maximum. For any other x coordinate of the reference peak the overlap is less, the products are smaller at each point, and the integral is smaller. A modest x displacement is enough to reduce the integral significantly, and a large displacement reduces it to zero.

If the reference has its peak at a different range than the signal, the overlap is also reduced. If it is far enough from the target, the integrated output is zero.

Thus, the correlation process consists of moving reference functions over the image plane and performing the required integration for each location. The output is a "spike" at the location of the target, with the width of the spike dependent on pulse width and shape and on the curvature of the hyperbolas (depends on target depth).

MATHEMATICAL DEVELOPMENT OF THE MATCHED FILTER

In the usual time-domain matched-filter process we consider a signal s(t) passing through a filter with impulse response h(t), obtaining

$$y(t) = \int_{-\infty}^{\infty} s(\tau)h(t-\tau)d\tau. \qquad (2)$$

The matched filter is defined as having $h(t) = s(-t)$. For symmetrical signal functions, however,

$$h(t-\tau) = h(t+\tau),$$

so the result is equivalent to correlation:

$$y(t) = \int_{-\infty}^{\infty} s(\tau)h(t+\tau)d\tau \qquad (3)$$

In this case, the image is "four-dimensional": three spatial dimensions (along- and cross-track and depth) and complex amplitude. Here we do not record the phase, so that fifth dimension does not apply. The correlation process takes place in the along-track direction and improves the resolution while eliminating the hyperbolas. The antenna beamwidth sets the cross-track contributions to the image; no cross-track dimension appears in the image because no scanning takes place in that direction. However, the correlation process provides modest narrowing of the beam in the cross-track direction because the different ranges to cross-track targets give different hyperbolas that do not match the reference well.

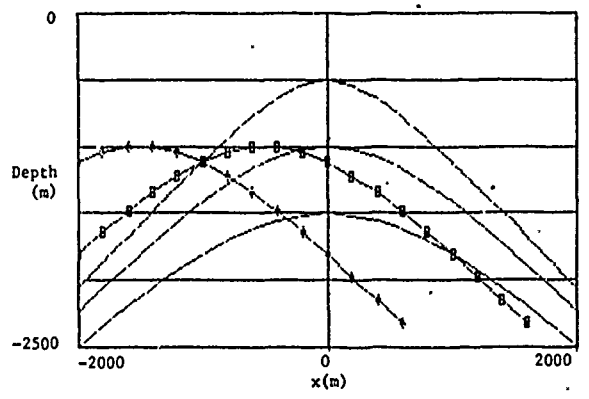


Figure 3. Illustration of hyperbolas for different depths and reference functions offset in the x direction. Target hyperbolic responses are shown at 500, 1000, and 1500 m depths. Offset reference hyperbolas are at 1000 m depth, with squares offset 500 m and pluses offset 1500 m.

The shape of the hyperbolas like that in Fig. 1 depends on depth to target, z. They are sharper at shallower depths. For this reason, the reference hyperbolas for different depths have different shapes, so a common reference function cannot be used. This point and the basic process are illustrated in Fig. 3. Here we assume the target is at 1000-m depth. Other targets (or references) are at 500 and 1500 m; the different shapes are clear. Also shown are two

offset references at 1000 m, one offset by 500 m and the other by 1500 m. Although references at different depths do not overlap a target response at the same x coordinate, they do intersect responses from deeper targets. Since these intersections are short (single points for the lines in the figure), they contribute little to the correlation for the deeper targets.

Since the radar output is in power, the correlation is of power functions (no phase considered). Let $f(x,z;z_o)$ be the received-power function for a target at depth z_o. We may generate a series of reference functions $g(x,z;z_1)$, where z_1 is the depth of a target for which g is matched. If a(t) is the output of the correlation of f and g,

$$a(x;z_o) = \int_{-\infty}^{\infty}\int_{-\infty}^{\infty} f(x,z;z_o)g[(x+\tau),z;z_1]\,d\tau\,dz. \qquad (4)$$

The correlation is only in the x direction, but the integral must include the z direction. This is the difference from ordinary correlations of simple time functions. Hence, we may use Fourier transforms to simplify the computation in the x dimension only:

$$A(k_x;z_o)=\int_{-\infty}^{\infty} e^{-jk_x x}\int_{-\infty}^{\infty}\int_{-\infty}^{\infty} f(x,z;z_o)g[(x+\tau),a;z_1]\,d\tau\,dz\,dx \qquad (5)$$

where A is the Fourier transform of a. Changing the order of integration and letting $\sigma = x + \tau$:

$$A(k_x;z_o)=\int_{-\infty}^{\infty}dz\int_{-\infty}^{\infty} f(x,z;z_o)e^{jk_x x}\int_{-\infty}^{\infty} g[(\sigma),z;z_1]e^{-jk_x\sigma}\,dx\,d\sigma.$$

The two inner integrals are Fourier transforms, so

$$A(k_x;z_o)=\int_{-\infty}^{\infty}dz F^{*}(k_x z;z_o)G(k_x x;z_1), \qquad (6)$$

where F and G are the transforms of f and g and F^* is the complex conjugate of F. This is different from the usual application of matched filtering to one-dimensional processes because of the z integral. It differs from two-dimensional matched filtering because one cannot transform with respect to z, but must integrate. The Fourier transform allows use of FFT routines for faster processing of long data sets.

The result of this is that one must evaluate these integrals for sets of z_1, taking the maximum value of A as the one corresponding to a match. This makes the computation lengthy, but in the case of large empty depth regions (frequently encountered in practice), the trial values of z_1 need only be near the locations of actual target returns.

To obtain the actual response, one then takes the inverse transform of $A(k_x;z_o)$ as given in (6) to obtain $a(x;z_o)$.

SIMULATED RESPONSES

The actual resolution of the University of Kansas radar is 6 m in the ice, but simulations were performed for a radar with 50-m resolution to allow the nature of the response to be shown better.

The two-dimensional response for a single point target is shown in Fig. 4. Its half-power width is 100 m in both directions; the width for a correlation function is usually on the order of twice the width of the initial pulse. Clearly the hyperbolic nature of the original signal was entirely removed.

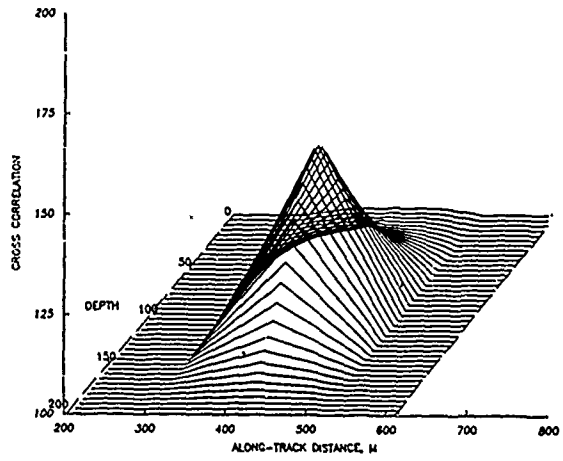


Fig. 4 Three-dimensional display of identified target at 100 m

Figure 5 shows an example of the response for a pair of targets separated 100 m along track and 75 m in depth. For this situation the targets are totally separated, although the weaker parts of the response for the shallower target would, without processing, overlap the weaker parts of the deep-target response.

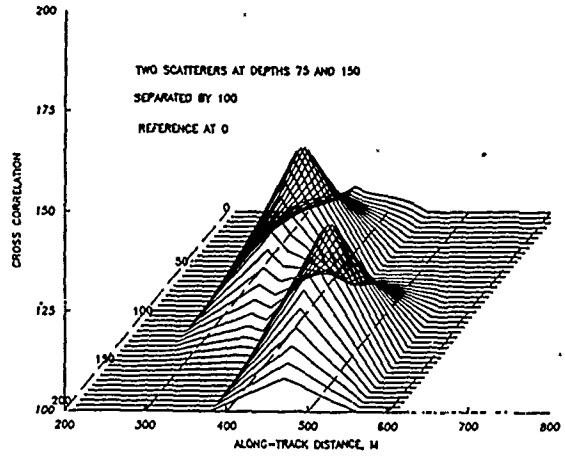


Fig. 5 Three-dimensional view of a transformed result

Figure 6 shows the correlation functions for different values of z_1 when the targets are 100 m apart along track and 50 m apart in depth. The effect of the target at 100-m depth is to broaden the response for z_1 = 50 m, but only at low levels. The 50-m deep target has a similar effect on the z_1 = 100-m response. The responses at 75-m and 80-m depths show that the outputs overlap in this region, but they are far enough below the peaks to make the presence of two targets obvious.

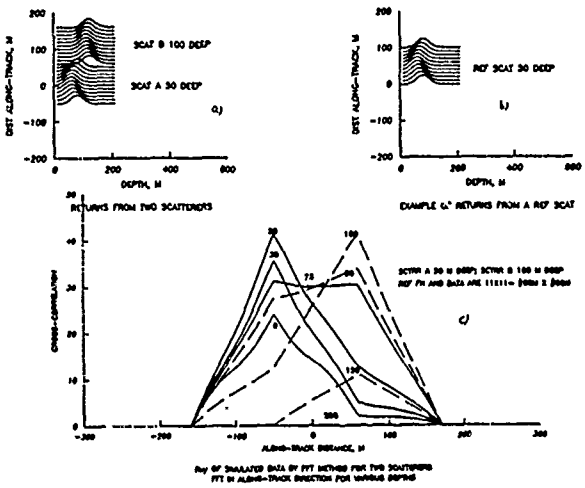


Fig. 6 Two targets at different depths and their locations identified after matched filtering

The appearance of these responses on an image is given in Fig. 7. This uses resolution comparable with that for the actual radar (5 m). The targets are only 5 m apart along track and 7.5 m in range. Unfortunately, the printed reproduction of the gray scales makes the targets appear to overlap more than they would in a good photographic rendition with adequate gray-level range.

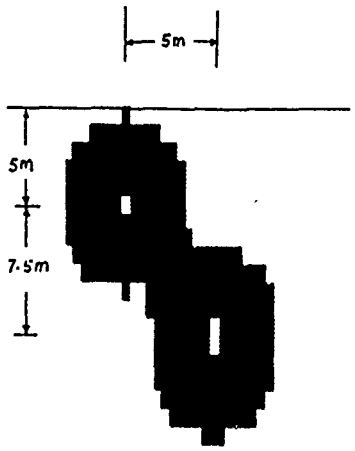


Fig. 7 Image of two well-separated targets identified by filtering

RESULTS ON REAL DATA

Our data from Ice Stream B in West Antarctica have few of the hyperbolas common in other Antarctic regions. Apparently the bottom is smooth enough that point targets are the exception in this area. However, many point scatterers seem to be present in the volume within 350 m of the surface. Figure 8 shows an example for a ground-based traverse made in 1987. Figures 8(a) and (b) show the original data as given on a computer printer with two different threshold levels. The vertical streaks extending through the image are artifacts from noise introduced into the system. One can see a rather jumbled array of volume-scatter returns in the upper layer, with some hints of hyperbolic shapes. However, when the processing algorithm was applied, the result is as shown in Fig. 8(c). Clear patterns of the nature of the upper-layer scatter emerge.

(a) Raw image (b) Raw image: same as (a) but with a different gray scale (c) Filtered image

Fig. 8 Effect of matched filtering applied to a raw image

CONCLUSION

The algorithm presented here allows removal of the hyperbolic confusion associated with point-target echoes in sounding through glacial ice. It made a significant difference in the appearance of the response in the upper layers of Ice Sheet B.

The algorithm uses a correlation, or matched-filter, approach to convert the hyperbolic point-target responses into spike responses. The nature of the correlation is different than in most applications because of the need to correlate in the along-track direction while integrating in the range direction. We believe this is a new approach that can significantly improve displays of the outputs of glacial-ice sounding radars and other radars with wide beamwidths.

ACKNOWLEDGEMENT

This work was supported by the National Science Foundation Division of Polar Programs under Grant DPP-8300450. Computation was supported in part by the University of Kansas Computation Center through the Engineering Computation Facility.

REFERENCES

In the interest of brevity, references have been omitted in this paper. Many references on the topic discussed are available from the authors.

RADAR REFLECTIONS FROM WATER INJECTED INTO AN ANTARCTIC GLACIER

R. H. Dean, C. H. Davis and W. Xin

Radar Systems and Remote Sensing Laboratory
University of Kansas Center for Research, Inc.
2291 Irving Hill Road
Lawrence KS 66045-2969, USA
913/864-4835
FAX: 913/864-7789, TELEX: 706352, OMNET: KANSAS.U.RSL

ABSTRACT

Most models of glaciers that have depths in excess of 500 m propose that the ice present in the deepest half of the glacier is totally compressed by the mass of the ice above it.

A 150-MHz coherent radar system was used to survey areas adjacent to a hot water drilling operation. The drilling was conducted on a 1030-m deep portion of Ice Stream B in West Antarctica. On two separate occasions a sequence of radar traverses observed strong specular reflections from a temporary horizontal water layer injected into the glacier by the drilling operation. This is the first radar observation of water placed deep into a glacier from a known source.

On the first occasion the layer was observed at a depth of 840 m, while on the second occasion the layer was observed at 780 m. In both cases the water layer refroze over a period of 8 to 14 hours. The horizontal openings that received the water from the drilling operation are a totally unexpected structural form at such great depths within a glacier. These two observations indicate that such structures do exist and suggest that these phenomenon require more detailed study for a satisfactory explanation.

ACKNOWLEDGEMENT

This work is supported by the National Science Foundation Division of Polar Programs, grant DPP-8716940.

BOTTOM CHARACTERISTICS OF THE ANTARCTIC ICE SHEET FROM RADAR

R. K. Moore, W. Xin & G. Raju

Radar Systems and Remote Sensing Laboratory
University of Kansas Center for Research, Inc.
2291 Irving Hill Road
Lawrence KS 66045-2969, USA
913/864-4835
FAX: 913/864-7789, TELEX: 706352, OMNET: KANSAS.U.RSL

ABSTRACT

An important quantity in the study of movement of the Antarctic glacial ice is the bottom roughness. The University of Kansas coherent Antarctic radar depth sounder (CARDS) has made 150-MHz backscatter measurements through the ice cap for two field seasons. The bottom echoes have characteristics similar to those received from a radar altimeter over land or sea and to the echoes used by astronomical radars to study characteristics of remote planetary objects.

We have used a simple scattering model to synthesize expected return pulses and compared it with those returned to CARDS from the bottom of the ice sheet. The model has two parameters: standard deviation of height and horizontal autocorrelation length of the bottom roughness. A least-squares fitting procedure allows determination of the appropriate parameters for the observed roughness. Variations in different locations beneath the ice sheet indicated by these results are of direct interest to glaciologists.

INTRODUCTION

Studies of the dynamics of the Antarctic ice sheets are important in that they may lead to information on long-term climatic trends. Moreover, these studies help in determining whether retreat of the ice sheets might lead to significant rises in sea level on a global scale. Bottom roughness is important in such studies because it is an indicator of the friction that impedes glacial flow.

Radar returns depend on functions of roughness and permittivity. Thus they may serve to measure roughness, a method widely used in radar astronomy [Daniels, 1961]. We may apply a similar method to the study of bottom roughness of an ice sheet. Echoes from the bottom of the ice have characteristics similar to those received from a radar altimeter over land or sea and radar telescope signals used to study characteristics of remote planetary objects.

The University of Kansas 150-MHz Coherent Antarctic Radar Depth Sounder (CARDS) made backscatter measurements through the ice cap of West Antarctica during the last two field seasons: 1987 at a downstream area and 1988 at an upstream site in ice stream B.

We used two simple one-dimensional scattering models to synthesize expected returned pulses. We then used a least-mean-square algorithm to compare them with pulses returned to CARDS from the bottom of the ice sheet. This provided an estimate of the characteristics of bottom roughness.

MODELING OF RADAR BOTTOM ECHO

To develop the model for the radar echo from the bottom surface, we consider a plane wave propagating toward a rough surface at angle of incidence θ. The scattering was described by a standard physical-optics model. The mean power returned from the rough surface depends for this model on the form chosen for the surface correlation function [Fung & Moore, 1966; Skolnik, 1975].

Two correlation functions frequently appear in the scattering literature: exponential and Gaussian. If the surface correlation function is exponential, it takes the form

$$\exp(-|x|/L) , \tag{1}$$

where x is the height relative to the mean level of the rough surface and L is its correlation length. Then the mean backscattered power from the rough surface is

$$p(\theta) = \frac{K\,\theta}{\cos^2\theta \sin\theta} \left(1 + A \frac{\sin^2\theta}{\cos^4\theta}\right)^{-3/2}, \tag{2}$$

$$A = \left(\frac{L}{4\pi\sigma^2}\right)^2. \tag{3}$$

θ is the incident angle of the plane wave, σ is the standard deviation of surface height, and K is a constant.

On the other hand, if the correlation function is Gaussian, it is

$$\exp(-x^2L^2). \tag{4}$$

The mean backscattered power can then be expressed as

$$p(\theta) = \frac{K}{\sin\theta} \exp\left(\frac{-L^2}{2\sigma^2} \tan^2\theta\right). \tag{5}$$

The meanings of θ, K, σ, and L are the same as in (2).

Since CARDS points vertically down, signals for the bottom echo are composed of multiple backscattering from different angles near vertical. At each angle,

the mean backscattered power is proportional to $p(\theta)$ as in (2) or (5). For CARDS operating on the surface, θ is function of time described by

$$\theta(t) = \cos^{-1}(h/r) = \cos^{-1}\left(\frac{h}{h + ct/2}\right) \quad t > 0, \qquad (6)$$

where h is the height of radar receivor above the rough bottom surface, r is the distance between the transmitter and the scatterer, c is speed of light in the ice, and t is delay time.

Considering the bottom echo as a weighted summation of continuous scattering from different angles, we can model the mean returned voltage pulse as a convolution of the transmitted pulse with a weighting function

$$v_r(t) = \int_0^\infty v_t(x)\, w(t-x)\, dx \qquad (7)$$

where $v_r(t)$ is the radar echo, $v_t(t)$ is the transmitted pulse, and $w(t)$ is a random weighting function. Under the assumptions that the signals backscattered from the rough surface at different angles are uncorrelated and that the antenna pattern is isotropic, $w(t)$ has the form

$$w(t) = p(\theta)^{1/2} x(t) \qquad (8)$$

Since θ is function of time, $p(\theta)$ can be written as $p(t)$. The fading is approximated by a zero-mean random process $x(t)$ with a correlation function

$$R_{xx}(t) = \delta(t).$$

Notice that $w(t)$ is not stationary because $p(\theta)$ depends explicitly on t. Its correlation function is

$$R_{ww}(t,u) = p(t)^{1/2} p(u)^{1/2}\, \delta(t-u). \qquad (9)$$

The mean power of $w(t)$ is

$$\sigma_{ww}(t) = p(\theta(t)).$$

With (8), we can estimate mean power of the bottom echo as

$$\hat{p}(t) = E\left(\int_0^\infty\int_0^\infty v(x)\, w(t-x)\, v(u)\, w(t-u)\right)$$

$$= \int_0^\infty\int_0^\infty v(x)\, v(u)\, R_{ww}(t-u,\, t-x)\, dx\, du$$

$$= \int_0^\infty v(x)^2\, p(t-x)\, dx, \qquad (10)$$

which is a convolution of the power of a transmitted pulse and the mean returned power of plane waves.

For CARDS, the transmitted pulse $v(t)$ has an approximately Gaussian shape:

$$v(t) = \exp\left(-(t/q)^2\right),$$

where q is a scale factor used to match the 3-dB pulse width. The mean-square error of estimation is defined as

$$MSE = \int_0^\infty \left(p(t) - \hat{p}(t,L,\sigma)\right)^2 dt, \qquad (11)$$

where $p(t)$ is the mean power of the bottom echo measured by CARDS, and $\hat{p}(t,L,\sigma)$ is its mean power estimated from (10) with parameters L and σ^2. However, the returned power only depends on the ratio of L and σ^2 for exponential correlation as in (2) or the ratio of L and σ for Gaussian correlation as in (5).

An iterative procedure was used to determine the model parameters that minimized the MSE. For this purpose $p(t)$ was the mean of 10 adjacent pulses.

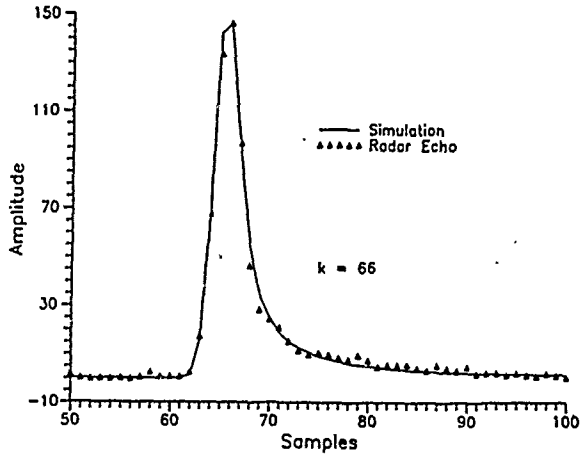


Fig. 1a

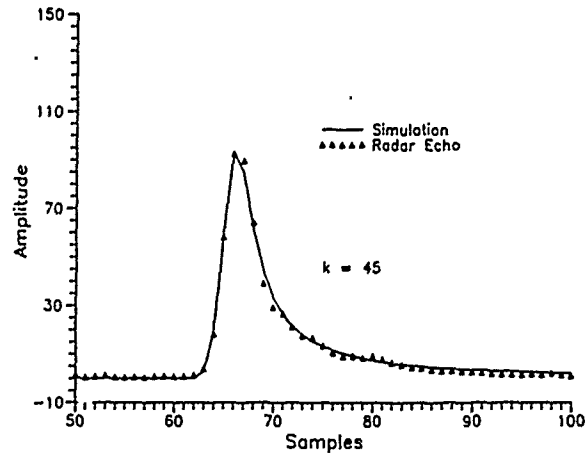


Fig. 1b

Fig. 1 Comparison of radar bottom echoes and computer simulations using a scattering model based on an exponential correlation function. 1a and 1b show the bottom echoes from two different places in the downstream region of Ice Stream B in West Antarctica. The data are from 1987. k is the smoothness factor L/σ^2.

RESULTS OF COMPUTER SIMULATION

Using (10) and (11), we compared the computer simulations with the data collected during 1987 and 1988 from Antarctica. We found that the model developed with the Gaussian correlation function fits the measured data poorly. With the exponential-correlation model, the simulated returned pulses and the data agree well.

Figures 1 and 2 show these results for examples from 1987 (downstream B) and 1988 (upstream B), respectively. Figures 1a and 1b compare the measured radar bottom echoes from two different places at the downstream location and their computer simulations. The bottom surface of 1a is smoother than that of 1b since the ratio $k = L/\sigma^2$ is considerably higher than for 1b.

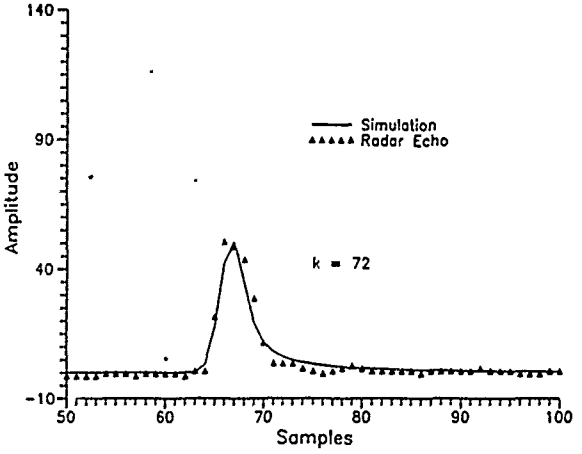


Fig. 2a

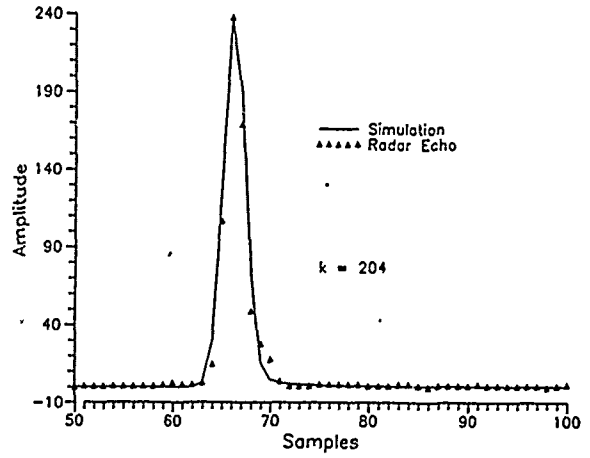


Fig. 2b

Fig. 2 Comparison of radar bottom echoes and computer simulations using a scattering model based on an exponential correlation function. 2a and 2b show the bottom echoes from two different places in the upstream region of Ice Stream B in West Antarctica. The data are from 1988. k is the smoothness factor L/σ^2.

The bottom of the ice is typically smoother at upstream B than at the downstream site. Figure 2a is atypical for the upstream area, and Figure 2b is typical. The smoothness factor k for Figure 2a is only slightly greater than those for Figure 1. For the more typical case shown in Figure 2b, it is much greater. Indeed, the surface in this case is nearly a specular reflector, although it is still rough.

Unfortunately, one cannot fully describe the roughness from such measurements. With the exponential roughness correlation function, the shape parameter for the echo is

$$k = L/\sigma^2.$$

so the correlation length and standard deviation of height cannot be separated. The rms slope of a surface is related to σ/L, and this is certainly a measure of roughness in one sense. The standard deviation σ is a measure in another sense. Thus the smoothness descriptor k may be thought of as

$$k = \frac{1}{(\text{sd of roughness})(\text{rms slope})}. \qquad (12)$$

Further work is needed to establish the relationship between this parameter and sliding friction, the quantity of glaciological interest.

SUMMARY

This study hows that rough-surface modeling may be very useful for determining the bottom roughness of the Antarctic ice sheet from radar echoes. However, the model presented here does not account for differences in roughness with direction, since the scattering function is based on an isotropic model for the correlation function. Although the antenna beam shape was not taken into account in the computer simulation, the wide beams associated with low-frequency radars mean that this factor is unimportant--the beamwidth of the scattering is much narrower than that of the antenna. The coherent integration used in CARDS reduces the effective beamwidth in the direction of travel, however. Hence, an improved model will be developed to account for the effect of the synthetic beam.

The smoothness factor that arises from the successful use of the exponential correlation function for bottom-surface roughness includes both rms slope and height standard deviation. We do not see how these two components of the roughness may be separated using radar echoes. However, this method provides the first fine-scale bottom-roughness estimates for the Antarctic glaciers. Relating the smoothness parameter to sliding friction is a task for the future.

REFERENCES

Daniels, F. B., "A theory of Radar Reflection from the Moon and Planets," J. Geophys. Res., vol. 66, no. 6, pp. 1781-1788, 1961.

Fung, A. K. & R. K. Moore, "The Correlation Function in Kirchhoff's Method of Solution of Scattering of Waves from Statistically Rough Surfaces," J. Geophys. Res., vol. 71, pp. 2939-2943, June 1966.

Skolnik, M. I., ed. Radar Handbook, Chapt. 5, McGraw-Hill, Inc.: New York, 1975.

ACKNOWLEDGEMENT

Research supported by National Science Foundation, Division of Polar Programs, Grant DPP-8300450.

SLOPE CORRECTION BY RELOCATION FOR SATELLITE RADAR ALTIMETRY

A.P.R. Cooper*

Scott Polar Research Institute
University of Cambridge
Lensfield Road
Cambridge
U.K.

Abstract

Because of their high accuracy, satellite altimeter data give the promise of being able to resolve many questions concerning the stability or otherwise of the large ice sheets of the world, and especially to show whether these ice sheets are increasing or decreasing in volume. The drawback of current and proposed altimeters is that they have a large footprint, of the order of 20 km across, and measure the range to the nearest point on the surface within that footprint. This results in apparent errors of up to 100 metres in surface elevation measurements, dependent on the slope of the surface. This paper addresses the problem of converting the range measurements from altimeter data into true surface elevations, using the relocation method of Harrison(1970). An algorithm is presented that takes into account the geometry of the satellite case, and which uses the inevitable across track drift of nominal repeat orbits in order to determine the correction in both the along and across track direction. Using this method, no external surface slope data are needed, and local surface undulations on length scales of 1 km or less can be resolved. The accuracy of the method is of the same order as that of the original measurements in all except a minority of cases, and these cases can be detected by the algorithm. The algorithm is tested using a model ice sheet surface, and accuracy and quality assurance are discussed. Although the method is discussed in the context of ice sheets, it is equally applicable to other land surfaces over which the altimeter retains lock.

Keywords

Altimetry, Slope Correction, Ice Caps

Introduction

Satellite altimeter data have the potential for measuring very small (~20 cm) changes in the elevation of suitable terrain surfaces. One of the main problems with these data is the difficulty of comparing them with data obtained by other methods, e.g. Geoceivers, GPS etc., because the altimeter ranges to the closest point of the surface within its footprint, rather than to the nadir point. Various methods of converting the measured range to a nadir range have been proposed (Brenner and others, 1983), but these rely on a variety of assumptions, the most obvious being that there is constancy of slope between the nearest point on the surface and the nadir point. A further difficulty is that of taking into account cross-track slopes. This paper presents a method of performing slope correction by relocating the measurement point to the point nearest to the satellite, and correcting the range appropriately. In order to obtain the cross track component of the correction, several orbits are used, and the natural scatter of these orbits is used to compile a correction in both along and across track directions.

* Now at British Antarctic Survey, High Cross, Madingley Rd, Cambridge CB3 0ET, Great Britain.

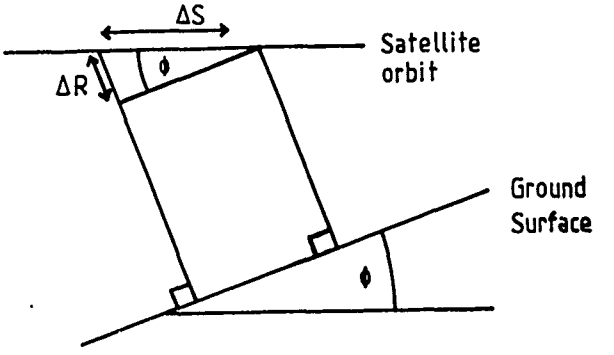


Fig 1 Diagram to illustrate the relationship between surface slope and range rate.

Theory

The method is adapted from that of Harrison (1970), with amendments for the 3-dimensional case, and for the somewhat different geometry that must be considered for the satellite case. If, at a particular sub-satellite point, the rate of change of range of the ground from the satellite with respect to distance along the ground is dR/dS, and the direction of maximum slope with respect to the chosen coordinate system is θ, then considering figure 1 we obtain :

$$\Phi = \sin^{-1}(dR/dS)$$

Then, considering figure 2, we see that :

$$\Gamma = \sin^{-1}(R.\sin\Phi/R_s)$$

and

$$R_c = R_s(1.0-\sin(\Phi-\Gamma)/\sin\Phi)$$

and hence, resolving Γ into components along and across track we have:

$$\delta x = \Gamma.\cos\theta$$

$$\delta y = \Gamma.\sin\theta$$

Where Φ is the angle of slope of the surface, R_c is the corrected range from the spherical shell containing the satellite orbit, R_s is the radius of the satellite orbit, R is the range from the satellite to the surface, Γ is the angular distance from the sub-satellite point to the corrected position, and δx and δy are the angular distances along and across track from the sub-satellite point to the corrected position.

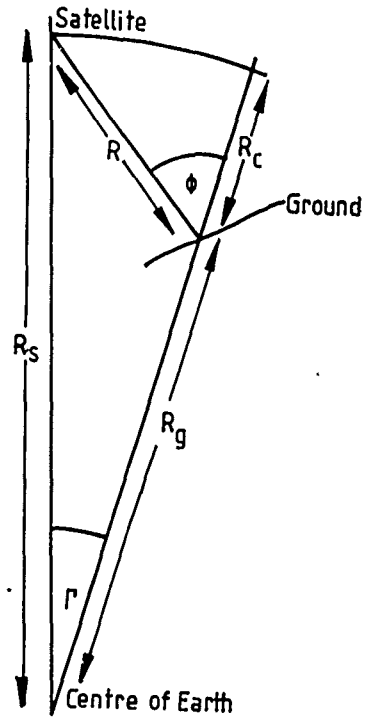


Fig 2 Diagram to illustrate the geometry of slope correction in the satellite case.

Implementation

In implementing this method, the most critical element is to estimate the value of dR/dS and θ for each satellite ground position. The method finally chosen after experiment was to fit a plane to the data from the satellite, using a weighted least squares method. The weighting used was proportional to the inverse of the square of the distance from the required point, and normalized to the average spacing of the points. If this resulted in a weight greater than 1, then the weight was set to 1. Then, the equation of the plane

$$R = Ax + By + C$$

was obtained by solving the simultaneous equations :

$$\begin{bmatrix} \Sigma Rx \\ \Sigma Ry \\ \Sigma W \end{bmatrix} = \begin{bmatrix} \Sigma Wx^2 & \Sigma Wxy & \Sigma Wx \\ \Sigma Wxy & \Sigma Wy^2 & \Sigma Wy \\ \Sigma Wx & \Sigma Wy & \Sigma W \end{bmatrix} \times \begin{bmatrix} A \\ B \\ C \end{bmatrix}$$

where W is the square of the weighting value, x is the along track coordinate in metres, y is the across track coordinate in metres, and R is the range value in metres. It was observed that better numerical stability is obtained if the origin of coordinates was taken as the point at which the slope is to be determined. Care also must be taken in calculating along and across track distances when near the start of the ground track, as very small angles are involved, and these may cause difficulties if the computer has limited numerical precision.

We then obtain the slope and direction of slope as follows :

$$dR/dS = \sqrt{(A^2 + B^2)}$$

$$\theta = \mathrm{Tan}^{-1}(B/A)$$

Data Organization

In this study, a very simple means of organizing the data has been used, for a proof of the concept. The mean start and end of the data segments are found, and are used to define a great circle which is used as the mean track. Offsets from this track are then calculated and used as inputs to the equations above.

In a real implementation this would not be feasible, as the satellite may not remain in one repeat pattern for sufficient repeats to give across track slopes. A further disadvantage would be that the additional data available at crossing points would not be used. This is being addressed in current studies, and will probably require the use of a hierarchical cell structured database, possibly based on tesseral addressing (Diaz, 1986 and Lusby-Taylor, 1986). Such a database organization would allow a correction to be made even in regions where the satellite tracks were sparse, although a significant improvement in quality is expected for more repeat cycles.

Accuracy and Quality Assurance

The accuracy of this method is dependent on the accuracy of individual range measurements, and on the accuracy with which the satellite position is known in three dimensions. All these can be regarded as inaccuracies in the range figure. A rough approximation of the error in the estimate of dR/dS induced by range errors is :

$$\epsilon_g = \epsilon_n/(D.\sqrt{N})$$

Where ϵ_g is the error in the estimated gradient, ϵ_n is the error in the range measurement, D is a characteristic distance over which the estimate takes place, and N is the number of measurements used. Thus, if we take reasonable approximations for the case of Seasat and ERS-1 of $D \approx 1$km, $\epsilon_n \approx 1$m and $N = 30$, then $\epsilon_g \approx 0.0002$.

In addition to this overall estimate of the errors, it would also be possible to estimate the errors due to the fitting of a plane to the data, using standard statistical techniques.

As well as random errors in range determinations, errors are introduced by the breakdown of the assumptions of the model used to estimate the slope at a particular point. The particular assumption that will cause the breakdown of the method is that of continuity of slope. dR/dS is discontinuous if the radius of curvature of the surface is less than the range to the satellite (Harrison, 1971), and this causes the assumption that the surface can be modelled by a plane to break down. It is possible to be aware of the breakdown of this assumption by calculating the distance between successive relocated points along the same track, and normalizing them to the distance between the un-relocated points. The resulting figure is an indication of the quality of the result, with values increasingly greater than 1 indicating increasing likelihood of the presence of a discontinuity in the range surface. Values less than 1 occur at the crests of rises in the surface, and indicate a high degree of confidence in the result.

Test procedure

In order to test the slope correction algorithm, a test procedure was devised. Test input data were created by finding the minimum range from a series of satellite positions to the surface of the ice-cap model defined by Drewry et al (1985), within the bound of the altimeter's beam limit. In order to overcome the problem of local minima in the range, five different starting points were used for each determination, and the lowest minimum was used as the test data. In addition to returning the minimum range, the position on the ground of that minimum was also returned, so both the vertical and horizontal accuracy of the algorithm could be tested.

This procedure, using modelled data, was used in preference to using Seasat data because of the lack of ground truth for Seasat surface elevations.

After producing these test data, three forms of error were introduced into the data, to simulate inaccuracies in both the determination of the satellite orbit, and in the range determination. In the case of errors in the orbit determination, it was assumed that the errors are uncorrelated between tracks, but constant along a particular track. They were applied as a standard deviation for the radial component of the orbit, and a standard deviation for the horizontal component of the satellite track. The range errors were assumed to be uncorrelated both along track and between tracks, and were applied by specifying a standard deviation, as for the other errors.

The ice model was chosen as the test surface in order to produce data most like that which will be obtained over real ice caps. However, this model surface includes such features as cliffs (at the margin of the ice shelves), steep slopes (near the margin of the ice cap) and a wide variety of gently undulating terrain at various distances from the centre of the ice cap.

Test Results

Three sets of test data were used. These were chosen to illustrate the performance of the algorithm over the three different types of terrain mentioned above. In addition to this, for each location, the test was run with negligible error (not zero, because of computing problems), and with errors similar to those expected from the actual mission, i.e. range measurement error with a standard deviation of 0.1 m, radial component of the orbit with a standard deviation of 0.25 m and horizontal component of the orbit with a standard deviation of 5.0 m. These results are presented as a series of histograms and cumulative frequency diagrams (fig. 3a-c).

All the results show a peaked distribution, with a mode at or very near zero. The distributions tend to be skewed, as a result of errors in the slope estimation which is systematically too low. There is also a broad tail on the distribution, which is explained by errors in the slope estimation due to discontinuities in the range rate which occur on concave upwards portions of the surface.

In all the figures 3a-c, it will be seen that the cumulative frequency diagrams often do not reach 100%. This is because the data have been truncated to 2.0 m, to make the diagrams uniform. This only loses a significant amount of data in figure 3c. In these figures, there is a second peak at a very large error (80 m!), which is not shown. This arises because of severe underestimation of the slope at a major discontinuity. It would also occur for a cliff higher than that used in figure 3b, which was 50 m high.

The shape of the distribution is such that the standard deviation of the errors is a poor estimator of the likelihood of a result with a particular error. The cumulative frequency curves portray the results more clearly, and it can be seen that, in the worst case, 70-80% of the results lie within 1 m of the correct value, and that more typical results would be 70% within .5 m. The error figures used in these tests are rather high, and the cross track error for the final orbit determination is unlikely to be as high as 5 m. Thus, these errors give a slightly pessimistic view of the accuracy likely to be attained with real data.

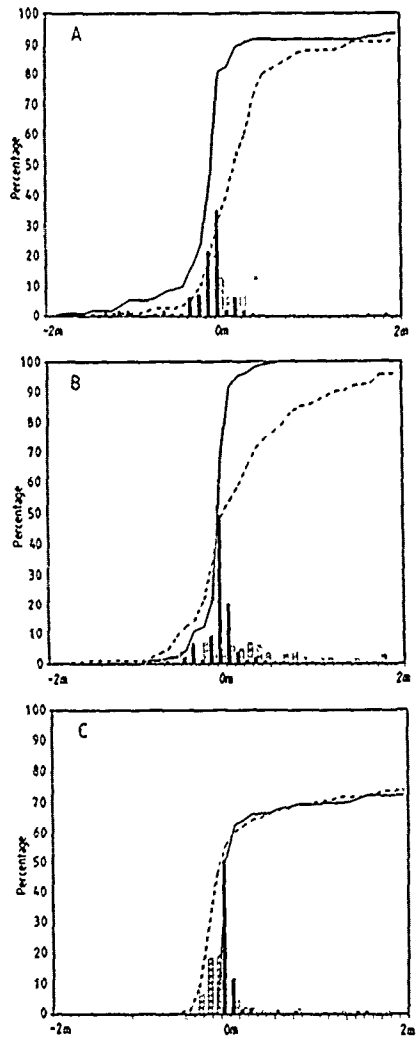


Fig 3 Error distributions for three test cases. In all cases, the full line and the filled bars are for test data with no added noise, and the empty bars and dashed lines are for data with 0.1 m standard deviation in range determination, 0.25 m in the satellite elevation and 5 m in the satellite position. A is for an undulating surface, B is for a surface with a 50 m cliff, and C is for a surface with a very steep slope.

Future Improvements

It is clear that the worst source of error remaining in the method is poor estimation of the surface slope in

the vicinity of discontinuities in the range rate, caused by concave upwards surfaces. This can be overcome by detecting such discontinuities, and by rejecting data from the side of them opposite from the sub-satellite point for which the correction is being computed. Work is in progress to provide an algorithm for detecting such discontinuities. When this is implemented, it is hoped that the error distribution will be considerably narrowed, and be entirely due to numerical errors and to errors caused by poor slope estimation caused by either uncertainties in the measured parameters or by unavoidable errors in the slope estimate.

Conclusions

A method has been presented for correcting the slope induced errors in satellite altimeter data. This method is geometrically correct, and is capable of producing both along and across track corrections, without any prior assumptions concerning the across track slope, or any iterative techniques. Methods are available for assessing the accuracy of the result, and flagging invalid results.

Acknowledgements

This work has taken place under a contract placed with SPRI by the UK ERS-1 Product Support Team, to develop algorithms for the processing of data from the ERS-1 satellite. I must also thank Dr. W.G. Rees for many helpful discussions, and Dr. E. Novotny for her aid in producing test data.

References

1. Brenner A.C., Bindschadler R.A., Thomas R.H., Zwally H.J., 1983, Slope induced errors in radar altimetry over continental ice sheets. Journal of Geophysical Research, Vol 88, No C3, pp 1617-1623.

2. Diaz, B.M., 1986, Tesseral addressing and arithmetic - overview and theory. In Spatial Data Processing using Tesseral Methods, Ed. Diaz, B.M. and Bell, S.B.M.. Natural Environment Research Council, U.K., pp 1-10.

3. Drewry D.J., McIntyre N.F., Cooper A.P.R., 1985, The Antarctic Ice sheet: a surface model for satellite altimeter studies. In Woldenburg M.J. Ed. Models in Geomorphology. Allen and Unwin 1985.

4. Harrison C.H., 1970, Reconstruction of subglacial relief from radio echo sounding records. Geophysics, Vol 35, No. 6, pp 1099-1115.

5. Harrison ' C.H., 1971, Radio echo sounding: focussing effects in wavy strata. Geophysical journal of the Royal Astronomical society. 24, pp 383-400.

6. Lusby-Taylor, C., 1986, A rectangular tesselation with computational and database advantages. In Spatial Data Processing using Tesseral Methods, Ed. Diaz, B.M. and Bell, S.B.M.. Natural Environment Research Council, U.K., pp 391-402.

COHERENT RADAR CONTOUR MAPPING OF ICE STREAM THICKNESS

C. H. Davis, R. K. Moore, G. Raju, and W. Xin

Radar Systems and Remote Sensing Laboratory
University of Kansas Center for Research, Inc.
2291 Irving Hill Road
Lawrence KS 66045-2969, USA
913/864-4835
FAX: 913/864-7789, TELEX: 706352, OMNET: KANSAS.U.RSL

Research supported by National Science Foundation Division of Polar Programs: Grant DPP-8300450

ABSTRACT

Because ice is transparent to electromagnetic waves with frequencies less than 300 MHz, low-frequency radar can provide a rapid image of a glacier from the surface to bedrock. Such a radar can be used to obtain a contour map of the glacier's ice thickness. Glaciologists use this contour map to calculate the total mass of the ice at specific points. They then can use this information along with strain measurements to determine the shear stress at the base of the glacier.

Previous ice-sheet "radio-echo-sounding" systems used vacuum-tube technology and did not use coherent integration. We developed a 150-MHz radar system designed specifically for probing the continental ice of Antarctica. This system uses coherent integration and solid-state technology. Thus, only 20 W of peak output power can produce signal-to-noise ratios comparable to those of older high-power systems.

We used our system to conduct a survey on Ice Stream B in West Antarctica in December 1988. The survey concentrated on a 2-km by 10-km grid surveyed by Ian Whillans of The Ohio State University. The measurements allowed us to produce an ice-thickness map with a 10-m contour interval. This contour map will be used by glaciologists to estimate the basal shear stress of the glacier and thereby estimate future glacial movement.

INTRODUCTION

Glaciologists in the Siple Coast Project are studying Ice stream B in West Antarctica because it is moving at a relatively high speed even though its surface gradient is very small. In the 1988-89 field season a team of scientists from several organizations conducted measurements on ice stream B in an ongoing effort to determine the glacial dynamics. Measurements made over a period of a decade will allow observations of short term variations within the West Antarctic ice sheet that might be associated with a global warming trend.

In conjunction with this effort, our Antarctic radar was used to conduct a survey on a strain grid at the Upstream B field camp (83° 28'S, 138° 5'W) in December 1988. This grid was set up in prior years by The Ohio State group under Prof. Ian Whillans to allow multi-year monitoring of the ice movement. The ice thickness along the strain grid can be determined from the radar traverses. These data allow calculation of the mass balance of the grid, which, when combined with the strain measurements, will provide an estimate of the shear stress at the ice/bedrock interface.

DESCRIPTION OF RADAR SYSTEM

The Coherent Antarctic Radar Depth Sounder (CARDS) has been under development for the past three years. It consists of an RF section (actually separate transmitter and receiver units), a digital data-handling section, and a microcomputer with storage on Bernoulli-box cartridges.

RF Section

The radar operates at a center frequency of 150 MHz. Its system design provides for high signal-to-noise sensitivty with moderately low RF transmitter power. This is achieved by using a pulse expansion/compression technique and full coherent integration (stacking) of the radar returns from multiple pulses. Figure 1 shows the overall system in block-diagram form, and the RF system is depicted in Fig. 2.

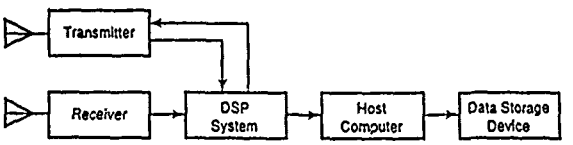


Fig. 1 Overall System Block Diagram

The basic pulse is about 60 ns wide, which corresponds to a range resolution of 5 m in ice. It modulates a stable carrier at 150 MHz (wavelength = 2 m in air). The carrier signal is obtained by frequency-multiplying the output of a stable crystal oscillator operated at 9.275 MHz. The modulated pulse is expanded to 1.6 μs in a surface-acoustic-wave (SAW) expander. The expanded pulse is amplified, gated to reduce range sidelobes, and amplified further in several stages to a peak RF power of 20 watts. A pair of four-element dipole arrays serve as the transmitting and receiving antennas. Because the wavelength is 1.75 m in ice, the antennas are a reasonable size for radar traverses on the ground and mounting on a small twin-engine aircraft.

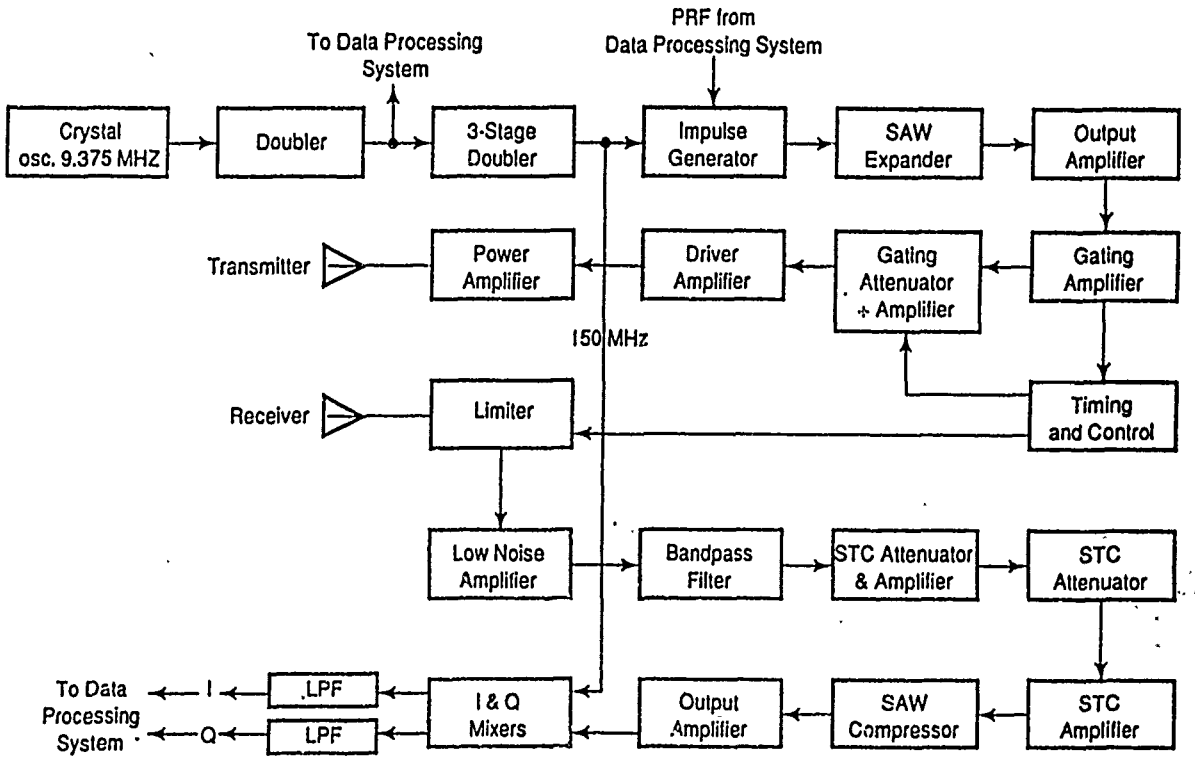


Fig. 2 Functional Block Diagram of RF Section

The receiver includes a low-noise preamplifier and an RF power limiter to protect the preamplifier from overdriving. This is followed by several RF amplifier stages to provide adequate power to drive the SAW that compresses the expanded pulse. The expansion/compression process reduces the peak power requirements by spreading the transmitted energy over 1.6 μs instead of 60 ns. The 5-m range resolution is preserved because of the complementary compression back to the 60-ns basic pulse.

Voltage-controlled attenuators are located between the amplifier stages. Their attenuation can be changed by a time-varying voltage to vary the sensitivity of the receiver. This Sensitivity Time Control (STC) is necessary to prevent the receiver from saturating due to the strong near-range returns, while maintaining the maximum sensitivity for the weaker long-range returns from the bottom of the glacier. After the pulse is compressed, it is amplified and heterodyned to baseband in in-phase and quadrature mixers. The resulting signals are digitized in a pair of A/D converters for further processing and coherent and incoherent integration in the digital subsystem.

Digital Section

A block diagram of the digital part of the system is shown in Fig. 3. The hard-wired digital controller

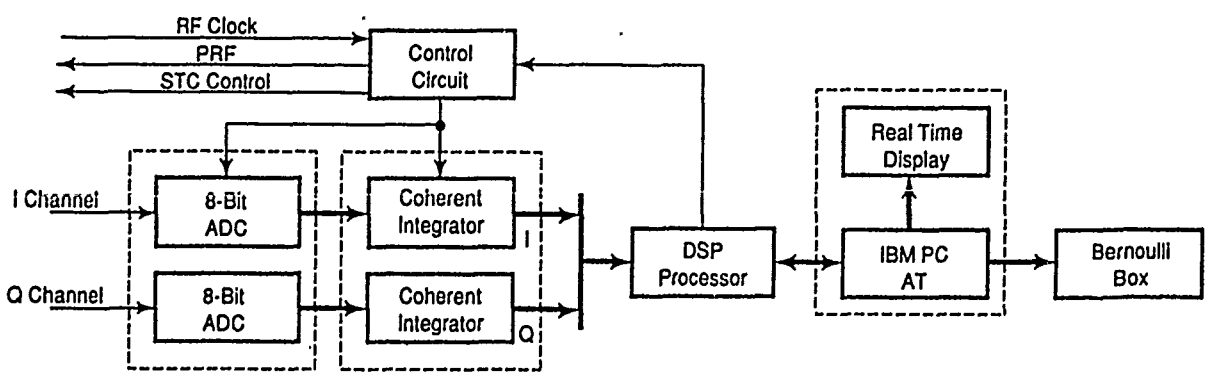


Fig. 3 Data Processing System

coordinates the operation of the entire radar system, and hence insures fully coherent operation of the system. The RF portion is synchronized by the PRF signal that triggers the radar transmitter and the carrier derived from the same oscillator. The controller also provides the digital signals to the STC gain-control modules and triggers the A/D converters at the proper times.

The I and Q analog signals from the radar system are digitized by two 20-MHz, 8-bit, A/D converters. The digital outputs go to two hardwired high-speed digital coherent integrators. High-speed RAM provides circular buffers in which the current results of the coherent integration are stored. As each new sample of the return pulse is available, it is added to the contents of the corresponding cell in the buffer.

When one set of coherent integrations is complete, the next set utilizes the second pair of circular RAM buffers. While the coherent integration continues in the second buffers, the contents of the first set is transferred to the TI ADSP-2100 signal-processing microcomputer. The DSP computer performs additional coherent integrations, square-law detection, and incoherent integrations if desired. The results pass to the Compaq PC for display and transfer to the Bernoulli box external storage device. Data acquisition is continuous because all of the above operations proceed in parallel and each operation requires less time than the preceding one.

The PC provides for program loading and user interaction with the radar. The operator selects the radar parameters and the type of signal processing from menus displayed on the screen. Provision is also made for the display of selected signals throughout the system for trouble shooting, as well as insuring that valid data are being collected.

Software

The software running the Compaq Portable II computer permits the user to set the operating parameters of the system by a series of menus. Reasonable default parameters are provided to expedite the system setup for normal operation.

The software provides for real-time control of the operation of the entire radar system. In addition, it allows input of inertial navigation and flight-condition data at intervals as short as two seconds. This information is embedded into the data stream for later use in reducing the radar data to maps. Markers can also be put into the data stream by pressing a key on the keyboard of the computer to correlate with exogenous events.

The software provides several display modes to allow examination of the data while the system is operating. This insures that the system is operating properly and that valid data are being collected. The operator can observe the I and Q data after coherent integration and square law detection on an A-scope display. A gray-scale display shows amplitude vs. depth and distance traveled along the track.

An NEC multi-sync, high-resolution color display can be interfaced to the system to provide for more informative displays of the collected data in non-real time. The software provides for the selection and display of segments of a radar traverse. The amplitude of the signal is encoded into a color map. This display shows depth vs. distance along the track. Figure 4 shows an example of such a display reproduced in black and white. The display window can be scanned through the data set to examine different segments of a traverse. The coordinates of features of interest can be extracted by means of a mouse-driven cursor and

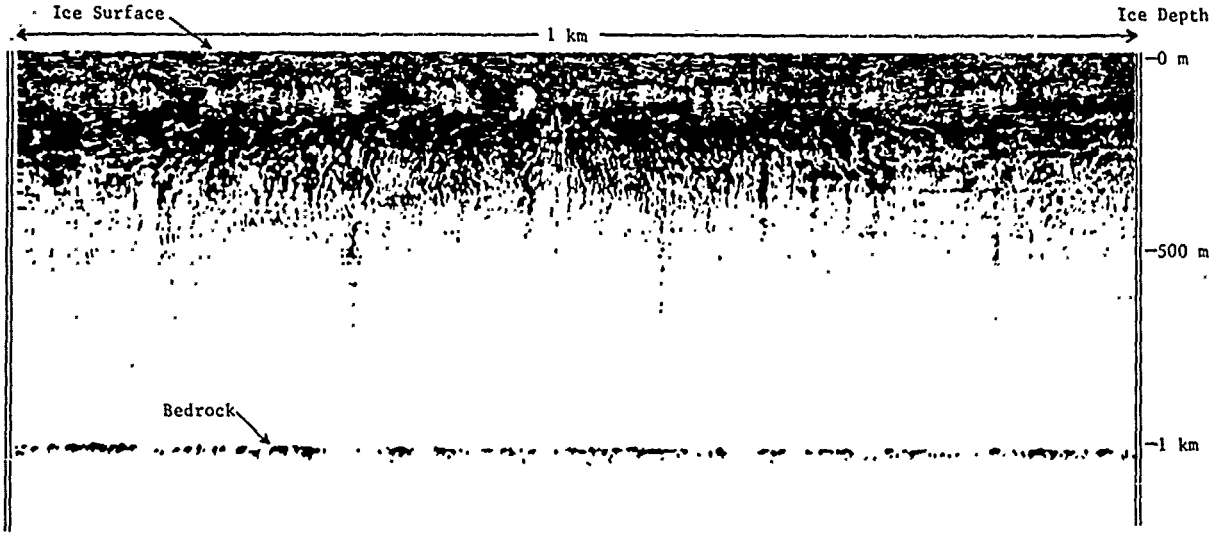


Fig. 4 Sample Along Track Distance vs. Ice Depth "Image" of Ice Stream B

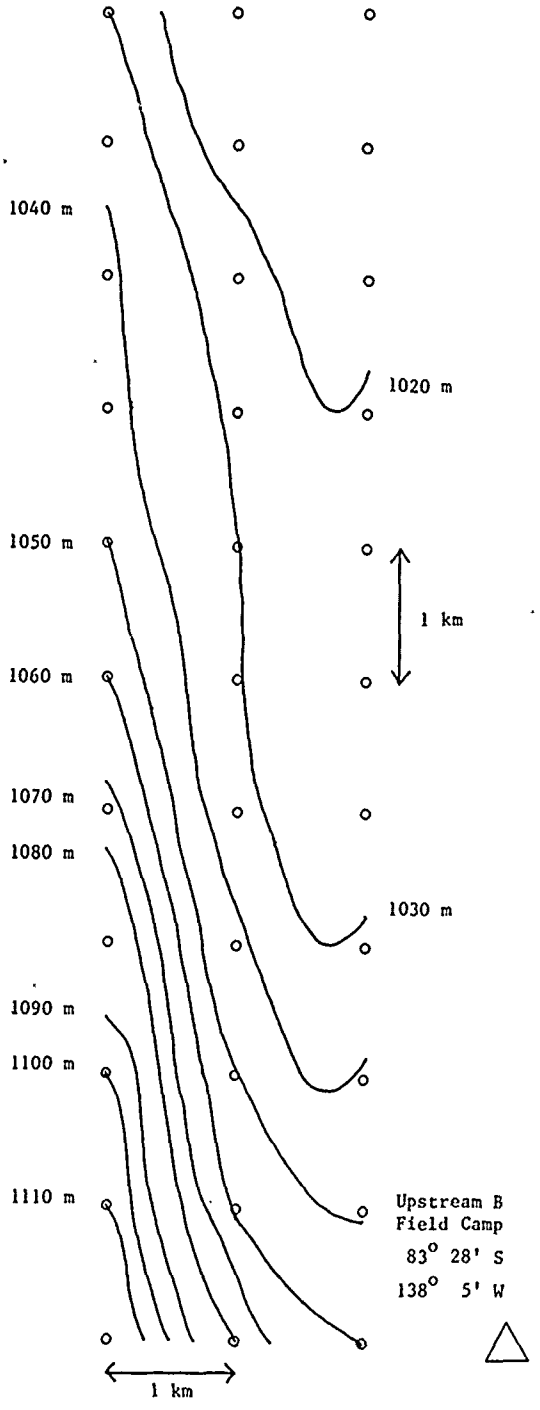


Fig. 5 Ten-meter Contour Interval of Ice Thickness

written to a new data file. Contour maps of the bottom echoes can then be produced from this data file.

Field Work--1988

We conducted a radar survey from December 1-15, 1988, at the Upstream B field camp on ice stream B in West Antarctica. The radar was rack mounted inside a Spryte tracked snow vehicle. The antennas were mounted on two sleds that were pulled behind the vehicle. A generator on another sled provided power to the system during the traverses.

The primary test area was a 2-km by 10-km strain grid that had been used previously by The Ohio State University personnel under Prof. Ian Whillans. We traversed this grid in both upstream and across-stream directions. Many of the measurements were repeated to verify the system performance.

RESULTS

No system downtime occurred during the entire survey of over 150 km of traverses. Traverses were repeated to verify the system reliability, and in all cases field analysis of the images on duplicate tracks were identical.

The primary purpose of the survey was to provide the ice-thickness data for use by Prof. Whillans and other glaciologists. However, the images show unusual features within the glacier itself. The image shown in Fig. 4 has several circular regions in the first 300 m that exhibit no scattering. Normally the first 300 m is characterized by continuous scattering from inhomogeneities in the ice and from buried crevasses. These circular regions must be concentrations of homogenous material that does not allow volume scattering. This is an unexpected and so far unexplained phenomenon. Other images show layered regions of scattering that undergo substantial change over a period of several hundred meters along the glacier.

The flexibility incorporated into the software allowed processing of the data in the field. As a result, we produced the contour map of the ice thickness for the entire 2-km by 10-km strain grid in the field. The grid and resulting 10-m contour map of the ice thickness is shown in Fig. 5. This contour map was analyzed by glaciologists in the field and is now being used to provide the mass balance information needed to produce an accurate estimate of the shear stress at the base of the glacier.

CONCLUSION

The field operation of the University of Kansas Coherent Antarctic Radar Depth Sounder was successfully demonstrated during the 1988-89 Antarctic field season. The images produced by the radar traverses showed many interesting features within the glacier itself, which were unexpected. The flexibility of the system software allowed the production of the ice-thickness contour map in the field for immediate analysis.

Cover Photo:

Geocoded Landsat-SPOT composite of Vancouver, British Columbia, Canada.
Produced by MacDonald Dettwiler and Associates, Richmond, B.C., Canada.
Landsat TM imagery courtesy of Canada Centre for Remote Sensing,
Energy Mines and Resources Canada.
SPOT imagery ©CNES, 1987.