ADA 118236

Effect of the Ionosphere on Radiowave Systems

Editor-in-Chief:
John M. Goodman

Assoc. Editors:

Frank D. Clarke
Jules Aarons

Based on
Ionospheric Effects Symposium

DTIC
AUG 1 1 1982
H

Held in
Old Town, Alexandria, Va.
April 14-16, 1981

Jointly Sponsored by the

Naval Research Laboratory
Office of Naval Research

and the

Air Force Geophysics Laboratory

APPROVED FOR PUBLIC RELEASE
DISTRIBUTION UNLIMITED

ORGANIZING COMMITTEE

Chairmen: Dr. John M. Goodman
Dr. Jules Aarons

Coordinator: Mr. F.D. Clarke

Secretary: Mrs. R.A. Smithson

Session Chairmen:	Session No.
Dr. J.M. Goodman	I
Dr. J. Kelso	II
Dr. J.V. Evans	IIIA
Dr. K. Davies	IIIB
Dr. E. Fremouw	IVA
Mr. G. Lane	IVB
Dr. J. Aarons	V
Dr. S. Ossakow	VI

Accession For
NTIS GRA&I
DTIC TAB
Unannounced
Justification
By
Distribution/
Availability Codes
Dist Special
A

DTIC
COPY
INSPECTED
2

Library of Congress Catalog Card No. 82-600510.

"A" Statement
8/5/82

PREFACE

IES '81 is the third Ionospheric Effects Symposium to be sponsored by the Naval Research Laboratory, the first having been held in 1975 and the second in 1978. The most recent symposium was co-sponsored by the Office of Naval Research and the Air Force Geophysics Laboratory. I am especially pleased that AFGL officially participated in IES '81, inasmuch as it has been my aim since the inauguration of these conferences that all DoD research and development activities play a key role.

The purpose of this symposium, as in the two previous IES conferences, was to improve the information transfer between system architects, managers, and designers on the one hand, and ionospheric physicists and propagation specialists on the other hand. Although the military (DoD) interest associated with various topics presented at the conference was transparent, the commercial and scientific research areas were also in evidence.

The conference was keynoted by Dr. Hans Mark, who, fresh from his participation in the first launch of the space shuttle, provided the attendees with his perceptions of the future direction of the space program. A special address was presented by J.N. Birch that highlighted the ionospheric research needs of present and future-planned DoD systems. A banquet, held on the evening of April 15, had as its guest speaker Prof. J.A. Van Allen, who discussed "The Magnetospheres of the Planets."

The conference itself covered various topics of current interest to the ionospheric research community. Session topics included the following: "Ionospheric Modification", chaired by J.M. Goodman; "General Reviews and Total Electron Content", chaired by J. Kelso; "Equatorial Scintillation Studies", chaired by J. V. Evans; "HF Propagation/Remote Sensing", chaired by K. Davies; "High-Latitude Scintillation", chaired by E. Fremouw; "Sub-HF Propagation and System Effects", chaired by G. Lane; "Ionospheric and Propagation Models", chaired by J. Aarons; and "Future Plans and Programs", chaired by S. Ossakow.

One of the areas of interest emphasized in the conference was ionospheric modification. There were 13 papers presented on this topic alone. The papers dealt with rocket plume effects, chemical releases, optical diagnostics, in-situ active experiments, ionospheric heating and its various manifestations, and possible applications of modification to the communication research community.

Several review papers were presented at the conference, including "Recent Developments in Artificial Ionospheric Heating", by C.M. Rush; "Ionospheric Predictions—A Review of the State of the Art", by K. Davies; "New Forecasting Methods of the Intensity and Time-Development of Geomagnetic and Ionospheric Storms", by S.I. Akasofu; "Recent High-Latitude Improvements in a Computer-Based Scintillation Model", by E.J. Fremouw and J.M. Lansinger; and "Effects of the Ionosphere on HF Radar Propagation", by D.B. Trizna and J.M. Headrick.

I would like to thank all of the session chairmen for their assistance in the program planning and for the careful review of papers which were submitted. Dr. Aarons, the co-chairman of IES '81, performed a yeoman's job in filtering through the oversupply of abstracts which were submitted to the organizers, and his efforts translated directly into a balanced and workable conference agenda. The organizing committee was quite pleased with the response to the call for papers, and it is indeed unfortunate that many excellent papers could not be provided space, and that a number of others had to be combined. In all, 73 papers are contained in this conference proceedings. In addition to 5 review papers, Dr. Aarons and I have reviewed "The

Radiowave Propagation Environment—Science and Technology Objectives for the 80's", the purpose of which was to set a tone for the conference. The symposium was opened by Dr. Alan Berman, the NRL Director of Research, and the text of his address is contained within this volume. I would like to acknowledge the continuous support and encouragement for the series of IES Conferences which have been provided by Dr. Berman and others at NRL, including Dr. T.P. Coffey and Dr. P. Mange. In addition, Dr. H. Mullaney and Dr. J. Dimmock are thanked for promulgation of the ONR sponsorship.

It is not possible to thank all individuals who assisted me during the course of the IES '81 activity. However, as in the past, a major share of credit goes to Mr. F.D. Clarke, the symposium coordinator, who orchestrated the entire affair. He was assisted by an able and competent staff including Mrs. Rhea Smithson, who was the official conference secretary, Mrs. Gailyn Nocente and Mrs. Jane Clarke. The technical Information Division at NRL is acknowledged for document preparation and printing services. Mrs. D. Robbins and her staff at TID produced an excellent preprint document which has been widely acclaimed by the attendees, and they also assisted in the layout and editing of this final volume. The chief projectionist for IES '81 was Mr. Dominic Panciarelli. The staff of the Ramada Inn in Old Town, Alexandria was extremely helpful in accommodating the special needs specified by the IES '81 organizers, including the symposium banquet held at the conclusion of the technical agenda on the second day.

Finally, I would like to thank all of the attendees of IES '81 for making the conference a memorable as well as worthwhile event.

Frank Clarke, Jules Aarons and I would like to congratulate the contributors whose papers are contained within this volume. On the whole, the contents reflect the state-of-the-art in ionospheric research and system effects. Hopefully this book will be a useful reference text for researchers and system designers alike.

John M. Goodman
Co-chairman and Editor-in-Chief
IES '81

TABLE OF CONTENTS

Preface iii

Welcoming Address xi
Alan Berman

Introductory Paper

The Radiowave Propagation Environment — Science and Technology Objectives for the 80's xxi
John M. Goodman and Jules Aarons

SESSION I
Ionospheric Modification — I

Ionospheric Effects of Rocket Exhaust Products — Skylab and HEAO-C 1
J. Zinn, C.D. Sutherland, L.M. Duncan, and S.N. Stone

Rocket Exhaust Expansion at High Altitudes and Ionospheric Holes 8
M.H Reilly

Occulation of the ATS-3 Satellite by the AVEFRIA Barium Ion Cloud 15
T.J. Fitzgerald, D.J. Simons and M.B. Pongratz

In Situ Active Experiment Techniques to Study the Ionosphere 23
M.B. Pongratz, R.A. Jeffries, D.J. Simons and G.M. Smith

Stimulation of ULF Magnetic Pulsations by Rapid Change in Ionospheric Conductivity 28
P.B. Morris, T.J. Rosenberg, and L.J. Lanzerotti

Enhancement of Radio Propagation with Reactive Gas Releases into the Ionosphere 38
P.A. Bernhardt

Recent Developments in Artificial Ionospheric Heating 50
C.M. Rush and K.C. Allen

Ionospheric Modification by High-Power Radio Waves 62
L.M. Duncan

Artificial Irregularities Generated by Ionospheric Heating and Their Effects on Transionospheric Propagation 72
Santimay Basu, Sunanda Basu, S. Ganguly, J.A. Klobuchar, and C.M. Rush

Vertical Incidence Pulse Absorption Measurements During High Power Radio Wave Heating of the D-Region 81
A.A. Tomko, A.J. Ferraro, and H.S. Lee

SESSION II
Ionospheric Modification II, General Reviews, TEC

Wireless Generation of ELF/VLF Radiation in the Ionosphere 91
C.L. Chang, V. Tripathi. K. Papadopoulos, J. Fedder, P.J. Palmadesso, and S.L. Ossakow

Ionospheric Heating Effects on LF Propagation 100
F.J. Kelly and D.N. Madhusudhana

Project HIPAS — High Power Active Stimulation of the Auroral Plasma 105
A.W. Wong, J. Roederer, R. Hunsucker, G. Dimonte, P. Viezbecke, J. Gregorio, and J. Santoru

Ionospheric Predictions — A Review of the State-of-the-Art 110
K. Davies

New Forecasting Methods of the Intensity and Time Development of Geomagnetic and Ionospheric Storms 132
S.I. Akasofu

Recent High-Latitude Improvements in a Computer-Based Scintillation Model 141
E.J. Fremouw and J.M. Lansinger

Enhancements and Sharp Depletions of Total Electron Content in the Nighttime Equatorial Ionosphere 155
D.E. Donatelli, P.M. Walsh, R.S. Allen and J.A. Klobuchar

Some Results of Ionospheric Total Electron Content and Scintillation Observations at Lunping 163
Yinn-Nien Huang

SESSION IIIA
Equatorial Scintillation Studies

Scintillation of Satellite Signals at Guam for Two Elevation Angles and Two Frequencies 174
M.R. Paulson

C-Band Ionospheric Scintillation Measurements at Hong Kong Earth Station During the Peak of Solar Activities in Sunspot Cycle 21 181
D.J. Fang

Microwave Equatorial Scintillation Intensity During Solar Maximum 193
J. Aarons, H.E. Whitney, E. MacKenzie, and S. Basu

Ionospheric Scintillation Observations at NATAL 202
K.C. Yeh, J.P. Mullen, J.R. Medeiros, R.F. da Silva, and R.T. Medeiros

VHF Scintillation Near the Equatorial Anomaly Crest in the Indian Sector 210
A. Dasgupta and A. Maitra

Cumulative Distribution of Scintillation Index 215
A.I. Trivedi, M.R. Deshpande, and H.O. Vats

Predicting the Equatorial Spread-F/Scintillation Environment by Theoretical and Numerical Simulation Techniques 219
S.T. Zalesak, S.L. Ossakow, and P.K. Chaturvedi

The STP/S3-4 Satellite Experiment: Equatorial F-Region Irregularities 228
M. Singh, E.P. Szuszczewicz, and J.C. Holmes

The Global Distribution of Equatorial Spread-F in the Topside Ionosphere During Equinoctial Periods 235
E.R. Young, W.J. Burke, F.J. Rich, R.C. Sagalyn, and M. Smiddy

Remote Detection of the Maximum Altitude of Equatorial Ionospheric Plasma Bubbles 244
R.F. Benson

Amplitude and Phase Scintillation Measurements Using the Global Positioning System 253
J.A. Klobuchar, C.L. Rino, and M.D. Cousins

SESSION IIIB
HF Propagation/Remote Sensing

Ionospheric Effects on HF Over-The-Horizon Radar 262
D.B. Trizna and J.M. Headrick

Measurement of Sea and Ice Backscatter Reflectivity Using an OTH Radar System .. 273
W.F. Ring and G.S. Sales

Adaptive Utilization of HF Paths: A Way to Cope with the Ionospheric Limitations Affecting Mid-Latitude and Transauroral Short-Wave Links 282
A.K. Gupta and M.D. Grossi

Multipath Measurements in the Athens-Salisbury T.E.P. Link 297
G. Stephanou, C. Caroubalos, and N. Corallis

An Empirical Model for the Propability Distribution of the Lowest Observed Frequency 305
D.B. Sailors

Accounting for Ionospheric Variability and Irregularity in High Frequency Direction Finding 315
M.G. Heaps

Communications in the "Ship Zone" Using Back Scattering from Meteor Trails and E-Layer Turbulences 325
H.E.R. Jones

Seasonal Variations in the Aurora Ionosphere 339
B.W. Reinisch, J. Tang, and H. Xuegin

Remote Sensing of High-Latitude Ionization Profiles by Ground-Based and Spaceborne Instrumentation 345
R.R. Vondrak

Application of X-Ray Imaging Techniques to Auroral Monitoring 354
D.M. Rust and P. Burstein

Ionospheric and Auroral Measurements from Space Using Vacuum Ultraviolet Emission 364
R.E. Huffman, D.E. Paulsen, F.J. LeBlanc, and J.C. Larrabee

SESSION IVA
High Latitude Scintillation Studies II

The Role of Curvature of Ionospheric Irregularities in Transequatorial Propagation at VHF 372
J.A. Ferguson

Scintillation Studies in the Ionospheric Trough 381
J.W. MacDougall and J.A. Koehler

Occurrence of F-Layer Irregularities in the Polar Cap 391
A.L. Johnson, J. Aarons, J. Buchau, J.P. Mullen, H.E. Whitney, and E. Weber

The Current Convective Instability and its Relation to Diffuse Auroral Scintillation Causing F-Region Irregularities 397
M.J. Keskinen, P.K. Chaturvedi, S.L. Ossakow, and B.E. McDonald

Observations Concerning the Relationships Between Electron Flux and Electric Field Variations and Thermal Plasma Irregularities in the Polar Cap 407
W.J. Burke and M. Smiddy

The Night-Time Auroral E-Layer: Particle Production, Latitudinal and Longitudinal Structure and Dynamics 417
J.R. Sharber and J.A. Whalen

The STP/S3-4 Satellite Experiment: High Latitude 431
P. Rodriguez, M. Singh, D.N. Walker, E.P. Szuszczewicz, and J.C. Holmes

The Role of In-Situ Measurements in Scintillation Modelling 440
S. Basu, Santimay Basu, and W.B. Hanson

SESSION IVB
Sub-HF Propagation/Radar and System Effects

MF Skywave Propagation 450
B.G. Melander and R.H. West

Lateral Bending Effect at the Ionospheric Height Transition 460
D. Mannheimer

Results of a Differential Omega Experiment 470
F.G. Karkalik, L.E. Abrams, and T.M. Watt

Modelling the C and D-Regions 483
W. Swider

The Application of an Ionospheric Model for Predicting
Radar-Propagation Effects 490
G.H. Millman

Anomalous Ionospheric Refraction Associated with the Auroral Zone 497
J.V. Evans and R.H. Wand

The Combined Effect of Faraday and Squint Rotation for an
Electronically Scanned, Satellite-Borne Radar 508
G.B. Goldstein

Limitations of VHF Tracking Caused by Ionospheric Correction Uncertainties 515
S. Pallaschke

Temporal Behavior of Pulses After Propagating Through a Turbulent Ionosphere 525
C. Yang and K.C. Yeh

SESSION V
Ionospheric/Propagation Models: Prediction and Forecasting

Prophet — An Emerging HF Prediction Technology 534
R.B. Rose

HF Propagation Assessment Studies Over Paths in the Atlantic 543
D.R. Uffelman

A Global Model for Wideband HF Skywave Propagation 553
A. Málaga

The Use of the International Reference Ionosphere (IRI-78) Models for
Calculating Long Path Fields at VI F 565
F.P. Snyder and J.A. Ferguson

A Semi-Empirical Model of the Ionospheric Electron Density 575
K.A. Pfitzer, S.J. Scotti, and W.P. Olson

Empirical F-Region Model Development Based on S3-1 Satellite Measurements 584
C.R. Philbrick

Simulated Propagation Effects on Transionospheric Radio Waves 591
K.C. Yeh and C.H. Liu

High-Latitude Ionospheric Model: First Step Towards a Predictive Capability 599
R.W. Schunk, W.J. Raitt, and J.J. Sojka

The Applicability of Using Theoretical Models to Improve Ionospheric Maps 610
D.N. Anderson, C.M. Rush, M. Pokempner, and F.G. Stewart

Satellite Orbits to Relieve Ionospheric Scintillation 620
P.F. Christopher and D.C. Rogers

Influence of Solar Sector Boundaries on Ionospheric Variability 630
M. Mendillo and G. Iuliano

SESSION VI
Future Plans and Programs

Waves in Space Plasmas Program 638
R.W. Fredricks and W.W.L. Taylor

DARN: A Radar Network for Studying the Large-Scale Structure and Dynamics of Ionospheric Electric Fields 645
R.A. Greenwald

An Advanced Mission to Map the Worldwide Topside Ionosphere 654
H.R. Mathwich, D.E. Aubert, A.F. Martz, Lt. D. Lewis, K. Bibl, and B. Reinisch

Global Real-Time Ionospheric Monitoring 664
J.W. Wright

WELCOMING ADDRESS

Dr. Alan Berman
Director of Research
Naval Research Laboratory

Good Morning. It is indeed a pleasure to welcome all of you to the 1981 Symposium on the Effect of the Ionosphere on Radiowave Systems. This is the third such conference which NRL has organized, and on this occasion, along with the Office of Naval Research, we are fortunate to co-sponsor the event with the Air Force Geophysics Laboratory.

The ionosphere has, of course, been the subject of considerable study for many decades, and there are many radiowave systems which are profoundly influenced by the natural plasma which surrounds the earth. I note that the brochure which advertised this event indicated that the following topics will be covered: effects on various systems including C^3I, navigation and positioning, direction finding, radar, and surveillance. In addition, specified ionospheric phenomena and techniques will be covered including: ionospheric modification, amplitude and phase scintillation, ionospheric forecasting and assessment, ionospheric models and total electron content. From an inspection of the abstracts received, it is clear that the papers to be presented are all of high quality, and a considerable effort has obviously been invested by the authors in the preparation of material furnished to the organizers of this conference.

One of the main objectives of IES '81, as was the case for its two predecessors, is development of a closer coupling between system architects and designers on the one hand, and propagation specialists and scientists on the other. This, I feel, makes IES unique and distinguishes it from otherwise excellent conferences sponsored by organizations such as AGU and URSI, for example. Speaking for the organizers of this conference, and NRL in particular, I extend my best wishes for a successful symposium. Thank you.

KEYNOTE SPEAKER

Dr. Hans Mark

BANQUET SPEAKER

Dr. James Van Allen

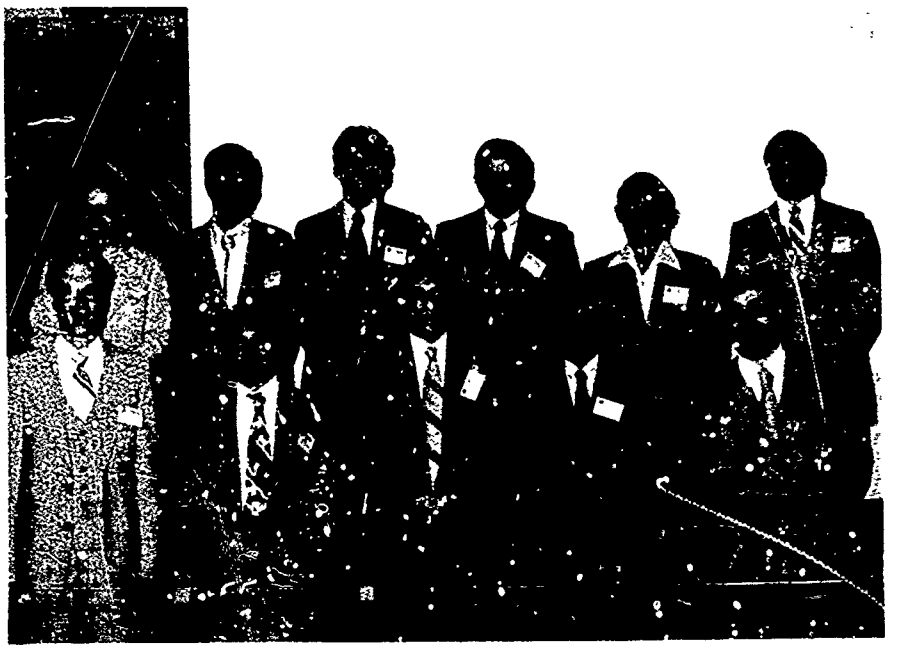
IES 1981 STAFF

Dr. J. Aarons
Chairman

Dr. J. M. Goodman
Chairman

Mr. F. D. Clarke
Coordinator

Mrs. R. A. Smithson
Secretary



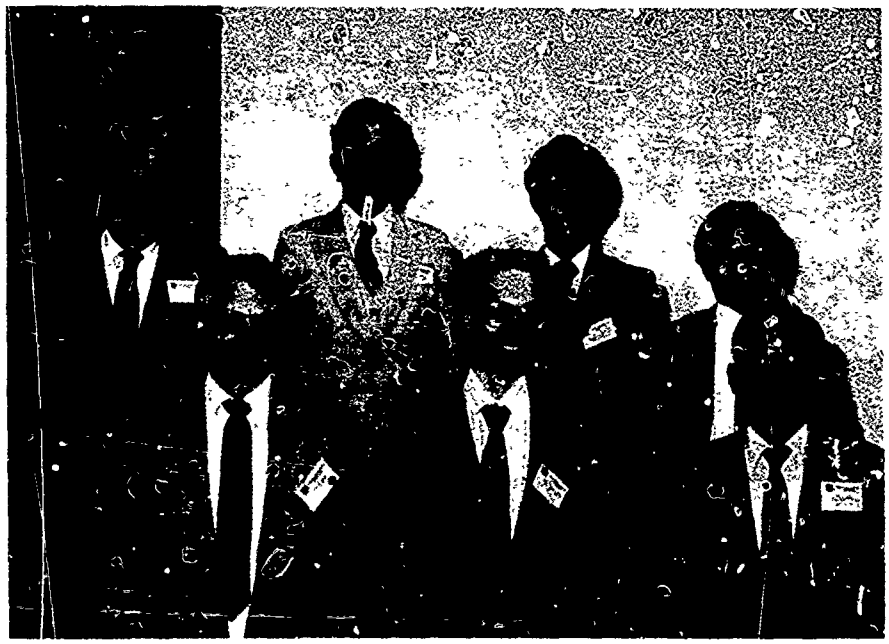
Back Row (Left-to-Right):

J.M. Goodman (session chairman), J. Zinn, M.H. Reilly, T.J. Fitzgerald, M.B. Pongratz, and P.B. Morris

Front Row (left-to-right):

P.A. Bernhardt, C.M. Rush, L.M. Ducan, S. Basu, and A.A. Tomko

GENERAL REVIEWS AND TOTAL ELECTRON CONTENT



Back Row (Left-to-Right):

J.M. Kelso (session chairman), F.J. Kelly, A.W. Wong, and S.I. Akasofu

Front Row (Left-to-Right):

E.J. Fremouw, R.S. Wolff, and Y. Huang

Back Row (Left-to-Right):

J.V. Evans (session chairman), M.R. Paulson, D.J. Fang, J. Aarons, J. Mullen, and A. dasGupta

Front Row (Left-to-Right):

H.O. Vats, S.T. Zalesak, M Singh, E.R. Young, R.F. Benson, and C.L. Rino

Back Row (Left-to-Right):

K. Davies (session chairman), D. Trizna, G.S. Sales, A.K. Gupta, D.B. Sailors, and M.G. Heaps

Front Row (Left-to-Right):

H.S. Jones, B. Reinisch, R.R. Vondrak, D.M. Rust, and R. Huffman

HIGH LATITUDE SCINTILLATION

Back Row (Left-to-Right):

E.J. Fremouw (session chairman), P. Rodriguez, J.R.Sharber, and A.L. Johnson

Front Row (Left-to-Right):

M.J. Keskinen, W.J. Burke, J.W. MacDougall, J.A. Ferguson, and S. Basu

Back Row (Left-to-Right):

G Lane (session chairman), B.G. Melander, D. Mannheimer, and G.H. Millman

Front Row (Left-to-Right):

G B Goldstein, S. Pallaschke, and C. Yang

IONOSPHERIC AND PROPAGATION MODELS

Back Row (Left-to-Right):

J. Aarons (session chairman), R.B. Rose, D.R. Uffelman, C.R. Philbrick, A. Malaga, and P. Rodriguez

Front Row (Left-to-Right):

R.W. Schunk, D.N. Anderson, P.F. Christopher, M. Mendillo, and M. Dubin

FUTURE PLANS AND PROGRAMS

(Left-to-Right):

S. Ossakow (session chairman), W.W.L. Taylor, R.A. Greenwald, H.R. Mathwich, and K. Davies

THE RADIOWAVE PROPAGATION ENVIRONMENT - SCIENCE AND TECHNOLOGY OBJECTIVES FOR THE 80's

JOHN M. GOODMAN
E.O. HULBURT CENTER FOR SPACE RESEARCH
NAVAL RESEARCH LABORATORY
WASHINGTON, D.C. 20375

JULES AARONS
AIR FORCE GEOPHYSICS LABORATORY
L.G. HANSCOM FIELD
BEDFORD, MA. 01730

ABSTRACT

Because the ionosphere is so crucially important to the operation of RF systems employed by the commercial world and DoD, the topics covered in various topical conferences such as IES, and others organized by components of URSI, COSPAR, AGU, IEEE, AND NATO-AGARD have remained active for several decades. Early emphasis in ionospheric propagation research was on HF but present emphasis is away from this band - both above and below, especially in the industrialized nations. Even so, the HF band remains important in many applications and will be a significant if not major driving force for ionospheric research for some time to come. This paper outlines the perceived science and technology objectives for the 1980's based upon the C^3I systems which are reckoned to be in place as we approach the 21st century.

1.0 General Introduction

The ionosphere has a strong influence on the personality of Radiowaves which propagate beneath, within, or through it; and this influence derives from the spatial and temporal non-uniformity in the refractivity of the magneto-ionic medium. A number of texts dealing with ionospheric radio propagation have been published, and general reviews of propagation effects are available in the literature, especially dealing with earth-space propagation. The fundamental importance of the ionosphere to DoD systems as well as others is evidenced by a substantial number of topical conferences sponsored by the wave propagation panel of the Advisory Group for Aerospace Research and Development (AGARD) which is affiliated with NATO. In addition, the U.S. Navy has sponsored two Ionospheric Effects Symposia which have surveyed the field and have illuminated relevance issues. Non-DoD conferences and workshops dealing with all aspects of the geoplasma and radiowave propagation abound, and references in this area are far too voluminous to be cited in a brief introduction. The role of prediction, forecasting, and assessment of the environment for a priori estimation of effects on propagating systems has received considerable attention over the years. As a result a plethora of ionospheric models have been developed for use in system design, and in some cases, system operations. As a minimum, models have found utility in specifying margins over which propagating systems must be designed to adapt to the changing ionosphere. The robustness of systems typically increases with increasing frequency for transionospheric propagation and in this case the ionosphere is simply a nuissance. Nevertheless the diminution of deleterious effects does not occur as rapidly as one would like. Ionospheric scintillation is a case and point.

Before the advent of the space age, the attention in this field was naturally directed toward the radio frequency domain below VHF, principally centered on the HF band for communication and surveillance purposes but also including VLF through MF which is of use in communication and navigation. Recently, ELF has been considered as a prime candidate for strategic communication by virtue of its considerable penetration into sea water. At HF and below, the ionosphere may be regarded, more or less, as part of the propagation channel - a very substantial part in most cases. Indeed, knowledge of the detailed character of the lower ionosphere is fundamental to the intelligent design and efficient operation of many systems which employ this frequency domain. The

precariousness of HF to the ionosphere is well known and the problems are largely understood. There physical understanding, however, does not always translate into useful forecasting or prediction techniques. The problem of sudden ionospheric disturbances has been attacked by system approaches involving the monitoring of 1-8 A solar x-ray flux but the successful prediction of magnetic storm (ionospheric storms) effects is not yet in hand, although most assuredly significant strides have been made. The problem of available bandwidth as well as ionospheric effects made the HF band less attractive than higher frequencies for line of sight (LOS) propagation but in the absence of satellite technology there was little choice for Beyond-Line-of-Sight (BLOS) or Extended-Line-of Sight (ELOS) requirements.

Indeed the use of the HF band is currently growing despite the large expenditures for satellite-dependent communication systems. This has come about for a variety of reasons including the sensitivity of military planners to satellite vulnerability and, above all, system cost. At low latitudes especially, HF point-to-point communication systems are quite attractive since this region exhibits higher maximum usable frequencies (MUFs) which increases the available bandwidth, and it is relatively uneffected by particle precipitation events and magnetic storm effects. The principal transient problem is the absorption arising from solar flare x-rays during daytime and certain transequatorial propagation effects associated with the Appleton anomaly. In addition many developing countries occupy the low latitude region and the expense of satellite systems might preclude their use there. Furthermore SATCOM scintillation is more pronounced near the geomagnetic equator. HF communication is somewhat less attractive at temperate and high latitudes but will nevertheless be a significant component in military thinking. The DoD officially regards the HF band as simply a back-up to SATCOM and because of this a considerable amount of expertise has been lost over the years because scientific and engineering talent has been directed toward the development of satellite systems at the expense of HF. This trend is now changing and conscious efforts are underway to improve our understanding of the HF channel. Effort is directed in two main areas: (i) development of advanced modems and adaptive systems which account for the ionosphere automatically and (ii) development and testing of improved short-term forecasting and assessment technologies. The first approach depends upon some understanding of the average behavior of the channel and may justify sympotic studies of the ionosphere for use in long and intermediate term prediction schemes. The second approach depends upon a somewhat more superficial understanding of the basic physics but may depend heavily upon real-time assessment schemes for success. More will be said about this later on.

Even though SATCOM will likely be the primary system for long haul communications in the future with an evolutionary development toward x-band or EHF, HF has considerable charm in the surveillance arena and will be important for some time in the intelligence world (i.e., the "I" of C^3I). HF will likely be a backup long haul communication system but will maintain a strong potential for intra-task communications via ground wave. The wise and judicious use of this mode will protect against detection by unfriendly forces while at the same time allowing communication to the desired terminals. This requires a knowledge of ionosphere. Conversely a good knowledge of the ionosphere over denied areas may assist in the intelligent allocation of signal intercept resources. Thus the use of elaborate electronic warfare technology by friend and foe alike may undoubtedly involve real-time knowledge of the radiowave propagation environment; and the ionosphere plays a key role in this game.

As satellites came of age, the radio frequencies employed for command and control of space systems as well as those used for long distance communications via satellite became progressively higher to avoid the obviously annoying effects introduced by the ionosphere. Nevertheless small-scale ionospheric inhomogeneities of electron concentration posed a potential threat to SATCOM over the magnetic equator and near the polar regions especially at UHF. Much of the commercial world has advanced to higher frequencies in the GHz domain but even here scintillation of signal phase and amplitude sometimes result in system perturbations, although the effects are not operationally significant. There are still important systems which operate a UHF and the DoD will utilize this band for Fleet Broadcast for some years to come. The primary ionospheric problem associated with the earth-space path is undoubtedly amplitude and phase scintillation. This not only must be accounted for by SATCOM system designers but also by designers of advanced satellite surveillance systems which require minimization of phase fluctuations and by advanced navigation systems such as NAVSTAR/GPS. Studies have shown that scintillation at L-Band, a frequency band employed by GPS, may be deleterious to operation in the important equatorial and low latitude regions. Single-frequency users of GPS also require more adequate models of total electron content to satisfy navigation requirements.

One of the most fascinating new fields of ionospheric physics is embodied in modification technology. This technology, once refined, may ultimately be of use in opening the bandwidth available for HF communications, extending the range of HF-OTH radar, thwarting naturally occurring

scintillation, as an assist in overcoming anti-communication nuclear burst scenarios and may be important in the generation of ELF/VLF signals for communication purposes. Also it has been shown that ionospheric inhomogeneities produced by powerful HF heaters will allow the BLOS transmission of facsimile at VHF and UHF. The theory and experimental aspects of HF heating is contained in proceedings of special symposia. A number of HF heater experiments are now underway or are planned in the near future. Sites for controlled experiments include the EISCAT facility in Northern Europe, a facility near Arecibo Puerto Rico, and a facility in Alaska. Several facilities are also located in the USSR. The proposed Satellite Power System (SPS) is a system which by its very nature may also be a potential contributor to unwanted ionospheric modification through the atmospheric/ionospheric "heating" process. Chemical reagent releases, ion cloud releases, and RF heating have been explored to study various aspects of ionospheric phenomenology. Very large holes in the ionosphere have been produced (inadvertantly) by the launch of rockets. The most noteworthy example is the launch of Skylab and more recently HEAO-C. Planned experiments in which water (H_2O) is released have also been successful in producing various ionospheric effects. The inadvertant modification caused by rocket launches is currently being investigated because of the possibility that ionospherically-supported radio links may suffer impairments from such effects.

2.0 Requirements and Needs

2.1 General Remarks - The perceived DoD requirements for satellite monitoring of the earth's space environment are driven in large measure by the need to command and control friendly forces and to gather intelligence concerning hostile forces. This driving function, termed C^3I for Command, Control, Communication and Intelligence, must be operable in all stages of conflict from crisis to all-out nuclear war. Force multiplication via C^3I is imperative in the future because of decreased reliance on foreign bases and the force advantage enjoyed by the enemy in certain postulated engagements. However, the degree to which such mulitplication may be accomplished is dependent upon adequate communication - the "glue" which holds it all together. The outcome of crisis or warfare engagements is also dependent upon adequate navigation/positioning of U.S. forces as well as ocean surveillance and oceanic weather/sea state information. These and the aforementioned components of C^3I are thought to be best handled through use of space assets to obtain global and synoptic coverage and to enhance connectivity between the actual forces and the force strategists and managers. The efficient use of terrestrial sensors and ionospheric channel "probes" has not been fully explored in this context and requires attention. The direct communication of information between the various components of the C^3I systems (including navigation and surveillance) requires a cooperative environment for the channel. Such "cooperation" is dependent upon the nature of the medium including: the troposphere, ionosphere, and magnetosphere - all of which are geographically complex and time-varying. The total system, therefore, requires a component capable of deriving specifics concerning the total hierarchy of solar-terrestrial interactions. At best, the target system would benefit the force managers through a comprehensive timely and accurate predictive capability vis-a-vis the performance of communication, navigation, surveillance, and weather-gathering systems for both U.S. and hostile forces. This system would provide the maximum response time for command and control. As a minimum the target system would provide a real-time assessment capability providing a minimum response time for command and control of forces. The minimum system must of necessity be more accurate and reliable. It is emphasized that satellites may provide the best capability to assess the aerospace environment on a synoptic and global basis and to provide early warning of changes in the interplanetary and near-earth space environments. Of course, ground-based diagnostic tools such as vertical incidence sounders, magnetometers, total electron content polarimeters, etc. may be orchestrated to provide considerable information especially if tied to models which are capable of update. The primary problem associated with the utilization of any specified hierarchy of environmental observables is one of real or quasi-real time data collection, rapid data dissemination, subsequent fusion of the information, and processing for the ultimate customer in a convenient form. There is also a need for feedback and interrogation for quality control. Thus the environmental assessment target system is itself a C^3I system. The intelligence (information) must be gathered, the information systems must be controlled to limit data which might overload the system, relevant data must be available on command, and the whole system must be tied together with a reliable communication system. Those changes in the geoplasma environment which affect existing communication systems in space such as FLEETSATCOM and DSCS (i.e., scintillation) and ground-based systems such as HF (i.e., absorption and coverage variation) and ground-based navigation systems such as LORAN and OMEGA are also affected by solar-driven ionospheric effects, and the satellite-based TRANSIT and NAVSTAR/GPS systems are not immune to perturbations in performance as a result of the non-uniform ionospheric personality; viz., scintillation producing inhomogeneities. Just as HF communication is held hostage to the vagaries of the ionospheric channel, we find the HF-OTH radar and HF-DF systems are also dependent upon the nature of the lower ionosphere. Future surveillance satellite systems may have performance degradations in certain regions because of propagation disturbances and ionospheric clutter arising

from sub-auroral forms and field aligned ionization. Satellite synthetic aperture systems (SAR) will encounter diminution in aperture gain because of ionospheric inhomogeneities which distort the surfaces of constant wave phase. The sun is responsible for increasing the rate at which satellite orbits decay since satellite drag is related to thermospheric temperature which is enhanced during periods of high solar activity. Furthermore solar monitoring is necessary to precisely specify orbits of navigation or low orbiting surveillance satellites. Spacecraft charging may be particularly damaging to systems such as NAVSTAR/GPS since the constellation operates in a rather virulent segment of the earth's radiation belts. The character of the magnetospheric changes radically during magnetic storms (substorms) and it is possible that system malfunction may result from spacecraft-magnetospheric interactions during disturbed periods. Radio noise generated during solar disturbances will reduce the efficiency of space-borne transponders and space radiation may also damage solar cells and sensors. Clearly the case is strong for developing operational systems for monitoring the solar-terrestrial environment. The nature of all the effects thought to be important sources of malfunction (or performance reduction) in systems of importance to the DoD is not known quantitatively except in a few instances; more research is needed to characterize the ionosphere and magnetosphere under all conditions of interest. Indeed, it is the role of space research to develop diagnostic tools to follow and forecast ionospheric behavior and response, both natural and artificial. This entails the design and use of new instrumentation to probe the neutral, ionized, and magnetically-controlled atmospheric components whether by direct probing, or by remote sensing of atmospheric, ionospheric and magnetospheric radioactive emissions. It also includes interpretive modelling, the testing of models against observed natural events, and improvement of forecasts of probable behavior including that of the sun, of the interplanetary medium and the earth's ionosphere. The second objective of space research is to develop a fuller understanding of the electromagnetic radiation environment which may affect the near-earth space medium and the military users of space and/or the space medium. This includes the delineation of the physical location of radiating sources for all spectral regions; from the radio band, through the infrared and visible, to the ultraviolet, x-ray and gamma ray ranges, together with their spectral and temporal behavior. The space environment, because of the fact that it exhibits control or influence over many systems of critical importance to the DoD in the C^3I, navigation, and surveillance arenas, is of fundamental importance to the military strategist. A-priori knowledge of the performance of radiowave propagation systems as deduced indirectly from solar-environmental monitoring, or more directly by propagation channel assessment may be critical in the development of tactical warfare scenarios. Its importance in the commercial world is also established although less critical.

3.0 Basic Research Objectives in Environmental Monitoring

Together with the operational space assets, there is a need for long term support of environmental monitoring systems, specifically those related to the understanding of ionospheric/magnetospheric phenomenology and space radiation effects. It is well understood that the program-related areas for such basic research includes both communications and navigation but the support for surveillance functions and various exploitation scenarios is not as well advertised.

Basic research carried out by various government laboratories in support of long-term requirements in specified C^3I areas is encouraged, to ensure that relevant breakthroughs in the physical sciences are exploited to provide new or improved techniques. In addition to research in the broad field of physical electronics, equally important research is underway in the areas of electromagnetic wave propagation and solar-magnetospheric-ionospheric research. The effects of the ionosphere on electromagnetic wave propagation are not limited to the lower frequency bands, but extend in the form of scintillations to frequencies used by communication satellites and transionospheric radar. Other effects such as spacecraft charging and satellite drag are also important system constraints for space systems and are being pursued at this basic research level. Some examples of relevant basic research programs in these areas are cited below. Together a generic description:

3.1 Ionospheric Research - There is a need to fully understand how to relate ionospheric influences to the design and operation of communication, navigation, and surveillance systems. This implies support for a complete research program which is aimed toward describing the ionosphere's total behavior beginning with the solar energy input and ending with the ionospheric response which ultimately controls radio propagation. A basic understanding of ionospheric dynamics is crucial to predicting the performance of both operational and conceptual systems. However, it is often not productive to attempt to relate ionospheric research to a particular system or even a frequency band because the same ionospheric phenomena is often responsible for performance degradation of quite different systems. It is considered more appropriate to study the ionospheric processes from a physical viewpoint. On the other hand it is imperative that the basic researcher be cognizant of the problems encountered by specific systems. In this way he may

transition his basic research ideas more efficiently and may provide more useful guidance to the system designer. It is remarked that many basic researchers are involved in exploratory and advanced exploratory research. This may be a pre-requisite for success.

Current programs are aimed at a better understanding of solar-magnetospheric-ionospheric interactions over the entire solar spectrum from x-rays to radio waves; the interplanetary magnetic field; the propagation of particles and waves; ionospheric changes with time; and ionospheric reaction to various natural and man-made perturbations. Once these phenomena are sufficiently understood system designers will be able to take full advantage of the environment to produce more reliable, more efficient systems. For operational systems, timely forecasts and assessments of propagation conditions based on inputs from real-time ionospheric systems and solar observing stations is essential.

3.2 Magnetospheric Research - Programs to study the spatial and temporal distribution of energetic particle fluxes, and their correlation with various measures of solar activity and ionospheric phenomena is underway. Trapped and precipitating particles, and electric fields which may energize ambient charged particles in the magnetosphere are being investigated. The results of these studies are required to determine the ion motions in the magnetosphere and the configuration and intensity of field aligned currents which link the source region in the magnetosphere to the ionosphere. The work is essential to understand and possibly predict ionospheric disturbances at high latitude.

3.3 Interplanetary Field and Solar Wind Research - The sun's atmosphere expands outward into space continuously to envelop the earth and other planets with the region between the sun and the earth characterized by its magnetic field. The sun's atmosphere is a fully ionized gas whose outward expansion produces the solar wind which directly modulates all radiation impinging on the earth. Extensive observations of the solar wind have been made near the earth and although its general properties are well established the plasma parameters fluctuate greatly with time and position. Presumably the gross fluctuations reflect changing conditions on the sun, but an appreciable fraction of them are generated by dynamic processes within the plasma itself. Satellite observations and theoretical studies are continuing to provide the necessary scientific base for reliable forecasts of ionospheric perturbations.

3.4 Solar Radiation and Magnetic Field Research - The driving force for all natural ionospheric activity is the sun, and theoretical and experimental research is being conducted to determine the gross dynamic motions on the solar surface using optical and radio techniques. While the physics of the sun, per se, is not a primary concern of these programs it is intrinsically included in the attempt to trace the effect of the sun's input to the interplanetary medium and the impact of its radiation and plasma flow on the terrestrial atmosphere. The aim of this work is to identify those centers of solar activity that give rise to flares capable of significantly affecting radio transmission on the earth. This requires the study of mechanisms for the production of x-rays, UV and energetic particles, as well as the intense magnetic fields characteristic of flare activity.

3.5 Propagation Studies - Various research establishments are currently involved in direct monitoring of solar-driven perturbations on the communication channel through observation of the amplitude and phase scintillation imposed on satellite signals. Data at VHF, UHF and L-band are obtained and analyzed for determination of first order signal statistics and for refinements in the knowledge of the geomorphology of scintillation producing regions. The ultimate goal is to derive a suitable model of amplitude and phase scintillation for use in prediction of events using readily derivable and accessible parameters as the driving functions. Currently there is inadequate understanding of the scintillation problem from a point-of-view of basic physics although considerable progress has been made in the numerical simulation of the processes thought responsible for scintillation over the magnetic equator. Progress has been made in explaining high latitudes effects as well.

Recently renewed interest in the HF band for communication has stimulated more experimental studies tied to improvement in HF channel assessment. Because the HF band is strongly affected by the ionosphere, remote sensing instruments which use HF are excellent diagnostic tools. Various space experiments are envisioned upon activation of the space shuttle. Experiments involving ELF/VLF/LF transmissions are continuing and emphasis is being directed toward the interaction/generation of this band of frequencies with a modified ionosphere.

4.0 Operational Systems Which Basic Research Will Support

Any system which requires the transmission of an electro-magnetic signal through or under the ionosphere may be affected by changing environmental conditions in the ionosphere. Long range

communications, over the horizon radars, and navigation systems are the most sensitive to changing ionospheric conditions. However, satellite-borne systems in the UHF/SHF frequency range are by no means safe from degradation or disruption as a result of scintillation problems. Even during apparently benign solar activity conditions, ionospheric scintillation may be an important factor.

4.1 Communications Systems - are vital to the success of various DoD missions. They transform a collection of units or moving platforms into an integrated weapon system and permit the commitment of this weapon system as a cohesive and credible element of national policy. Even with many billions of dollars invested in communication systems, the total system is not totally adequate for either the needs of national command policy or the efficient operation of the services. It is unreliable, slow and prone to errors. It is neither totally secure nor covert, and is susceptible to jamming. Future SATCOM systems operating at EHF may partially resolve this problem. However HF will continue to operate as a back up for long haul communications and will be employed for intra-task force communications. Further HF will be important in the surveillance in the world for some time to come. The futility of developing fully self adaptive systems over uncooperative links dictates that ionospheric effects will be a major contributor to communication ineffectiveness.

4.2 Navigation Systems - The OMEGA navigation system operating in the VLF band is essentially propagation limited. The system operates in the hyperbolic mode in which phase differences are measured between signals received from various pairs of transmitters eliminating the need for precise time at the receivers. Normal fix (RMS) error is 1 nautical mile however perturbations caused by ionospheric disturbances can raise the fix error by an order of magnitude which is unacceptable for both peacetime and wartime operations.

The existing Navy TRANSIT Navigation satellite with an unperturbed fix accuracy of 1/2 nautical mile has registered errors of up to five (5) nautical miles during disturbed solar conditions.

The NAVSTAR/GPS constellation was planned to be a group of twenty-four satellites (8 satellites in three orbital planes) the purpose of which is to provide precise time and ranging data to users. The scaled-down version will contain 18 satellites in the constellation. Ionospheric compensation (with respect to group-path-delay error) is removed by utilization of two closely spaced L-band frequencies, but L-band scintillation may still present a problem in the neighborhood of the geomagnetic equator ($\pm$ 15°) causing system dropouts. Further, single frequency (disadvantaged) users will also have to compensate for the ionospheric group-path-delay through utilization of an appropriate model of the total electron content (TEC).

4.3 Surveillance - Over-the-Horizon (OTH) radars employ either a ionospheric forward scatter (receiver lies in front of the transmitter) or backscatter (part of the target reflections bounce back to the combined receiver/transmitter site) mode. Operating in the HF band these systems are vulnerable to ionospheric variations and these variations seriously degrade their usefulness both for fleet defense or for measuring ocean surface conditions in real-time at ranges from 600 to 2200 nautical miles.

HF direction finding equipment is clearly affected by the ionosphere. Travelling ionospheric disturbances produce angle-of-arrival (AOA) errors not only following magnetic storms but also at unexpected times. Day-night ionospheric (tilts) effects are well known.

5.0 Environmental Effects/System Deficiencies by Frequency Band

5.1 ELF Frequency Band (30-300 Hz)

5.1.1 Background - The ELF band is of much current interest, because it is proposed to operate a communications system in this band to provide one way radio communications primarily with submarines cruising at patrol depth anywhere in the world. It is proposed to operate the system on frequencies between 40 and 80 hertz. ELF is relatively unaffected by nuclear bursts.

5.1.2 Deficiency - Energy in this band propagates in an effective parallel-plate wave guide, whose lower boundary is the earth's surface and whose upper boundary is the D and E regions of the ionosphere. This upper boundary is susceptible to large scale ionospheric irregularities and anisotropy due to the geomagnetic field and diurnal effects such as

BI-DIRECTIONAL PROPAGATION (long/short path interference)
- Geomagnetic non-reciprocity
- Day-night asymmetry
- Transequatorial paths (dark path irregularities)

MULTI-LAYER RESONANCE EFFECTS
Anomalous high attenuation (due to standing waves set up above the nightime D and lower E regions)
Dispersion

IONOSPHERIC DISCONTINUITIES
Height changes at twilight zone
Interposed conducting layers such as Sporadic E
Solar and nuclear perturbations

NORTHERN LATITUDE PHENOMENA
Solar x-ray flares
Solar charged particle fluxes
PCA Events (increased attenuation)
Magnetic Storms

5.1 3 Scientific Needs - Specification of D-region ionosphere parameters are needed; electron density, neutral density, and atmospheric composition. Relationship of D and E region parameters to geophysical phenomena. The importance of Bremsstrahlung in low altitude ionization production and the geomagnetic late time effect produced by particle precipitation from the magnetosphere needs to be explored.

5.2 VLF BAND (3-30 KHz)

5.2.1 Background - The VLF band is of major importance of DoD for communications and navigation. The VLF radio system to submarines has good availability and adequate capacity but it does not permit submarines to operate at optimum depths and speeds while communicating. The VLF navigation system, OMEGA, is a world-wide system which can provide a geographic position fix to aircraft, ships, including submerged submarines with a RMS fix error of 1 nautical mile.

5.2.2 Deficiency - Degrading effects on communication are due to changes in signal attenuation, phase, and phase rate brought about by perturbations in the propagation medium. For acceptable communications there must be enough stability in the transmission medium from one information pulse to the next.

In the navigation use deviations from the assumed earth-ionosphere waveguide will cause phase deviations from the predicted values resulting in fix errors up to over an order of magnitude (10 nautical miles) over the unperturbed condition.

VLF disturbances consist of sudden phase anomalies which occur simultaneously with the visible flare and are caused by solar x-rays and result in sudden increase in phase velocity and change in amplitude of the signal due to increased ionization below the normal D region. Polar Cap events occur from a few minutes to hours after some larger flares and are caused by protons (and possibly heavier particles) precipitating into the high latitude ionosphere causing increased ionization below the D region. These are predominately daytime effects and may result in phase and amplitude changes. Magnetic storms are characterized by a sudden commencement 1-2 days after a solar flare (but many do not appear flare related). Effects at mid-latitudes affect the phase and amplitude of the signal.

5.2.3 Scientific Needs - Need to understand the dynamic behavior of the D region in order to predict its behavior. Excellent mathematical tools exist for modelling VLF propagation, the use of which requires specification of D region parameters which are inadequately known at present. Specification of electron density, neutral density and atmospheric composition is needed as well as ionization mechanisms and acceleration mechanisms of particles precipitating from the magnetosphere.

5.3 MF/HF BAND (300 KHz - 30 MHz)

5.3.1 Background

Communications - The use of high-frequency radio links as a principal element of the Navy Communication System will be required until at least 1990 because satellite systems of sufficient capacity will not be available until that period. Even after 1990 some residual capability must be retained as a backup system and for communication with allied navies. Because of satellite vulnerabilities just now being illuminated and because of the cost of space assets, HF may have a significant role in the future DoD communication architecture. The two greatest liabilities of the current HF radio system are the poor availability (about 65% [illegible] ease with which an enemy can

locate U.S. ships by long-range HF direction-finding. Furthermore with the greater utilization of SATCOM the hands-on expertise for making HF communications work properly is gradually eroding.

Surveillance - Over-the-Horizon-Detection radars utilize the processes of ionospheric reflection. System operation depends on either ionospheric forward scatter (the receiver lies in front of the transmitter) or the backscatter process (part of the target reflections bound back to a combined receiver/transmitter site). Backscatter OHD systems are more effective. Because the returning target signal is submerged in the scatter return of the surrounding target area, high transmitted power, pulse coding, and advanced data correlation are required. Since most targets of interest have higher velocities than their surrounding environments, doppler shifts on the order of a few Hz are of prime interest. The key to successful employment of this system concept is real-time knowledge of ionospheric characteristics over the search path, which can involve any number of great-circle paths, thousands of km long, some of which traverse the polar regions and the auroral zones.

5.3.2 Deficiency - The HF band is by far the most sensitive to ionospheric variations due to solar disturbances since it requires all layers of the ionosphere for propagation. Primary reflection occurs in the E and F regions with the D region acting as a variable attenuator.

There are several quasi-global disturbance phenomena that can degrade long-range HF systems. Three types are:

Very short term	Sudden frequency deviation (SFD)
Short term	Short-wave fade (SWF) or D-region enhancement
Long term	Ionospheric storm and PCA

Sudden Frequency Deviation (SFD) - is an impulsive deviation in the frequency of radio waves reflected from the F-region of the ionosphere. This phenomenon is caused by ionization enhancements in the upper D, E, and F. It is characterized by a sharp peak deviation and a decay back to zero deviation. The duration is approximately 15 minutes and is coincident with event onset. This frequency deviation could produce false targeting on OTH radars depending on system gating and correlation settings.

Short Wave Fade (SWF) - A short-wave fade is a decrease or loss in signal strength of radio waves, caused by increased absorption to x-ray enhancement in the D-region and the lower E-region. Onset is in close coincidence with the solar eruption and the effects can last from several minutes to several hours, depending on the magnitude of the x-ray burst and the HF frequency of operation.

Ionospheric Storms - The ionospheric storm is a long-term phenomenon resulting from a disruption of the earth's magnetic field by particles (electrons and protons) ejected from a solar flare. Onset can be anywhere from 8 to 48 hours after a solar event and primarily affects the upper F-region. It is characterized by degraded HF radio propagation, high latitude radio blackout, and sporadic E. The effects can last anywhere from 8 to 72 hours. At high latitudes, ionospheric storms create such phenomena as auroral displays and PCA events.

Polar Cap Absorption (PCA) - The PCA is a result of high-energy particle influx from a solar event and is evidenced by enhanced absorption of radio waves in the polar regions. The PCA starts from 1-4 hours after the solar event and can last from 1-3 days.

5.3.3 Scientific Need - There exists today gaps in the current state of the knowledge in defining certain ionospheric constituents, ionization, recombination and transport processes. Although theories abound, few have been rigorously tested experimentally. Specifically some of the areas where a definite lack of knowledge exists include:

(a) ionospheric winds and large scale motions,
(b) auroral to midlatitude and equatorial coupling,
(c) E and F region interactions,
(d) D-region constituency and recombination rates,
(e) F-region dynamics during ionospheric and geomagnetic storm disturbances,
(f) the formation of F-region irregularities,
(g) E and F-region coupling,
(h) ionospheric and magnetospheric coupling,
(i) solar flare mechanisms and prediction,
(j) ionospheric irregularity characteristics.

Various propagation prediction models are now being improved to allow for update. These models

provide for quasi-real-time prediction/forecasting of the optimum frequencies for transmission over cooperative links. However, ionospheric turbulence is inadequately described by these models and approaches. The presence of Spread-F and large scale irregularities are difficult to account for and introduce uncertainties. There is a need to improve our understanding of the ionospheric properties which are important in wideband HF as well as narrow band systems.

5.4 UHF/SHF BAND (300 MHz - 30 GHz)

5.4.1 Background - Beyond-Line-of-sight radio communications among surface forces and aircraft can be achieved by means of UHF/SHF communications satellites and/or high-altitude UHF communication relays.

5.4.2 Deficiency

Scintillation - The occurrence of irregularities in the ionization distribution in the ionosphere has been known since the advent of ionospheric radio sounding techniques in the 1920's. The early observations were in the form of a spreading and smearing of the F-region reflections recorded by an ionosonde, which led to terminology, "spread F", now classifying this phenomenon. More recent observations from satellite-borne experiments, and satellite beacon monitoring have led to the conclusion that irregularities may occur anywhere within the ionosphere but are most prevalent at F-region altitudes in the equatorial and polar-auroral zones.

The irregularities that form in these spread-F zones are typically severe enough to have a detrimental effect on any system depending upon electromagnetic-wave propagation off or through the disturbed regions. The growing use of satellite-borne instrumentation for navigation and communication networks necessitates a better understanding of the gross dynamics of the irregularity zones and the intrinsic characteristics of the irregularities themselves. Ionospherically-caused scintillations imposed on satellite transmissions will need to be considered in the design of future satellite systems especially if operating through the equatorial regions ($\pm$ 20^{o}) and in the auroral or polar regions.

5.4.3 Scientific Needs - Assess the deficiencies in knowledge about ionospheric caused scintillations, how they are formed and their relationship to ionospheric conditions. Many theories have been suggested for the formation of the irregularities and none completely acceptable. Promising theories for equatorial scintillation have been advanced and there is an intriguing relationship between plumes (holes) observed by radar, spread F irregularities and the position of the equatorial anomaly. Mid-latitude effects have not been adequately identified yet but, collection and analysis of data available indicates that there may be occasional problems in that region also. Fading caused by equatorial irregularities effects frequencies as high as 6 GHz or more with fading in excess of 4-6 dB and it is significantly greater at lower frequencies. There was almost no information available regarding phase coherence during scintillation until WIDEBAND DNA-002 was launched. Now the situation has improved considerably. Nevertheless it has been discovered that pronounced phase scintillation does occur in the absence of amplitude fading. Variation in propagation time delay due to scintillation sources may be small, but variations in total electron content at different look angles and at different times probably needs to be considered in the design of navigation satellite systems. Additional propagation studies of scintillation effects, particularly in equatorial and polar regions, appear to be needed. This will require a coordinated program, with observations and analysis. Theoretical development of strong scatter, as opposed to weak scatter for which there is theory, is also needed to complement the extrapolation of experimental data for system design consideration.

As indicated above, the need to examine scintillation geomorphology or to strive for a coordinated program of assessing competing theories of cause and effect is based upon continued operations at frequencies at L-Band and below principally. Arguments that effects at UHF (i.e., FLTSATCOM) can be controlled by time and space diversity technology are only partially true. To develop these techniques, and to test their validity under realistic conditions comprehensive ionospheric climatologies would have to be produced as a minimum. Unless the worldwide climatology is developed system designs must allow for worst-case Rayleigh fading conditions. It is noteworthy that oceanic scintillation data is virtually non-existent.

6.0 Conclusion

There is already a need to study the ionosphere and its coupling to the magnetosphere above and the troposphere below in order to more fully develop the insight required to specify the radiowave propagation effects introduced by these various media. This venture must be closely coupled to studies of properties associated with both the sun and the interplanetary medium in order that the

potential for an adequate predictive and forecasting technology may be realized in full. Current empirical models of the ionosphere are inadequate except perhaps for system design guidance and more effort could fruitfully be spent in this area. They are especially poor over oceanic areas and in zones where experimental observations are sparse. Physical models serve to fill this void in some instances but they also provide predictions which are not useful in most operational scenarios. In addition, the more sophisticated scientific models - whether they be empirical or physical in nature - may overburden the computational capacity of the operational C^3I system. As a result the future trend may be directed toward the development of more simplified operational models which may be updated with real-time observables. It is emphasized however that user agencies must recognize that the scientific models must precede the operational models to fully identify by relationships involved.

Most of the regular macroscopic features of the ionosphere are reasonably well understood although certain details remain as perplexing problems to the user community. However, with respect to short-term prediction or forecasting capability we encounter far more serious deficiencies and the day-to-day and other short-term temporal variations in electron density contribute to this. There are also some uncertainties associated with the geomorphology of ionospheric inhomogeneities of all scales although considerable progress has been made during the previous decade through comprehensive experimental and theoretical studies. In short we may assert that the irregular properties of the ionosphere are clearly inadequate for purposes of C^3I both from the point of view of phenomenology and driving or triggering functions. This deficiency affects both trans-ionospheric and ionospheric-reflected propagation assessment in profound ways. The definition of the irregular ionosphere and our ability to predict its impact upon radiowave systems in near-real time could be the single most important contribution in the decade of the 80's.

Programs directed toward modification of the radiowave propagation environment will also receive basic research attention in the 80's. The implications of this research is not insubstantial and will be followed with great interest.

IONOSPHERIC EFFECTS OF ROCKET EXHAUST PRODUCTS - SKYLAB AND HEAO-C

J. Zinn, C. D. Sutherland, L. M. Duncan, and S. N. Stone
University of California, Los Alamos National Laboratory
Los Alamos, NM 87545

ABSTRACT

This paper is about ionospheric F-layer depletions produced by chemical reactions with exhaust gases from large rockets. It describes a 2-dimensional computer model of the ionosphere, and it compares model results with experimental data on the structure and variability of the natural ionosphere, as well as data on ionospheric holes produced by the launches of Skylab (May, 1973) and HEAO-C (September, 1979). It also describes measurements that we made in conjunction with the HEAO-C launch, in cooperation with several other experimental groups.

The computer model includes an approximate representation of thermospheric tidal winds and E fields in addition to vertical motions associated with diurnal changes in temperature. The computed ionospheric structure is sensitive to all the above. For a small number of cases, we compare results of computations of the normal diurnal variations of ionospheric structure with incoherent scatter and total electron content data.

Computations of ionospheric depletions from the Skylab and HEAO-C launches are in satisfactory agreement with the observations. The winds appear to be essential for interpretation of the Skylab results.

INTRODUCTION

The main purpose of this paper is to describe a new computer model of the mesosphere, thermosphere and ionsophere, and to compare some computed results with experimental data. We developed the model with certain experimental data in view, and adjustable parameters were adjusted to fit. Then, to the extent that the computations and the data agree, our description of the physics and chemistry in the model may be equivalent to a description of the controlling physics and chemistry of the phenomena observed. If this is true, it is good. Moreover, the values that we establish for the adjustable parameters may be equivalent to a determination of those parameters, which are not otherwise known.

The model was developed in the course of a study of ionospheric depletion effects of large rocket launches through the F layer. For that study it was necessary to treat the chemistry and diffusion of neutral species, as well as ions, as integral parts of the interactive problem. Moreover, a time-dependent two-dimensional or three-dimensional model was required. Ours is two-dimensional with a one-dimensional option. This paper describes a set of one-dimensional computations of the prelaunch ambient ionosphere and of the ionospheric effects of geomagnetic storms. Also included are two-dimensional computations of F-region hole formation from the launches of Skylab I and HEAO-C.

The first observation of a large scale ionospheric hole produced by a rocket launch was reported by Mendillo et al. (1975), and occurred with the launch of Skylab I in May 1973. We performed and reported a number of 2-D model computations of this event (Zinn and Sutherland, 1980a, 1980b, Zinn et al., 1980c); however, the main intent of this paper is to describe subsequent more detailed computations of the normal ionosphere, which have led to a better description of the state of affairs prior to the Skylab launch. Two-dimensional computations of the Skylab and HEAO-C launch effects will also be described.

The data reported by Mendillo et al. consisted of total electron content (TEC) measurements, primarily from the Sagamore Hill Observatory at Hamilton, MA, on a line of sight to the communications satellite ATS-3. Data from other stations and other lines of sight were also reported. It was noted by Mendillo et al. that a major magnetic storm commenced on May 13, the day before the Skylab launch, and the ionosphere was in the recovery phase on May 14.

In our first attempts to model the F-layer depletion produced by the launch we discovered that the observations seemed to require the presence of a southward-blowing wind that carried the exhaust gas cloud out of the several observational lines of sight of the TEC measurements; otherwise the duration of the depletion would have been much longer than observed. Having said that, however, we felt some compulsion to account in detail for the probable wind components involved, along with the E fields and other transport processes.

In this paper we describe the model and our efforts to treat the special phenomena associated with the May 13-14 geomagnetic storm. Section II gives the general outlines of the model. Section III contains comparisons of computed results with data and a discussion of things that the comparisons may have taught us about thermospheric convective mixing, magnetic-storm-driven winds, and electron temperatures. Section IV describes the two-dimensional computations of Skylab and HEAO-C.

THE COMPUTER MODEL

The computer code uses a two-dimensional array of Eulerian mesh cells in Cartesian coordinates, x horizontal (in the geomagnetic meridian plane), and z vertical. The range of z normally extends from 50 to 1050 km, using 15 logarithmically spaced mesh layers. The range of x is specified for each problem, usually 500 to 2000 km on either side of the rocket trajectory plane. There are 21 discrete mesh columns in x, for a total of 21 x 15 = 315 cells.

Conditions are assumed to be uniform in the third Cartesian direction (y). For this reason, the code is best suited to problems where the launch plane is perpendicular to the magnetic meridian (x,z) plane.

The model includes the following physical and chemical processes.

A. Chemistry

The code integrates the chemical/photochemical kinetic equations for 30 individual chemical species in each of the 315 cells. The chemical and photochemical rate coefficients in each cell are periodically recomputed as the temperatures and solar zenith angles change.

B. Solar Radiation, Scatter UV, Cosmic Rays, and Precipitating Electrons

The nominal spectral intensity of sunlight, for quiet solar conditions, in each of 81 wavelength bins, is stored within the code. Attenuation of the solar radiation in each wavelength bin at each of the 15 altitudes, due to absorption by O, O_2, N_2, CO_2, and O_3, is computed from the computed concentrations of those species and the computed solar zenith angles. The attenuation algorithm allows for the sphericity of the earth and the altitude-dependent times of sunrise and sunset. The rate coefficients for photodissociation, photoionization, and photoexcitation reactions are computed from the computed spectral intensities folded with the tabulated wavelength-dependent cross sections.

Other ionization source terms are included in the model to simulate the effects of cosmic rays, precipitating Van Allen electrons, and scattered He 304-A, Lyman α, and Lyman β radiation in the night sector.

The basic input table of solar spectral intensities is a composite of data from Ackerman (1971), and Heroux and Hinteregger (1978), and represents a condition of low solar activity. Recent EUV data by Torr et al. (1979) indicate a significant increase in intensity after the onset of solar cycle 21 for wavelengths shorter than 1050 A. The computed photoionization frequencies show a corresponding increase, almost in direct proportion to the measured 10.7-cm microwave flux. On the basis of these data, we have elected to scale all photoionization coefficients computed with the standard spectrum in proportion to the 81-day average 10.7-cm flux, which we specify as an input constant. The same scale factor is applied to N_2 photodissociation rates.

C. Diffusion and Transport

Diffusion rates are computed for each neutral species under the influence of gravity, for assumed conditions of steady flow, including advection by the neutral winds. Molecular diffusion and eddy diffusion are included as physically distinct processes. The ions and electrons are transported with a separate algorithm that represents the effects of electrostatic and v x B forces in addition to gravity and collisions with the neutral atmosphere. The electrostatic field components are computed self-consistently.

Collision frequencies and molecular diffusion coefficients are computed for each species as functions of altitude and time, using formulae given by Banks and Kockarts (1973) (Vol. A, Chapt. 9).

D. Temperatures

The array of neutral gas temperatures, functions of altitude and time, is generated with the formulae prescribed by Hedin et al. (1977). The latitude, time of year, solar 10.7-cm flux, its

81-day average, and the geomagnetic a_p index are specified as input parameters, in addition to the altitude and time of day. The chemical rate coefficients, diffusion coefficients, collision frequencies and vertical wind velocities are computed from these temperatures.

The model atmosphere expands and contracts in response to the changing temperatures. The computed F-layer ionization profile is very sensitive to this process, since it affects the height distribution of N_2 and O_2 molecules, which affects the rate of destruction of O^+ ions.

Electron and ion temperatures are generated with a set of algorithms that relates them to the neutral temperature, neutral molecular concentrations, and electron concentrations.

A major part of our task has been in the development of the self-consistent plasma transport algorithms. It is beyond the page limitations of this paper to include the equations. They will be described elsewhere.

Another major task was to arrive at a set of eddy diffusion coefficients consistent with known thermospheric composition data. Ultimately, we arrived at a very simple formulation with a constant eddy diffusion coefficient of 3×10^6 cm^2/s at all altitudes; however, special modifications are required for periods of strong geomagnetic activity.

E. E Fields and Horizontal Winds

Seasonal and tidal wind velocities are generated as functions of altitude, time of day, latitude, season, and solar activity, based on published models calibrated against experimental observations (Volland and Mayr, 1973; Forbes and Garrett, 1976, 1978; Garrett and Forbes, 1978; Roble et al., 1977; Harper, 1979; Amayenc, 1974). The neutral winds affect not only the drift and dispersal of contaminants in the ionosphere, such as rocket exhausts, but, through v x B forces, they exert an important influence on the vertical distribution of ionization and its normal temporal variations.

The plasma also drifts in response to E fields. The E field data are derived from a computer code supplied by A. D. Richmond et al. (1980), based on incoherent scatter measurements of ionospheric drift motions.

F. Geomagnetic Storms

Special parameter modifications are required for periods of large geomagnetic activity. Thermospheric temperatures are increased as described by the model of Hedin et al. (1977), which is included in our model as a subroutine. Enhanced equatorward and westward winds are developed (Miller et al., 1979; Mayr and Hedin, 1977; Blanc and Richmond, 1980). We represent these incremental winds as a quantity proportional to the a_p index, depending also on altitude and geomagnetic latitude. The mixing of heavy and light atoms and molecules is enhanced (Jacchia et al., 1976; Mayr and Hedin, 1977; Blanc and Richmond, 1980). We represent this effect by enhancing the thermospheric eddy diffusion coefficient by an amount calculated from the Jacchia et al. (1976) data.

ONE-DIMENSIONAL COMPUTATIONS AND DATA COMPARISONS

1. Midlatitude Ionosphere, May 12-14, 1973

Skylab I was launched on May 14, 1973. Total electron content data for the month of May 1973, recorded at the Sagamore Hill Observatory, Hamilton, MA, were published by Mendillo et al. (1975), showing the normal day-to-day ionospheric variability, the effect of the May 13-14 magnetic storm, and the effect of the Skylab launch. Ionosonde data were also reported for May 14.

Because of our interest in the Skylab launch effects, we made a concentrated effort to model the ambient ionosphere for the location and time of the Skylab F layer traversal (May 12-14, 1973; latitude 38°N, longitude 70.5°W). We computed the expected regular variations of the ionosphere and thermosphere for the "normal" 48-hour period between 1200 EST on May 11 and 1200 EST May 13, followed by a 24-hour geomagnetic storm period commencing at 1200 hrs May 13. The solar activity indices at this time were relatively low ($F_{10.7}$ = 87.4 and its three-solar-rotation-average $\bar{F}_{10.7}$ = 101.8). The magnetic activity indices prior to the storm were moderate; $\bar{A}_p$ = 14 for the preceding 24 hours. With these parameters, the MSIS temperature model produces local exospheric temperatures T_∞ which vary between 832°K at midnight and 962°K at noon.

A set of computed concentration profiles for neutral constituents at noon on May 13 is shown in Fig. 1. The O, N_2, O_2, and Ar profiles are in excellent agreement with those generated by the MSIS model (Hedin et al., 1977a). The H concentrations are lower than the MSIS values by a factor of 2; the He concentrations are high by a factor of 2. Computed ion concentration profiles for the same time are shown in Fig. 2.

Figure 3 shows a set of computed total electron content vs time plots for the 44 hour period 0000 hrs EST May 13 through 2000 hrs May 14, not including the effect of the Skylab launch. The TEC data for the same period, recorded at the Sagamore Hill Observatory in Hamilton, MA, for the ray path to

satellite ATS-3, are also shown. Values of the geomagnetic a_p index for the same period are plotted at the bottom of the graph.

The computed TEC vs time values agree moderately well with the measured values prior to the time of the Skylab launch. The data show a considerable variability with time, which correlate to some degree with the variations of the a_p index. The computed TEC curves show similar trends, suggesting that we have chosen a reasonable model for representing the effects of magnetic activity on the ionospheric composition.

In this model a large number of parameters can be regarded as given, based on published data. A smaller set of parameters are variable - i.e. not available directly as published data, but subject to our adjustment to produce the best possible agreement between our computed results and an aggregate of ionospheric/thermospheric composition data. Quantities that are given include the chemical/photochemical rate coefficients, collision cross sections, the solar spectrum, the neutral temperature profiles, geomagnetic-quiet-time wind profiles and E fields, geomagnetic activity indices, and solar 10.7 cm flux. The mesospheric eddy diffusion coefficient was also regarded as given, equal to 3×10^6 cm^2/s.

The assumed flat profile of the thermospheric eddy diffusion coefficient was arbitrary. The magnetic-quiet-time value was defined arbitrarily to be 3×10^6 cm/s, the same as the mesospheric value. This value is not inconsistent with MSIS neutral composition data. The degree of variability of the thermospheric eddy diffusion coefficient with magnetic activity is expressed through a fit to published Esro satellite neutral composition data (from Jacchia et al. 1976).

The geomagnetic activity influence on meridional wind profiles was also regarded as adjustable, through a constant of proportionality connecting the wind increment to the geomagnetic a_p index. This proportionality constant was adjusted to fit the magnetic storm TEC data.

No published models of electron temperatures are available. We generated our own T_e algorithm, based on data by J. V. Evans, 1973 and R. M. Harper, 1980, coupled with observations of night-time and twilight ionospheric electron content. No special variability of T_e with magnetic activity was assumed, other than that which enters through the neutral temperature variation, described through the MSIS model.

It is well known that ionospheric electron concentrations are influenced strongly by the ratios of concentrations of atomic oxygen to N_2 and O_2 in the regions where electron concentrations are large. In our model these concentrations ratios, at given thermospheric altitudes, are controlled by the neutral temperatures and also by the relative magnitudes of eddy and molecular diffusion coefficients. The altitude of the F2 peak and the degree of physical overlap between the ion/electron populations and the molecule populations are also influenced by the meridional winds, the east-west electric field, and the electron temperature. Hence, the electron concentrations are sensitive indicators of several processes about which there is a dearth of quantitative information. If correctly interpreted, measured electron density profiles can furnish useful data about those processes.

To date we have examined only a limited set of data, and obtained useful but limited information. Obviously, a broader range of data must be examined. At this stage we are optimistic that the computer model is performing well, and that it will be a useful tool for the study of ionospheric processes.

ROCKET LAUNCHES AND TWO-DIMENSIONAL MODEL COMPARISONS

As first enunciated by Mendillo et al. (1975) the plasma depletion effects caused by powered flights of rockets through the F2 region arise as the result of charge exchange reactions of H_2O and H_2, the main rocket exhaust species, with O^+, the normally dominant ion species. These reactions produce primarily H_2O^+, H_3O^+ and OH^+ ions, which recombine rapidly with electrons, reducing the overall electron/ion concentrations. The recombination rate coefficients for electrons with polyatomic ions are some five orders of magnitude larger than for electrons with O^+.

The data of Mendillo et al. (1975) showed a 50% reduction in TEC over a region some 1000 km across and persisting for at least four hours. In the burning of the Saturn V second stage engines some 10^{31} exhaust molecules were emitted. This would have been sufficient in principle to remove all the F layer ions over a region some 10^4 km across. The mitigating factor is gravitational settling, which removes many of the molecules from the F layer before they can react. The number of ions removed is determined by the competitive interaction between lateral diffusion, which brings the unreacted exhaust molecules into contact with more surviving O^+ ions, and gravity, which removes the molecules.

Our initial model computations of the Skylab effects did include these processes, and did lead to an ionospheric hole about 1000 km across. However, in the computations the hole survived past sundown and through the night, i.e., much longer than the four hours implied by the data. The difference was attributable to thermospheric winds, which carried the exhaust gases out of the instrumented lines of sight. Zinn et al. (1980a,b,c) showed results of computations with winds included, which agreed reasonably well with the Mendillo et al. data. We are currently in the process of running computations with a more detailed wind model that includes the effects of the May 13, 1973 magnetic storm.

Similar F2-region depletion phenomena occurred with the Atlas-Centaur launch of satellite HEAO-C in September 1979. By that time ionospheric researchers were aware of the phenomena to observe, and somewhat organized through the efforts of Mendillo et al. (1979). As a result, a large number of observations were conducted. Some were reported by Mendillo, Rote and Bernhardt (1980). Our own efforts included incoherent scatter measurements from Arecibo, PR, airglow measurements from the Florida peninsula, and computer model predictions, furnished in advance to participating observers.

In the case of HEAO-C the launch occurred at night, and a very pronounced ionospheric hole was formed. The electron column density was reduced by a factor of from 5 to 10 for about 5 hours, recovering at sunrise. The size of the hole was about 600 km in the cross-trajectory direction. The 6300 A airglow reached an intensity of 20 kilorayleighs and could be photographed with an unfiltered camera. The observations were in rather good agreement with computer model predictions (Zinn et al., 1980c).

REFERENCES

M. Ackerman, "Ultraviolet Radiation Related to Mesospheric Processes," in Mesospheric Models and Related Experiments, Fiocco (ed.), (D. Reidel Publishing Co., Dordrecht, Holland, 1971).

P. Amayenc, "Tidal Oscillations of the Meridional Neutral Wind at Midlatitudes," Radio Sci. 9, 281-293 (1974).

P. M. Banks, and G. Kockarts, Aeronomy (Academic Press, NY, 1973).

M. Blanc and A. D. Richmond, "The Ionospheric Disturbance Dynamo," J. Geophys. Res. 85, 1669-1686 (1980).

J. V. Evans, "Seasonal and Sunspot Cycle Variations of F Region Electron Temperatures and Protonospheric Heat Fluxes," J. Geophys. Res., 78, 2344-2349 (1973).

J. M. Forbes and H. B. Garrett, "Solar Diurnal Tide in the Thermosphere," J. Atmos. Sci. 33, 2226 (1976).

J. M. Forbes and H. B. Garrett, "Seasonal-Latitudinal Structure of the Diurnal Thermospheric Tide," J. Atmos. Sci. 35, 148-159 (1978).

H. B. Garrett and J. M. Forbes, "Tidal Structure of the Thermosphere at Equinox," J. Atmos. and Terr. Phys. 40, 657-668 (1978).

R. M. Harper, "A Semidiurnal Tide in the Meridional Wind at F-Region Heights at Low Latitudes," J. Geophys. Res. 84, 411-415 (1979).

R. M. Harper, unpublished work (1979b).

A. E. Hedin, J. E. Salah, J. V. Evans, C. A. Reber, G. P. Newton, N. W. Spencer, D. C. Kayser, D. Alcayde, P. Bauer, L. Cogger, and J. P. McClure, "A Global Thermospheric Model Based on Mass Spectrometer and Incoherent Scatter Data MSIS1. N_2 Density and Temperature," J. Geophys. Res. 82, 2139-2147 (1977).

A. E. Hedin, C. A. Reber, G. P. Newton, N. W. Spencer, H. C. Brinton, H. G. Mayr, and W. E. Potter, "A Global Thermospheric Model Based on Mass Spectrometer and Incoherent Scatter Data MSIS2 Composition," J. Geophys. Res. 82, 2148-2156 (1977a).

L. Heroux and H. E. Hinteregger, "Aeronomical Reference Spectrum for Solar UV Below 2000 A," J. Geophys. Res. 83, 5305-5308 (1978).

L. G. Jacchia, J. W. Slowey, and U. von Zahn, "Latitudinal Changes of Composition in the Disturbed Thermosphere from Esro 4 Measurements," J. Geophys. Res. 81, 36-42 (1976).

H. G. Mayr and A. E. Hedin, "Significance of Large Scale Circulation in Magnetic Storm Characteristics With Application to AE-C Neutral Composition Data," J. Geophys. Res. 82, 1227-1234 (1977).

M. Mendillo, G. S. Hawkins, and J. A. Klobuchar, "A Sudden Vanishing of the Ionospheric F Region Due to the Launch of Skylab," J. Geophys. Res. 80, 2217-2228 (June 1975).

M. Mendillo, J. Baumgardner, and J. A. Klobuchar, "Opportunity to Observe a Large-Scale Hole in the Ionosphere," EOS 60, 513 (July 1979).

Michael Mendillo, Donald Rote, and Paul A. Bernhardt, "Preliminary Report on the HEAO Hole in the Ionosphere," EOS 61, 529-530 (1980).

N. A. Miller, J. M. Grebowsky, H. G. Mayr, and I. Harris, "F Layer Positive Response to a Geomagnetic Storm - June 1972," J. Geophys. Res. 84, 6493-6500 (1979).

A. D. Richmond, M. Blanc, B. A. Emery, R. H. Wand, B. G. Fejer, R. F. Woodman, S. Ganguly, P. Amayenc, R. A. Behnke, C. Calderon, and J. V. Evans, "An Empirical Model of Quiet-Day Ionospheric Electric Fields at Middle and Low Latitudes," J. Geophys. Res. 85 (1980).

R. G. Roble, R. E. Dickinson, and E. C. Ridley, "Seasonal and Solar Cycle Variations of the Zonal Mean Circulation in the Thermosphere," J. Geophys. Res. 82, 5493-5504 (1977).

Marsha R. Torr, D. G. Torr, r. A. Ong, and H. E. Hinteregger, "Ionization Frequencies for Major Thermospheric Constituents as a Function of Solar Cycle 21," Geophys. Res. Lett. 6, 771-774 (October 1979).

H. Volland and H. G. Mayr, "A Numerical Study of Three-Dimensional Diurnal Variations within the Thermosphere," Ann. Geophys. 29, 61-75 (1973).

J. Zinn and C. D. Sutherland, "Effects of Rocket Exhaust Products in the Thermosphere and Ionosphere," Los Alamos Scientific Laboratory report LA-8233-MS (1980a).

J. Zinn and C. D. Sutherland, "Effects of Rocket Exhaust Products in the Thermosphere and Ionosphere," Space Solar Power Review 1, 109-132 (1980b).

J. Zinn, C. D. Sutherland, S. N. Stone, L. M. Duncan, and R. Behnke, "Ionospheric Effects of Rocket Exhaust Products - HEAO-C, Skylab, and HLLV," U.S. Department of Energy report, DOE/ER-0081 (1980c).

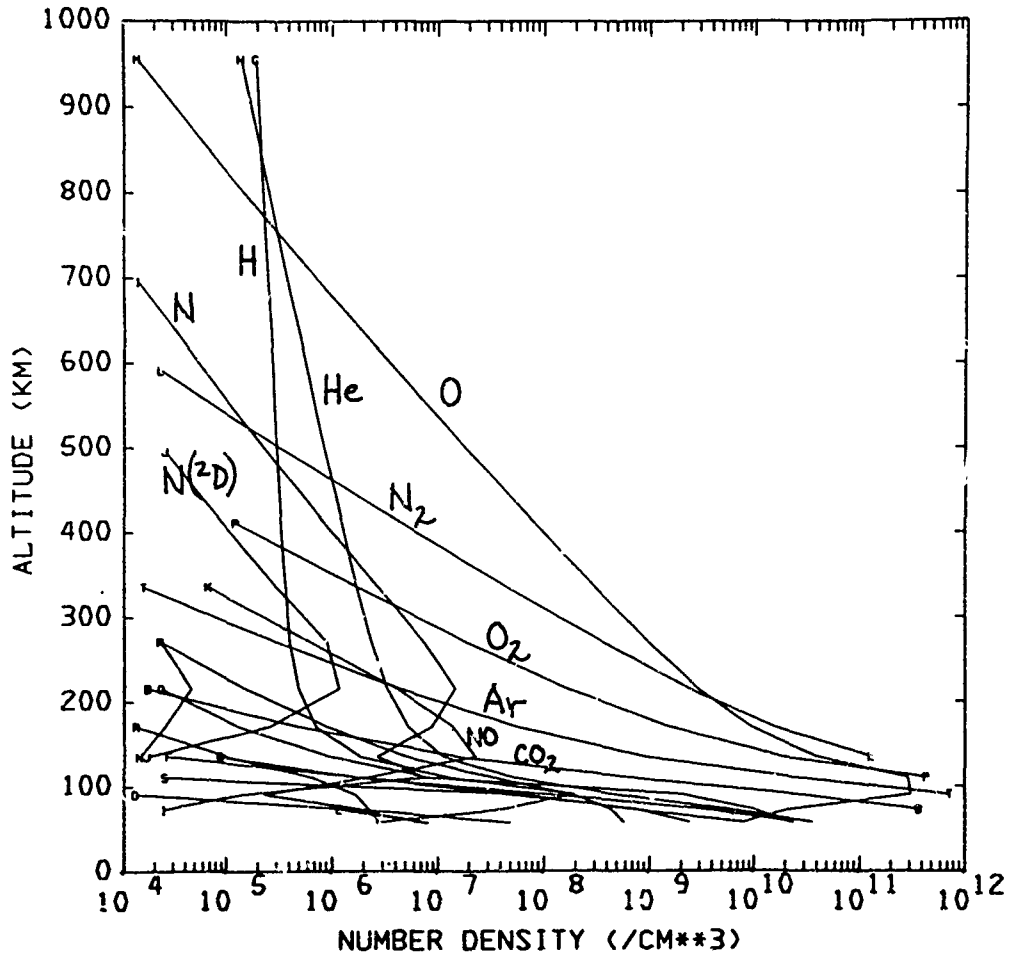


Fig. 1. Computed concentration profiles for neutral species at noon EST May 13, 1973 at Sagamore Hill, MA.

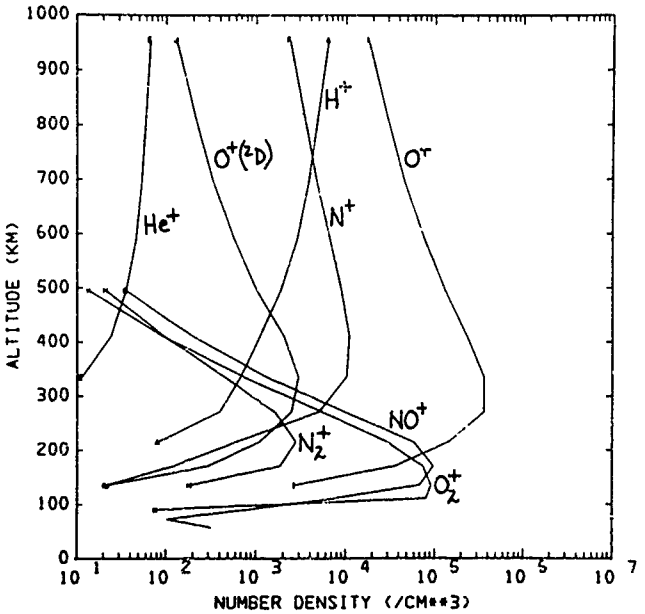


Fig. 2 Corresponding positive ion concentration profiles

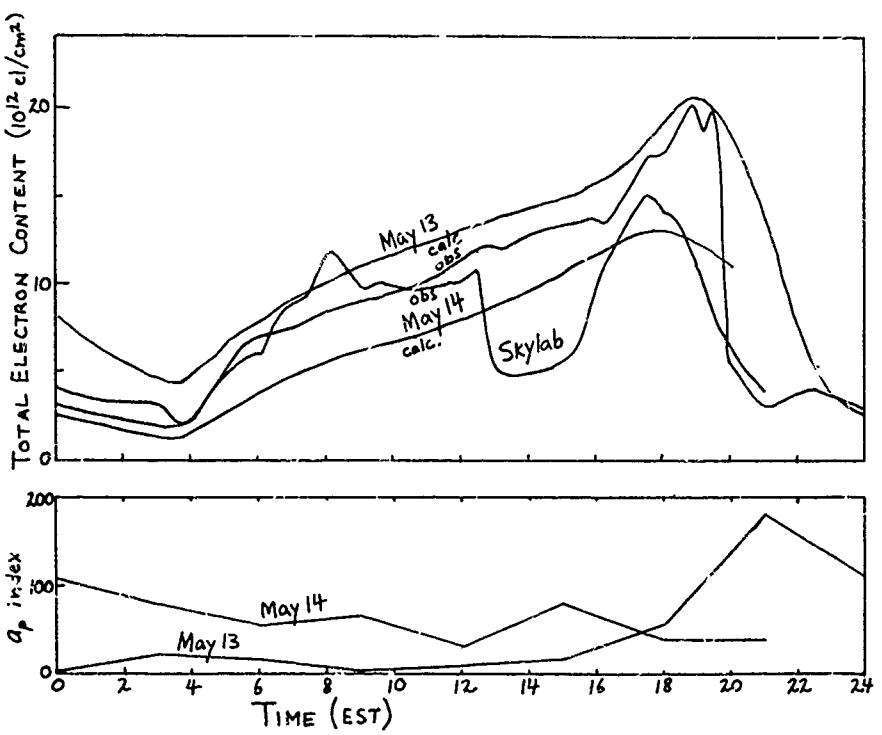


Fig. 3.
Top graph: Computed and measured total electron content vs time for May 13 and 14 from Sagamore Hill, MA. The measured curve for May 14 shows the effect of the Skylab launch. The computed curve does not include the launch.
Bottom graph: Planetary a_p index for the same period.

ROCKET EXHAUST EXPANSION AT HIGH ALTITUDES AND IONOSPHERIC HOLES

MICHAEL H. REILLY
E.O. HULBURT CENTER FOR SPACE RESEARCH
NAVAL RESEARCH LABORATORY
WASHINGTON, DC 20375

ABSTRACT

High altitude releases of certain molecules in booster rocket exhausts are known to deplete electron concentration in the ionosphere dramatically. These effects are exhibited in recent Faraday rotation measurements for the HEAO-C launch from Cape Kennedy on September 20, 1979 and the NOAA-B launch from Vandenberg on May 29, 1980, in which total electron content dropout responses to the booster rockets are evident. The early time development of the ionospheric holes in these data is a proving ground for mathematical models of rocket exhaust expansion in the ionosphere. A kinetic model for this expansion, appropriate for releases at high altitudes, is reformulated in terms of coupled integral equations of the second kind. These equations are obtained from the Boltzmann equation with a collisional relaxation model which allows for unrestricted temperature and velocity differences and satisfies the appropriate conservation conditions for particle number, momentum, and energy. The results are discussed and compared with previous work. Computational strategy is indicated.

INTRODUCTION

Extraordinary ionospheric effects are induced by the introduction of rocket exhaust molecules at high altitudes during the launch of space systems. The well-known SKYLAB effect, as investigated by Mendillo, Hawkins, and Klobuchar [1975], is a case in point. In this study Faraday rotation data for times including the SKYLAB launch were used to characterize the induced electron depletion region (i.e., ionospheric "hole") and to help identify the essential chemical mechanisms responsible for it. Booster rocket emission of such molecules as H_2O and H_2 in the upper atmosphere, say above 250 km where monatomic species are dominant, increases the probability of electron loss dramatically. Radiative recombination of electrons with ambient O^+ ions, the usual loss process for electrons at high altitude, is orders of magnitude less efficient than dissociative recombination of electrons with such molecular ions as H_2O^+ and H_2^+ which are formed by charge exchange collisions of rocket exhaust molecules with O^+.

Another opportunity to observe such chemical modification of the ionosphere [Mendillo, Baumgardner, and Klobuchar, 1979] was provided by the HEAO-C launch from Cape Kennedy on 20 September 1979. As in the preceding case, this launch burned within the F-region. This time, however, the launch was nocturnal. Coordinated experiments by several institutions were carried out to characterize the HEAO-C ionospheric hole and observe its effects on HF propagation paths. Preliminary results of these experiments were presented at a "HEAO-HOLE" workshop/symposium in November, 1979. These results have been summarized recently [Proceedings, 1980 a,b], [Mendillo, Rote, and Bernhardt, 1980]. The HEAO-C results will certainly assist the evaluation of environmental influence of booster rockets, e.g., for the proposed Solar Power Satellite program and for the Space Shuttle. They will also help the mathematical model development for description of the ionospheric hole dynamics, which is the subject of the present article.

Part of NRL's involvement in the HEAO-HOLE experiments [Proceedings, 1980 a,b] was the measurement from Bermuda of the Faraday rotation of VHF signals from the geostationary satellites ATS-3 and ATS-5, which were located at azimuth angles 231° and 187°, respectively, with respect to Bermuda. The eastward Centaur booster rocket trajectory passed under the raypaths associated with the VHF transmissions from these satellites, passing as close as 47 km to the ATS-3 raypath and 220 km to the ATS-5 raypath (to the receiver). The measurements enabled a determination of the total electron content (TEC) time profiles associated with these raypaths. TEC dropout responses to the booster rocket perturbation were dramatic, and these depletions lasted until sunrise. The early-time development of the ionospheric hole, i.e., the first ten minutes or so after the onset

of TEC reduction, is particularly detai ' i these results [Reilly et al, 1981]. This is the "hole formation" phase, during which time the rocket exhaust has expanded out to the raypath and is chemically removing electrons along it. Other processes which affect electron concentration along the raypath tend to occur on a substantially longer time scale. Hence, the early-time TEC reduction edge data is a proving ground for mathematical models of the rocket exhaust expansion, given the relatively well understood recombination chemistry. Recently, Zinn et al [1980] have employed their computation model to simulate the Bermuda TEC reduction edge results. Their model for the ionospheric hole dynamics is apparently two-dimensional and includes rocket exhaust expansion as thermalized mutual diffusion, augmented by an initial ballistic expansion regime. Their results for the ATS-3 (the closer raypath) TEC reduction edge are very good, although they underestimate the steepness of the TEC reduction edge for the ATS-5 raypath. The present author also attempted to simulate the TEC reduction edge data [Reilly et al 1981], using a brute force three-dimensional calculation in which rocket exhaust expansion was described as thermalized mutual diffusion in the neutral atmospheric density background. The agreement with the ATS-3 results was again very good, but the ATS-5 TEC reduction edge steepness was rather badly underestimated [Reilly, 1980].

There is no good theoretical justification for using a thermalized, mutual diffusion model for the rocket exhaust expansion for releases at high altitudes (e.g., $\geq$ 350 km), other than the simplification it provides. Schunk [1978] shows that such a model typically and incorrectly neglects non-linear acceleration, viscous stress, certain velocity and temperature-dependent terms in the collision frequency, and thermospheric winds. Furthermore, as Bernhardt [1979] mentions, only the last of various expansion regimes is described by the mutual diffusion model. The earlier self-continuum and collisionless expansion phases are incorrectly neglected, along with the processes by which the rocket exhaust becomes thermalized with the ambient background. All these effects, particularly important for high altitude releases, can be included by going back to the Boltzmann equation and obtaining better solutions of it. This was Bernhardt's approach [1979], and is the path followed here and reformulated next in terms of a different collisional relaxation model.

KINETIC MODEL

The Boltzmann equation for the distribution function $f_r(\underset{\sim}{x},\underset{\sim}{v}_r,t)$ for particle species r is

$$\frac{\partial f_r}{\partial t} + \underset{\sim}{v}_r \cdot \frac{\partial f_r}{\partial \underset{\sim}{x}} + \underset{\sim}{a}_r \cdot \frac{\partial f_r}{\partial \underset{\sim}{v}_r} = \left(\frac{\delta f_r}{\delta t}\right)_{coll} + P_r - L_r, \tag{1}$$

where $\underset{\sim}{x},\underset{\sim}{v}_r$, and t are coordinates for position, velocity, and time, and $\underset{\sim}{a}_r$ is the force per unit mass on particle r. The three terms on the right of this equation represent the contributions of collisions, external production (sources), and external losses. In principle, chemical reactions can be included in these terms (e.g., see Chapter 7 of [Burgers, 1969]), but since rocket exhaust expansion is primarily of interest here, this will not be done now. The concentration $n_r(\underset{\sim}{x},t)$, mean velocity $\underset{\sim}{u}_r(\underset{\sim}{x},t)$, and temperature $T_r(\underset{\sim}{x},t)$ are obtained from zero'th, first, and second moments of the distribution function in the usual way:

$$n_r[1,\underset{\sim}{u}_r,u_r^2 + 3kT_r/m_r] = \int_{-\infty}^{\infty} d^3\underset{\sim}{v}_r\, f_r[1,\underset{\sim}{v}_r,v_r^2]. \tag{2}$$

The collision term assumed here is of the form

$$\left(\frac{\delta f_r}{\delta t}\right)_{coll.} = \sum_s \kappa_{rs}\, n_s[n_r\, \phi_s - f_r] = \sum_s \left(\frac{\delta f_{rs}}{\delta t}\right), \tag{3}$$

where the terms on the right, except for $n_r(\underset{\sim}{x},t)$ and $n_s(\underset{\sim}{x},t)$, are functions of $\underset{\sim}{x},\underset{\sim}{v}_r$, and t, in general, and the Maxwellian distribution function is given as

$$\phi_s = (\underset{\sim}{x},\underset{\sim}{v},t) = (\pi c_s^2)^{-3/2} \exp[-(\underset{\sim}{v} - \underset{\sim}{u}_s(\underset{\sim}{x},t))^2/c_s^2],$$

where
$$c_s(\underset{\sim}{x},t)^2 \equiv 2k\, T_s\ (\underset{\sim}{x},t)/m_s. \tag{4}$$

Thus far, the collisional relaxation term is of the form used previously by Baum [1974,1974] and

Bernhardt [1979]. Baum's case was highly specialized, however, because of his assumption of constant collision frequencies. Whatever collisional relaxation model is employed, it has been traditional to require that particle number, momentum, and energy are conserved in short-range, elastic encounters. If the rate of change of number, momentum, and energy densities of type r particles in collisions with type s particles is given by

$$\frac{\delta}{\delta t}\left[(n_r)_s, (\underset{\sim}{p}_r)_s, (\epsilon_r)_s\right] = \int_{-\infty}^{\infty} d^3\underset{\sim}{v}_r \left(\frac{\delta f_{rs}}{\delta t}\right)\left[1, m_r\underset{\sim}{v}_r, \frac{1}{2}m_r v_r^2\right] \tag{5}$$

the aforementioned conservation conditions are

$$\frac{\delta (n_r)_s}{\delta t} = 0 = \frac{\delta (p_r)_s}{\delta t} + \frac{\delta (p_s)_r}{\delta t} = \frac{\delta (\epsilon_r)_s}{\delta t} + \frac{\delta (\epsilon_s)_r}{\delta t}. \tag{6}$$

If κ_{rs} is dependent on $\underset{\sim}{v}_r$, as in Bernhardt's collisional model, then it is not difficult to show that none of these conditions is satisfied in the general case. All of these conditions are satisfied in Boltzmann's original integral expression for collisions (e.g., see [Burgers, 1969]). This is also true for Eq. (3) when κ_{rs} depends only on $\underset{\sim}{x}$ and t and satisfies the auxiliary conditions

$$m_r \kappa_{rs} = m_s \kappa_{sr}, \tag{7}$$

as is easily verified. The so-called "BGK approximation" for collisions similarly neglects the velocity dependence of κ_{rs} in Eq. (3), although the function ϕ_s in this equation is then not quite that defined by Eq. (4) [Gross and Krook, 1956], [Burgers, 1969]. κ_{rs} is thus assumed to depend only on $\underset{\sim}{x}$ and t here, and it is chosen to obtain agreement with the momentum transport equation for unrestricted velocity and temperature differences, where collisions are included through the Boltzmann collision term treatment of hard elastic sphere forces between particles in appropriate Maxwellian distributions (see pg. 72 of [Burgers, 1969]). The final result is

$$\kappa_{rs}(\underset{\sim}{x},t), = \frac{8}{3\sqrt{\pi}} \frac{m_s}{m_r + m_s} \alpha_{rs} \pi \sigma_{rs}^2 \Phi(\epsilon_{rs}) \tag{8a}$$

where Φ is a function of

$$\underset{\sim}{u}_{rs} \equiv \underset{\sim}{u}_s - \underset{\sim}{u}_r \quad \text{and} \quad \alpha_{rs}^2 \equiv 2k\left(\frac{T_r}{m_r} + \frac{T_s}{m_s}\right) \tag{8b}$$

through $\epsilon_{rs} \equiv |\underset{\sim}{u}_{rs}|/\alpha_{rs}$:

$$\Phi(\epsilon) \equiv \frac{3\sqrt{\pi}}{8}\left(\epsilon + \frac{1}{\epsilon} - \frac{1}{4\epsilon^3}\right)\mathrm{erf}(\epsilon) + \frac{3}{8}\left(1 + \frac{1}{2\epsilon^2}\right)\exp(-\epsilon^2) \tag{8c}$$

$$\approx \left[1 + \frac{9\pi}{64}\epsilon^2\right]^{1/2}.$$

Here $\pi\sigma_{rs}^2$ is the hard sphere collision cross-section. Evidently, κ_{rs} satisfies Eq. (7). The approximation in Eq. (8c) is good to within 1.5% for all ϵ, and substantially simplifies computations. A similar approximation was suggested by Bernhardt [1979] for his collision frequency.

For the rocket exhaust expansion case it is assumed that there are only two particle types present, exhaust (e) and ambient (a). It is further assumed, for simplicity, that the ambient background paremeters are unperturbed during the expansion phase of the rocket exhaust. As discussed by Bernhardt [1979] this approximation is good when the mass density of the exhaust at the mean free path distance from the source is much less than the ambient mass density. Preliminary estimates indicate that the approximation should be not bad for the HEAO-C case, but poor for the SKYLAB case. Another consequence of the approximation is that self-collision effects can be ignored, except in the specification of the rocket exhaust source. Other source and sink effects for the rocket exhaust (e.g., chemical reactions) are also ignored, which should be valid for the early-time rocket expansion phase [Forbes and Medillo, 1976]. The simplified Boltzmann equation for the rocket exhaust is

$$\frac{\partial f}{\partial t} + \underset{\sim}{v} \cdot \frac{\partial f}{\partial \underset{\sim}{x}} + \underset{\sim}{a} \cdot \frac{\partial f}{\partial \underset{\sim}{v}} = \nu_s(\underset{\sim}{x},t)[n\phi_s - f] + P, \tag{9a}$$

where
$$\nu_s(\underset{\sim}{x},t) = \kappa_{es}(\underset{\sim}{x},t)\ n_s(\underset{\sim}{x}) \tag{9b}$$

is the collision frequency in terms of κ_{ea} given by Eq. (7). Except for this term, the subscript e has been dropped for ease of notation.

The source term P in equation (8) has the form

$$P(\underset{\sim}{x},\underset{\sim}{v},t) = Q(t)\psi(\underset{\sim}{v})\delta[\underset{\sim}{x}-\underset{\sim}{S}(t)], \tag{10}$$

where Q(t) is the number of exhaust molecules emitted per second by the engine at time t, $\underset{\sim}{S}(t)$ is the position coordinate of the source at time t, and $\psi(\underset{\sim}{v})$ represents the velocity distribution of the exhaust molecules at the source. For the rocket engine source, the exhaust molecules undergo isentropic expansion which converts their initial thermal energy into ordered, directed motion at a hypersonic limiting velocity along radial streamlines which emanate from a common source point. The transition from isentropic expansion to translational "freezing" along these streamlines has been described by Grundy [1969]. Transforming to the rest frame of the rocket, denoted by primes, so that

$$\underset{\sim}{v} = \underset{\sim}{v}' + \dot{\underset{\sim}{S}} \qquad \text{and} \qquad \psi'(\underset{\sim}{v}') = \psi(\underset{\sim}{v}),$$

one may then write in spherical coordinates

$$\psi'(\underset{\sim}{v}') = 2\pi(\underset{\sim}{v}'^2)^{-1}\,\delta(v' - w_s)\,f(\cos\theta'),$$

where w_S is the hypersonic limiting speed, and $f(\cos\theta')$ is the relatively sharp angular distribution about the axis along the direction of $-\dot{\underset{\sim}{S}}$ [Draper and Hill, 1966]. This is normalized according to $\int_{-1}^{1} dx f(x) = 1$. The temperature effect here is relatively unimportant. In the rest frame of the earth observer the corresponding angular distribution about the direction of $\dot{\underset{\sim}{S}}$ is found from the relation

$$\tan\theta = \sin\theta'/(\gamma - \cos\theta'); \quad \gamma \equiv S/w_s$$

For γ not too close to 1, relevant θ values are close to 0 ($\gamma > 1$) or π ($\gamma < 1$). This suggests a further simplification:

$$\psi(\underset{\sim}{v}) \approx \delta[\underset{\sim}{v} - \underset{\sim}{u}_s(t)],$$

where
$$\underset{\sim}{u}_s(t) \equiv \dot{\underset{\sim}{S}}(t)\ (1 - \gamma^{-1}) \tag{11}$$

This is a special case of a source distribution of the form of Eq. (4), which can also be used for exploding cannisters, where the temperature effect is important.

The solution of Eq's (8) - (11) can proceed in the manner described by Bernhardt [1979]. Coupled integral equations of the second kind are obtained for the concentration, mean velocity, and temperature. A particularly elegant way to arrive at these equations starts with the definition of new variables $\underset{\sim}{x}(t')$, $\underset{\sim}{v}(t')$, which take on the values of $\underset{\sim}{x}$ and $\underset{\sim}{v}$ when $t' = t$ and satisfy

$$\frac{d\,\underset{\sim}{x}(t')}{dt'} = \underset{\sim}{v}(t'), \quad \frac{d\,\underset{\sim}{v}(t')}{dt'} = \underset{\sim}{a} \tag{12a}$$

For constant $\underset{\sim}{a}$

$$\underset{\sim}{v}(t') = \underset{\sim}{v} + \int_t^{t'} dt_1\,\underset{\sim}{a} = \underset{\sim}{v} - \underset{\sim}{a}(t - t')$$

$$\underset{\sim}{x}(t') = \underset{\sim}{x} + \int_t^{t'} dt_1\,[\underset{\sim}{v} - \underset{\sim}{a}(t - t_1)] = \underset{\sim}{x} - \underset{\sim}{v}(t - t') + \frac{1}{2}\underset{\sim}{a}(t - t')^2. \tag{12b}$$

Eq. (8) becomes

$$\frac{df}{dt} + \nu_a f = P + \nu_a n \phi_a \equiv H \tag{13}$$

,where $\underset{\sim}{x}$ and $\underset{\sim}{v}$ have been replaced by the new variables in Eq. (12b). For a source which is switched on at t=0

$$f(t) = \int_0^t dt' \exp[-\int_{t'}^t dt_2 \nu_a(t_2)] H(t') \qquad \text{,or}$$

$$f(\underset{\sim}{x},\underset{\sim}{v},t) = \int_0^t dt' \exp[-\int_0^{t-t'} dt_2 \nu_a(t - t_2)]\ H(\underset{\sim}{x}(t'),\underset{\sim}{v}(t'),t')$$

Evidently, this suggests the substitution, $t_1 = t - t'$, so that

$$f(\underset{\sim}{x},\underset{\sim}{v},t) = \int_0^t dt_1\ G_2(\underset{\sim}{x},\underset{\sim}{v},t_1)\ H(\underset{\sim}{x} - \underset{\sim}{v}t_1 + \frac{1}{2}\underset{\sim}{a}t_1^2, \underset{\sim}{v} - \underset{\sim}{a}t_1, t - t_1)$$

,where

$$G_2(\underset{\sim}{x},\underset{\sim}{v},t) \equiv \exp[-\int_0^{t_1} \nu_a(\underset{\sim}{x} - \underset{\sim}{v}t_2 + \frac{1}{2}\underset{\sim}{a}t_2^2, t - t_2)dt_2] \tag{14}$$

This is essentially the expression obtained by Bernhardt [1979], with the exception of the collision frequency arguments for the different collision models. The integral equations for concentration, mean velocity, and temperature are obtained by an application of Eq. (2). The results are essentially those given by Bernhardt's Eq. (12), with slight corrections.

If one further specializes to the case $\underset{\sim}{a} = 0$ (e.g., neglect of gravity for early time behavior and treats the case of a point impulse source at the origin, so that

$$Q(t) = Q_0\,\delta(t) \qquad \text{and} \qquad S(0) = \underset{\sim}{0}, \tag{15}$$

then application of Eq. (2) to Eq. (14) gives

$$\begin{aligned} n[1, \underset{\sim}{u}, u^2 + 3kT/m] &= Q_0\, G_3^{(0)}(\underset{\sim}{x},0,t)\psi(\underset{\sim}{x}/t)t^{-3}[1, \underset{\sim}{x}/t, (\underset{\sim}{x}/t)^2] \\ &+ \int_0^t dt_1 \int_{-\infty}^{\infty} d^3\underset{\sim}{x}_1\ G_3^{(0)}(\underset{\sim}{x}, \underset{\sim}{x} - \underset{\sim}{x}_1, t_1)\ \nu_a(\underset{\sim}{x} - \underset{\sim}{x}_1, t - t_1)\ n(\underset{\sim}{x} - \underset{\sim}{x}_1, t - t_1) \\ &\times \phi_a(\underset{\sim}{x}_1/t_1)t_1^{-3}[1, \underset{\sim}{x}_1/t_1, (\underset{\sim}{x}_1/t_1)^2] \end{aligned} \tag{16a}$$

as the coupled integral equations for concentration, mean velocity, and temperature, where

$$G_3^{(0)}(\underset{\sim}{x},\underset{\sim}{x}',t_1) \equiv \exp\left[-\int_0^{t_1} dt_2\ \nu_a\left(\underset{\sim}{x} - \frac{t_2}{t_1}(\underset{\sim}{x} - \underset{\sim}{x}'), t - t_2\right)\right]. \tag{16b}$$

The first term on the right of the concentration equation in Eq. (16a) is denoted by n_0, and n_1 is defined by

$$n(\underset{\sim}{x},t) = n_0(\underset{\sim}{x},t) + n_1(\underset{\sim}{x},t)$$

where

$$n_0(\underset{\sim}{x},t) = Q_0\, G_3^{(0)}(\underset{\sim}{x},0,t)\ \delta[\underset{\sim}{x} - \underset{\sim}{u}_s t], \tag{17}$$

using Eq. (11) and the definition $\underset{\sim}{u}_s = \underset{\sim}{u}(0)$. From Eq. (16a) $n(\underset{\sim}{u}_s t,t) = n_0(\underset{\sim}{u}_s t,t)$, $\underset{\sim}{u}(\underset{\sim}{u}_s t,t) = \underset{\sim}{u}_s$, and $T(\underset{\sim}{u}_s t,t) = 0$. For $\underset{\sim}{x} \neq \underset{\sim}{u}_s t$, $n(\underset{\sim}{x},t) = n_1(\underset{\sim}{x},t)$, and this concentration, the mean velocity $\underset{\sim}{u}(\underset{\sim}{x},t)$ and the temperature $T(\underset{\sim}{x},t)$ are found as solutions of the following equation set:

$$f^i(\underline{x},t) \equiv n_1(\underline{x},t)\ [1, \underline{u}(\underline{x},t),\ u^2 + 3kT(\underline{x},t)/m] \qquad (i = 0,1,2,3,4)$$

$$= g^i(\underline{x},t) + \int_0^t dt_1 \int_{-\infty}^{\infty} d^3x_1\ K^i(\underline{x},t;\underline{x}_1,t_1)\ f^0(\underline{x} - \underline{x}_1, t - t_1)$$

$$g^i(\underline{x},t) \equiv \int_0^t dt_1\ b(t_1)\ \phi_s\ (\underline{V}) t_1^{-3}\ [1, \underline{V}, V^2]$$

$$\underline{V} \equiv (\underline{A}/t_1) + \underline{u}_s,\ \underline{A} \equiv \underline{x} - \underline{u}_s t \quad (\neq \underline{0})$$

$$b(t_1) \equiv Q_3\ G_3^{(0)}(\underline{x}, \underline{u}_s(t - t_1), t_1)\ G_3^{(t-t_1)}(\underline{u}_s(t - t_1), \underline{0}, t - t_1)\, \nu_a[\underline{u}_s(t - t_1), t - t_1]$$

$$K^i(\underline{x},t;\underline{x}_1,t_1) = G_3^{(0)}(\underline{x}, \underline{x} - \underline{x}_1, t_1)\, \nu_a(\underline{x} - \underline{x}_1, t - t_1) F^i(\underline{x}_1, t_1)$$

$$F^i(\underline{x}_1,t_1) \equiv \phi_a(\underline{x}_1/t_1) t_1^{-3} [1, \underline{x}_1/t_1, (\underline{x}_1/t_1)^2] \qquad (18)$$

DISCUSSION

It is planned to solve Eq. (18) for the point source problem and to demonstrate departures from a thermalized diffusion model. The full rocket exhaust effect can be included by summing point source puffs along the rocket trajectory. Results will be aimed at simulating the TEC reduction edges measured for both the HEAO-C launch and, more recently, the NOAA-B launch from Vandenberg on May 29, 1980 [Reilly et al, 1981]. The first step in these simulations is the determination of time profiles of rocket exhaust molecule concentrations along the associated raypaths.

There are many effects to unravel. While it may be thought that a thermalized diffusion model would tend to overestimate the rate of rocket expansion, due to the effects mentioned by Schunk [1978], there is at least one effect working oppositely. For example, it will be noted from Eq's (8a), (8b), and (9b) that, as thermalization of rocket exhaust takes place, the collision frequency increases along with α_{ia}. Hence, during early times the expansion rate tends to be underestimated by a thermalized diffusion model, because of this effect, which can be quite dramatic for light exhaust molecules (e.g., H_2).

At this writing, the intent is to approximate the integral in Eq. (18) by standard numerical quadrature techniques, and solve the equations iteratively for successively later times. Results will be reported at a later time.

REFERENCES

1. Bernhardt, P.A., "High-Altitude Gas Releases. Transition from Collisionless Flow to Diffusive Flow in a Nonuniform Atmosphere", J. Geophys. Res., 84, p 4341, 1979.

2. Burgers, J.M., "Flow Equations for Composite Gases", Academic Press, New York, 1969.

3. Forbes, J.M. and M. Mendillo, "Diffusion Aspects of Ionospheric Modification by the Release of Highly Reactive Molecules into the F-Region", J. Atm. and Terr Physics, 38, p 1299, 1976.

4. Gross, E.P. and M. Krook, "Model for Collision Processes in Gases: Small-Amplitude Oscillations of Charged Two-Component Systems", Phys. Rev., 102, p 593, 1956.

5. Grundy, R.E. "Axially Symmetric Expansion of a Monatomic Gas from an Orifice into a Vacuum", Phys of Fluids, 12, p 2011, 1969.

6. Hill, J.A.F. and J.S. Draper, "Analytical Approximation for the Flow from a Nozzle into a Vacuum", J. Spacecraft, 3, p 1552, 1966.

7. Mendillo, M., G.S. Hawkins, and J.A. Klobuchar, "A Sudden Vanishing of the Ionospheric F Region Due to the Launch of Skylab", J. Geophys. Res., 80, p 2217,1975.

8. Mendillo, M., J. Baumgardner, and J.A. Klobuchar, "Opportunity to Observe a Large Scale Hole in the Ionosphere", EOS, Trans. Amer. Geophys. Union, 60, p 513, July 1979.

9. Mendillo, M., D. Rote, and P.A. Bernhardt, "Preliminary Report on the HEAO Hole in the Ionosphere", EOS, Trans Amer. Geophys. Union, 61, p 529, July 1980.

10. Proceedings of the Workshop/Symposium on the Preliminary Evaluation of the Ionspheric Disturbances associated with the HEAO-C launch, with Applications to the SPS Environmental Assessment, M. Mendillo and B. Baumgardner, Ed's., D.O.E. Report Conf-7911108,1980a and Report for D.O.E. contract #31-109-38-5326,1980b.

11. Reilly, M.H., "Interpretation of Bermuda Polarimetry Data for the HEAO-C Ionospheric Hole", U.R.S.I. North American Radio Science Meeting, Session GI., Quebec, June 3, 1980.

12. Reilly, M.H., L.O. Harnish, and J.M. Goodman, "Polarimetry Studies of Ionospheric Modification by Rocket Boosters", NRL Memo Report, to be published 1981.

13. Schunk, R.W., "On the Dispersal of Artificially-Injected Gases in the Night-time Atmosphere", Planet. Space Sci., 26, p 605, 1978.

14. Zinn,J., C.D. Sutherland, S.N. Stone, L.M. Duncan, and R. Behnke, "Ionospheric Effects of Rocket Exhaust Products -- HEAO-C, SKYLAB, and HLLV", Los Alamos Scientific Lab. Report LA-UR-80-1160-REV, 1980.

OCCULTATION OF THE ATS-3 SATELLITE BY THE AVEFRIA BARIUM ION CLOUD

T. J. Fitzgerald
EG&G, Inc.
Los Alamos, NM 87544

D. J. Simons, M. B. Pongratz
Los Alamos National Laboratory
Los Alamos, NM 87545

and

J. R. Clynch
Applied Research Laboratories
University of Texas, Austin, TX

ABSTRACT

During the AVEFRIA DOS barium release experiment, sponsored by the Los Alamos National Laboratory and the Defense Nuclear Agency in May 1978, the line of sight from one of the ground observation stations to the ATS-3 satellite was occulted by the barium ion cloud for a period of approximately five minutes. Optical measurements of the structured barium ion cloud were made with intensified cameras using the 455.4-nm wavelength fluorescent ion line. These measurements have been related to barium ion column density. During the occultation, the amplitude scintillations of the 136.47-MHz signal from the ATS-3 satellite were monitored. The optical measurements have been used to correlate the barium column density with the total electron content measurements and to calculate the scintillation index, S_4, and the two dimensional intensity pattern for comparison with the measured amplitude scintillations.

INTRODUCTION

We have calculated the diffraction pattern induced on the 136 MHz signal from the ATS-3 satellite by an ionized barium plasma produced during the AVEFRIA barium injection of May, 1978.[1] The experiment was sponsored by the Los Alamos National Laboratory and the Defense Nuclear Agency. Neutral barium was injected at an altitude of 190 km over the Tonopah test range in Nevada before dawn when the solar depression angle was 10°. The neutral barium was photoionized by sunlight at the release altitude. The ionized barium was observed photographically through its fluorescence emission using filtered, intensified cameras at a number of stations on the ground. At one of the ground stations, Antelope Valley, Nevada, the signal from the ATS-3 geo-synchronous satellite was monitored by a group from the Applied Research Laboratories of the University of Texas at Austin under J. Clynch. For a period of about five minutes, beginning approximately three minutes after the release, the line of sight from the ground station to the satellite was occluded by the barium ion cloud. By that time the ion

cloud had developed a number of striations. Amplitude scintillations were observed in the 136-MHz signal during the occultation. We calculated the scintillation pattern based on the electron column densities as deduced from the photographic images of the ion cloud and compared the calculated intensity with measurements. We have assumed the ion cloud acted as phase changing screen and calculated the diffraction pattern at the ground.

PHASE SCREEN CALCULATION

Diffraction by a phase changing screen can be treated by the methods of scalar diffraction theory.[2] A particularly useful description can be obtained in terms of the angular spectrum, the Fourier transform of the scalar field. If $U(x,y,0)$ is field at $z=0$ after emerging from the phase screen, the angular spectrum, $A_0(f_x,f_y)$, is

$$A_0(f_x,f_y) = \int_{-\infty}^{\infty}\int U(x,y,0) \exp\left[-i2\pi (f_x x + f_y y)\right] dxdy$$

where f_x and f_y are the spatial frequencies. The effect of propagation through a distance, z, is to apply a linear filter to the angular spectrum. That is, the angular spectrum at distance z, $A(f_x,f_y,z)$, is

$$A(f_x,f_y,z) = A_0(f_x,f_y) \exp\left[i2\pi \frac{z}{\lambda}\sqrt{1 - (\lambda f_x)^2 - (\lambda f_y)^2}\right]$$

where λ is the wavelength of the radiation. The field at distance z is then the inverse Fourier transform of $A(f_x,f_y,z)$. The calculation of the diffraction pattern can be carried out using fast Fourier transform algorithms as described by Buckley for a one-dimensional screen.[3] The extension to two dimensions is straightforward. In our calculation, radiances are first converted to phase change so that we obtain a two-dimensional complex array representing the field emerging from the barium cloud. (We assume that the curvature of the field arising from the finite distance to the cloud has a negligible effect on the scintillation pattern.) We take the complex two-dimensional transform of the phase array. After applying the appropriate phase change to each element of the transform to represent the propagation to the ground, we take the inverse Fourier transform. Finally, we obtain the magnitude square of the complex field to represent the intensity at the ground.

OPTICAL DATA

The barium ion cloud was observed in 455.4-nm fluorescence light using radiometrically calibrated cameras. Peak radiances of approximately 3×10^{-9} W/cm^{2}sr were observed. Ten photographic images were scanned and processed to produce 256 x 256 arrays of radiance versus position. The exposure time of the cameras was .5 s; the time between the images processed was approximately 30 s. A reproduction of one of the processed images at 236 s after release is shown in Figure 1. The cross hairs indicate the direction of the satellite. The ion cloud image moved almost exactly horizontal across the film plane. The striations in the cloud, which were aligned with the geomagnetic field, appear tilted off the vertical by about 30°. The occultation therefore begins towards the top of the leading edge of the cloud, proceeds through the center of the cloud, and ends towards the bottom of the trailing region of the

cloud. Although the gross structure of the cloud remains relatively constant from frame to frame, there are significant changes in detailed features and also in the overall size of the cloud during the occultation. Figure 2 shows the cloud image at 295 s after release and can be compared to Figure 1 to indicate the motion and the development from the earlier time.

In order to calculate the intensity at a point on the ground versus time, we have had to assume that the intensity pattern calculated from a single frame can be translated with the overall motion of the cloud. Clearly, because of changes in the cloud, we do not expect this assumption to be rigorously true even for the 30 s between frames and we cannot expect continuity from frame to frame.

The relatively low signal-to-noise in the original image limits the extent to which the intensity pattern can be accurately calculated. The noise arises from film grain, inhomogeneities in the intensifier, and photon fluctuations. The signal content is limited by the inherent lack of structure in the barium cloud, the low intensity of the emitted light, and by lack of resolving power of the camera system. It is also true that the cloud as observed from the ground was folded back on itself, which decreased the contrast of the image.

We have attempted to compensate for the transfer function of the camera system and to reduce the noise by applying a Wiener filter to the image. Since the apparent structure in the cloud decreased with time, the parameters of the filter were changed to give more smoothing at later times. This filtering is important because the coherent radio wave, which is assumed to be incident on the cloud, accentuates the noise in the image. That is, the scintillation pattern is noisier than the image. The calculated scintillation index is therefore, to some extent, a function of the filtering.

Systematic errors up to 30% in the scintillation index could arise from systematic errors in the radiometric calibration. An error in the triangulated distance to the ion cloud does not significantly affect the calculated intensity pattern or the scintillation index.

RESULTS OF CALCULATION

The calculated intensity patterns corresponding to the images in Figures 1 and 2 are shown in Figures 3 and 4. Again, the cross hairs mark the direction of the satellite. Light areas correspond to high values of intensity. The circular diffraction patterns away from the cloud arise from star images or defects on the film.

Figure 5 shows the phase change along a line through the center of the cloud perpendicular to the striations starting from the right-hand side in the data shown in Figure 1. The maximum phase change is almost 30 radians above background. (This profile is through the center of the cloud and not in the satellite direction). The intensity profile corresponding to this phase change profile is plotted in Figure 6. We note that the structure in the intensity profile is more complicated and more pronounced than in the phase change profile. There is some evidence of focusing with a maximum of the phase change corresponding to a minimum of intensity. Similar profiles for the data at 295 s are shown in Figures 7 and 8. Comparing these results to those for a random phase screen obtained by Buckley, it appears that we are beyond the distance of strong focusing for the radio wavelength used and are in the regime

of scintillations with spatial scales smaller than that of the ion cloud. It also appears that the intensity pattern is controlled by phase variations of the order of a few radians that are superimposed on the larger phase changes due to the gross structure of the cloud.

COMPARISON WITH RADIO MEASUREMENTS

Figure 9 shows the electron column density along the line of sight to the satellite as derived from the image data assuming an optically thin ion cloud. These results are consistent in time development with the total electron content derived from Faraday rotation measurements but show a peak column density of only approximately 75% of that derived from the radio measurements. This discrepancy is within the uncertainty expected for the radiometry and will tend to cause the calculated scintillation index to be low. It is also possible that the assumption that the ion cloud is optically thin is not strictly true, and we are therefore underestimating the column densities.

We can compare the measured intensity of the radio signal versus time with our calculated intensity versus spatial distance. This is shown for data near the beginning of the occultation in Figure 10. We have translated spatial distance in our calculations into time using the apparent velocity of the ion cloud. We see in both curves an increase in intensity as the edge of the cloud reaches the satellite location and then a decrease as the full occultation begins.

It is apparent that our calculation lacks the resolution to follow the intensity changes with time scales of approximately 1 s. Our calculated values of the scintillation index will therefore be lower than those of the data. We also note that the calculated values do not follow the trend of the observed values after 220 s. This difference may be ascribed to the growth in spatial dimensions of the cloud, which tends to accelerate changes in the intensity pattern relative to what we calculate, assuming a stable ion cloud.

During the occultation, the growth rate of the cloud decreased with time, so that our calculated values tend to better follow the trend of the observations. This can be seen in the comparison shown in Figure 11 for the next set of data. Again, our calculations do not resolve the relatively fast fluctuations of intensity shown in the observations.

We have calculated the scintillation index, S_4, over 50-s intervals for the observations and for our calculations. These are plotted in Figure 12. S_4 reaches a maximum early in the occultation, when a region of the ion cloud with discernable striations passes in front of the satellite. Later, as regions with less apparent structure cross the signal path, the scintillation index decreases. The calculated S_4 values follow the observed values but at a lower level, which may be ascribed to the smoothing inherent in our lack of resolution.

CONCLUSIONS

We have shown that it is possible to calculate radio scintillation patterns from optical observations of barium clouds and to get general agreement with measurements. The advantage to calculating scintillations from optical data is that one gains the two-dimensional intensity pattern which, for a finite cloud, may be important. The limitations to the calculation arise from two sources that are apparent in the results presented here. First, if the barium cloud is highly structured or has steep gradients, the optical data may lack the resolution to reconstruct the relatively fine details of the diffraction pattern. This holds for our data particularly at the beginning of the occultation. Second, even if the ion cloud is not highly structured, the image may be so noisy the calculation is dominated by the scintillation pattern of the noise. This appears to be true in our results towards the end of the occultation.

ACKNOWLEDGMENTS

This work was supported by the Department of Energy under Contract No. DE-AC08-76 NV01183.

REFERENCES

[1]M. B. Pongratz, G. M. Smith, R. C. Carlos, H. G. Horak, D. J. Simons, and C. F. Lebeda, High Spatial Resolution Field-Aligned Observations of Gradient Drift Striation Formation in Barium Clouds. EOS Transactions AGU, 59, 1161 (1978).

[2]J. W. Goodman, Introduction to Fourier Optics, McGraw-Hill, 1968.

[3]R. A. Buckley, Diffraction by a Random Phase-Changing Screen: A Numerical Experiment, J. Atmos. Terr. Phys. 37, 1431 (1975).

By acceptance of this article, the publisher and/or recipient acknowledges the U. S. Government's right to retain a non-exclusive, royalty-free license in and to any copyright covering this paper.

Figure 1. AVEFRIA ion cloud at E + 236s.

Figure 2. AVEFRIA ion cloud at E + 295s.

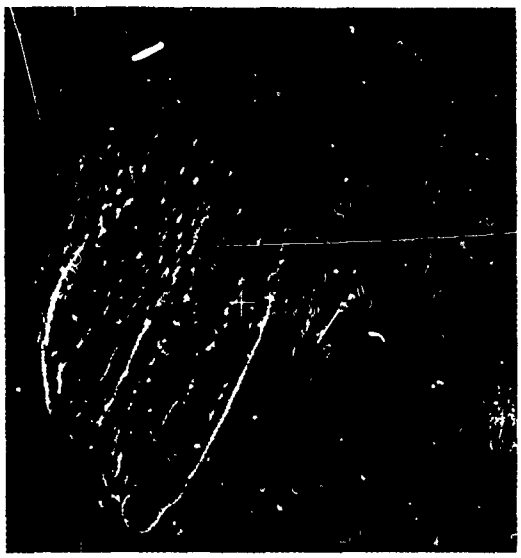


Figure 3. Calculated intensity at E + 236s.

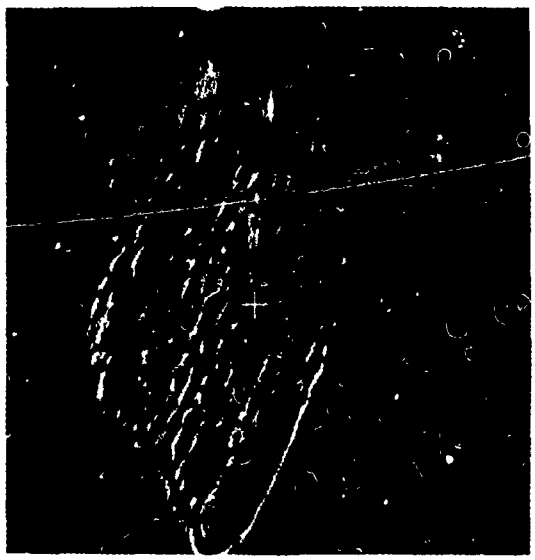


Figure 4. Calculated intensity at E + 295s.

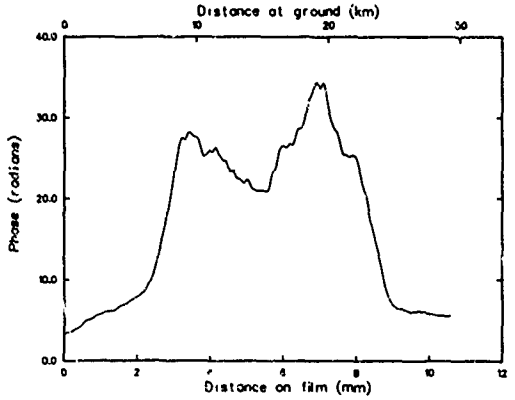


Figure 5. Phase change across cloud at E + 236s.

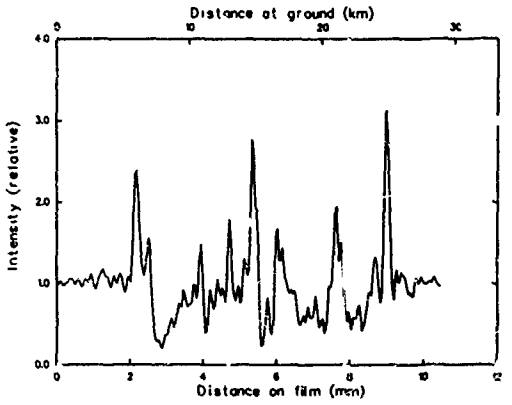


Figure 6. Calcualted intensity across cloud at E + 236s.

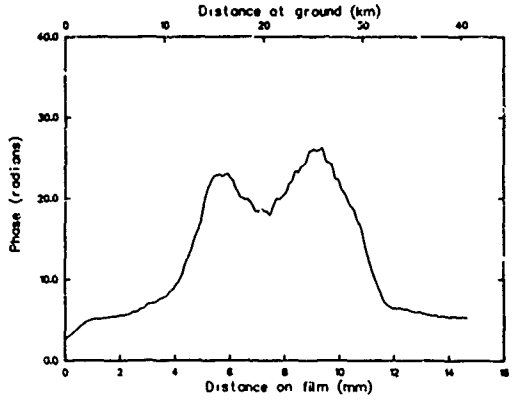


Figure 7. Phase change across cloud at E + 295s.

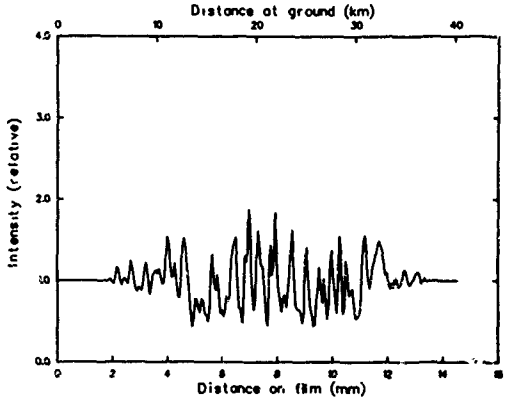


Figure 8. Calculated intensity across cloud at E + 295s.

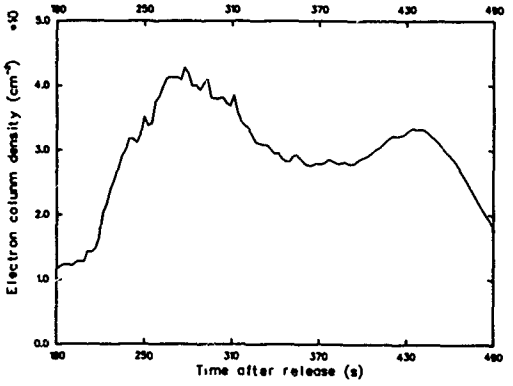


Figure 9. Electron column density in satellite direction.

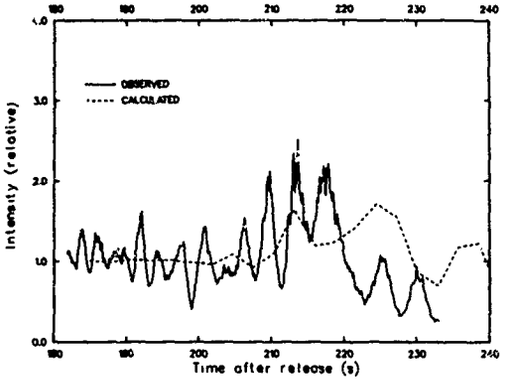


Figure 10. Comparison of intensity from data at E + 206s.

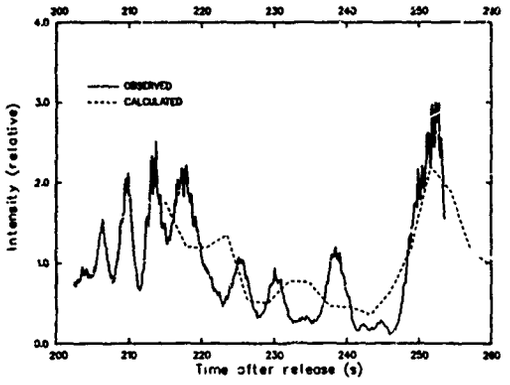


Figure 11. Comparison of intensity from data at E + 236s.

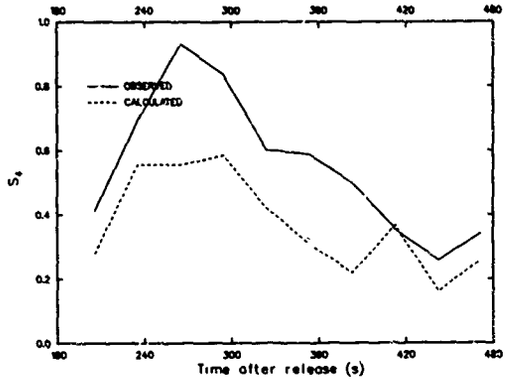


Figure 12. Comparison of calculated and observed S_4.

IN SITU ACTIVE EXPERIMENT TECHNIQUES TO STUDY THE IONOSPHERE

Morris B. Pongratz, Robert A. Jeffries, David J. Simons, and Gordon M. Smith
University of California, Los Alamos National Laboratory
Los Alamos, New Mexico 87545

INTRODUCTION

In situ active experiment techniques can be used to study the natural ionosphere, to simulate natural or artificial ionospheric disturbances, and to provide a test bed for radiowave propagation studies. We distinguish between active experimental techniques using radiowaves and active experiments conducted in situ such as barium injections or chemical releases.

The opportunity for remote optical diagnostics and a relatively short ionization time in full sunlight make barium the most popular ion-producing in situ experiment material. Low-density barium clouds can be used to visibly trace-out structures in the natural ionosphere (Gendrin's active observation category of active experiments [1]). We present results from the Periquito experiment which show low-altitude (500-km) dayside-cleft fields and structures. We also present results which show structure on auroral-zone field lines at several R_e altitude, observed during the CAMEO experiment.

Recent improvements in barium-cloud radiative-transfer calculations coupled with image-intensified optical diagnostics have permitted quantitative determination of the line-of-sight column density for dense barium clouds. Consequently, shaped-charge barium injections and thermite barium releases, used to generate dense plasmas unstable to structure producing mechanisms, can be quantitatively diagnosed for spatial wavelength-dependent growth and decay rates. Such experiments can also be used to produce a known ionospheric disturbance for use as a test bed for radiowave propagation studies. Results of the Avefria experiment that met these objectives will be presented.

Recent experiments at Los Alamos, in cooperation with Sandia National Laboratories and the National Research Council of Canada, have pioneered the technique of detonating high explosives in the F-region producing water vapor that, in turn, chemically reacts with atomic ions and produces a large depletion in the ionospheric plasma [2]. Ionospheric depletion experiments remove about 1000 times more ions than can be generated in present barium ionospheric-enhancement experiments. Our report on the Lagopedo experiments described the chemistry involved in depletions and gave results on efficiencies and scale sizes [3]. We present some observations from the Waterhole I auroral topside ionospheric-depletion experiment that "turned off" the aurora in a small space-time volume [4]. Such experiments lie within Gendrin's active perturbation category of active experiments [1].

ACTIVE OBSERVATION EXPERIMENTS FOR TRACING PLASMA IRREGULARITIES: PERIQUITO AND CAMEO

The active observation experiment utilizes a tracer material, such as barium ions, to trace or stain chemical or physical processes, facilitating diagnosis of the natural state of the space plasma. Care must be exercised to insure that the tracer material does not effect a significant perturbation to the natural state. Generally, the relative insignificance of the density or energy density of the tracer material are the basis for arguments regarding the experiment's passive nature.

The Los Alamos-sponsored Periquito Dos experiment utilized barium ions to trace out the "throat" convection pattern in the low-altitude dayside magnetospheric cusp region. The degree to which the barium clouds develop fine-scale field-aligned structures or striations largely determines the spatial shape of the space-time volume over which the barium cloud can be used to measure electric fields and convection velocities. During an earlier and similar polar-cusp injection experiment, the barium cloud did not develop fine-scale striations and consequently the high-velocity, high-altitude barium could be observed for an extended time interval and to high altitude; the spatial shape of the observed volume was tall and thin. During Periquito Dos, many fine-scale striations formed early, giving the low-velocity barium cloud the appearance of a persistent rayed auroral arc (see Fig. 1). The striation formation included a horizontal spreading, giving the spatial volume studied a wide

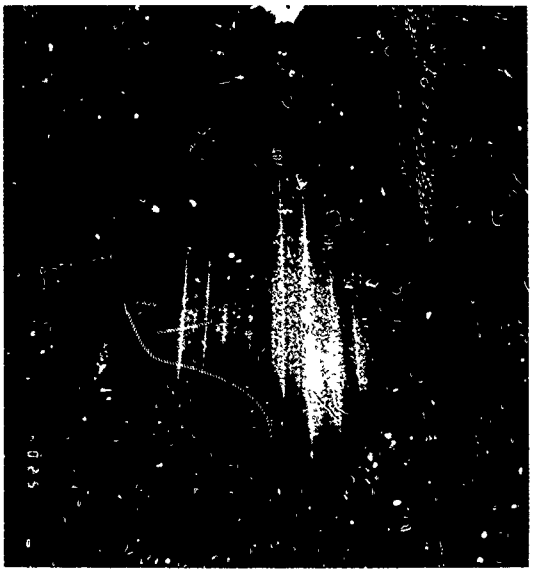


Figure 1. Periquito Dos rayed barium-ion cloud. Note that the individual striations (rays) can be identified, labeled and triangulated from time t_1 to time t_2.

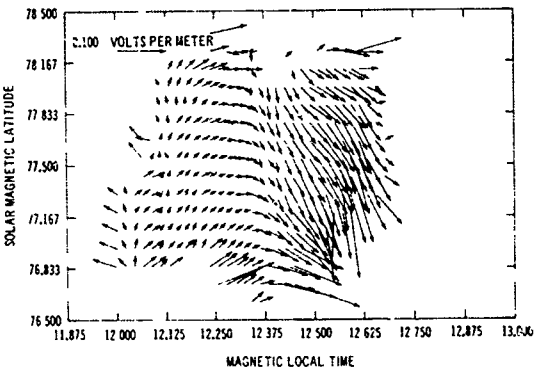


Figure 2. Periquito Dos convection electric field pattern. Under the E x B force ions convect along a direction 90° in azimuth less than the electric field aximuth. At about 12.3 magnetic local time the ions convect antisunward.

magnetic latitude and local time extent, but a short altitude interval. By identifying the persistent striations and triangulating their positions as a function of time, we were able to gain a "snapshot" of the ion-convection pattern in the polar cusp region (see Fig. 2). The component of the electric field perpendicular to B can be found from the velocity field by using the $E = -v \times B$ relationship. However, the velocity field (at a given magnetic potential) is determined by finding the distance along a surface of constant magnetic potential between the intersection of the striation and the potential surface at time t_1 and the intersection of the same striation with the same magnetic-potential surface at time t_2. The convection velocity obtained by dividing this distance by the quantity $(t_2 - t_1)$ is, of course, perpendicular to B. Only this $v_\perp$ velocity can be measured because positions of individual ions cannot be discerned. The Periquito Dos "snapshot" of the ion-convection pattern in the polar cusp region is very similar to the Hanson and Heelis model. Note that the Heelis et al. paper [5] which first identified the "throat" convection pattern, required many satellite passes through this region over an extended time interval, encompassing a myriad of geomagnetic conditions.

The CAMEO (Chemically Active Material Ejected from Orbit) experiment involved thermite barium releases from a satellite at 965-km altitude over the polar cap. A combination of magnetic mirror, gravity and parallel electric field forces resulted in the upwards acceleration of the barium ions to several R_e altitudes on auroral zone magnetic field lines [6]. Photographs of the high-altitude barium cloud show considerable structuring although the plasma is collisionless (see Fig. 3). Booker has speculated that because of geometrical considerations, such structures extending along the magnetic field to high altitudes could produce significant scintillations despite the low absolute density fluctuations [7].

THE AVEFRIA BARIUM INJECTION EXPERIMENTS: SIMULATION OF IONOSPHERIC DISTURBANCES

The Avefria experiments conducted by Los Alamos and Sandia in May, 1978 above the Tonapah Test Range had the dual objectives of studying the phenomenology of structure producing instabilities in space plasmas and providing a controlled (or at least diagnosable) ionospheric disturbance for use as a test bed for radiowave propagation studies. Sounding rockets with onboard attitude control systems were used to lift high-explosive shaped-charge barium injection payloads to approximately 195 km and orient the injection perpendicular to the geomagnetic field. As predicted by the theory of Simons et al. [8] such an injection orientation would result in the prompt formation of striations in the fast barium ion plasma.

This predicted but as yet unproven behavior prompted us to pre-position ground receivers for measurements of the downlink propagation from geostationary satellites. The observed prompt structuring coupled with a reliable rocket trajectory allowed all three fixed ground receiver sites to observe occulation of satellite downlinks by a structured barium cloud on both the Avefria Uno and Dos experiments (six out of six possible occultations). Furthermore optical diagnosis of cloud surface brightness by image-intensified cameras at each fixed ground site permitted an independent measurement of the spatial-frequency power spectral density (PSD) of the ionospheric irregularities comprising the occultation. Preliminary results indicate inconsistencies between observed downlink signals and those calculated using the measured PSD's with thin-phase screen propagation models [9].

Figure 3. The CAMEO barium-ion streak at thousands of kilometers altitude as viewed from Table Mountain Observatory in California. Note that field-aligned structures exist at high altitudes.

The Avefria experiments did serve to confirm both the prompt striations model of Simons et al. and the relatively slower gradient-drift instability model of Linson and Workman [10]. These two models each describe structuring in separate parts of the Avefria ion clouds. Figure 4 shows up the field-line views of the Avefria Dos barium-ion cloud between event plus 120 s and event plus 210 s. The left (east) portion of the ion cloud resulted from the fast ions in the shaped-charge produced jet and they structured promptly following the injection. As Fig. 4 shows the debris ion cloud on the right structured on a much slower time scale, in gross agreement with the gradient-drift model. This model is a Rayleigh-Taylor-like instability wherein the barium plasma cloud provided a conductivity gradient in which small density irregularities are unstable in much the same manner as the heavier fluid atop the lighter fluid is unstable in the classic Rayleigh-Taylor example. While the Avefria experiments demonstrated the essential validity of linear structure producing models, the nonlinear forward cascade observed in the Avefria PSD's challenges theoretical description.

"TURNING OFF" THE AURORA USING ACTIVE PERTURBATION TECHNIQUES: WATERHOLE *I*

Because present sounding rocket-launched ionospheric depletion experiments can remove about 1000 times as many ions as similar ionospheric enhancement experiments can produce, we expect that active perturbation experiments artificially perturbing the space plasma away from equilibrium, isolating and controlling preselected parameters, will favor depletion rather than enhancement techniques [11]. Waterhole I was the first such experiment. Project Waterhole was a topside auroral ionosphere modification experiment conducted by the National Research Council of Canada and the Los Alamos National Laboratory [4]. The topside density depletion was designed to reduce the charge carriers available

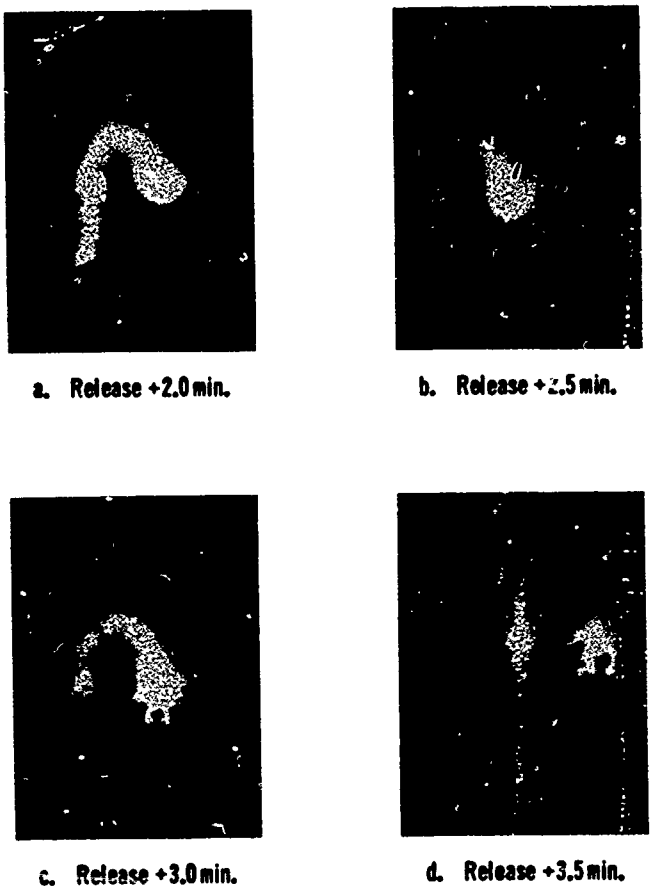


Figure 4. Up-the-field-line view of onset of Rayleigh-Taylor-like instability during the Avefria Dos experiment.

for carrying parallel currents. Parallel currents play a key role in many auroral particle energization mechanisms. By changing the impedance of the magnetoshperic-ionospheric current systems, we hoped our observations would enable us to deduce the auroral circuit dynamics. The experiment was conducted in April 1980.

The Waterhole I sounding rocket, launched from Churchill, Manitoba, carried in situ diagrostics and an 88-kg ammonium-nitrate/nitromethane explosive in a mother-daughter configuration. The explosive was detonated at 300 km altitude in the topside auroral ionosphere. The detonation occurred about 10 km north of flux tubes conjugate to visible aurora at lower altitudes. As expected, the charge-exchange/dissociative-recombination chemistry caused a rapid depletion of the thermal plasma (note that as described in Pongratz [11] the depletion technique is limited to altitudes where O^+ is the dominant ion). Yau et al. [4] have reported a factor of thirty reduction in the flux of 500-eV electrons, which coincided with the explosive detonation and persisted for about two minutes (see Figure 5). The signal is consistent with a vanishing of the energetic electrons throughout the depleted volume (about 25-km radius) and a return to pre-explosion flux levels as the depleted volume is excited by the in situ diagnostic package.

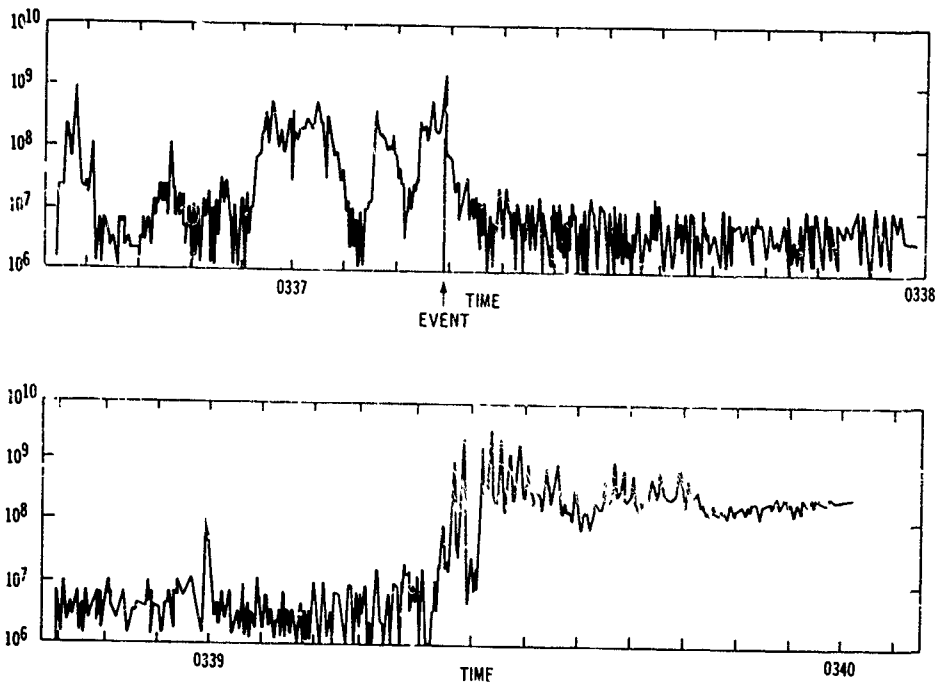


Figure 5. 500-eV differential electron flux measured during Waterhole I (figure courtesy of Dr. Andrew Yau).

REFERENCES

1. Gendrin, R., Space Science Reviews 15, 905 (1974).
2. Pongratz, M. B., G. M. Smith, C. D. Sutherland, and John Zinn, in Effect of the Ionosphere on Space and Terrestrial Systems, John M. Goodman, editor, U.S. Government Printing Office, Washington, DC, 1978.
3. Pongratz, M. B., G. M. Smith, H. G. Horak, D. J. Simons, C. D. Sutherland, J. H. Wolcott and J. Zinn, LA-UR-80-139, presented at Satellite Power Systems Program Review and Symposium, April, 1980.
4. Yau, A. W., B. A. Whalen, F. Creutzberg, M. B. Pongratz, and G. Smith, submitted to J. Geophys. Res, 1981.
5. Heelis, R. A., W. B. Hanson, and J. L. Burch, J. Geophys. Res., 81, 3803 (1976).
6. Heppner, J. P. Heppner, M. L. Miller, M. B. Pongratz, G. M. Smith, L. L. Smith, S. B. Mende, N. R. Nath, NASA Goddard Tech Note X-690-80-31, to be published in J. Geophys. Res., 1981.
7. Booker, H. G. and J. A. Ferguson, J. Atmos. Terr. Phys., 40, 803, 1978.
8. Simons, D. J., M. B. Pongratz, and S. P. Gary, J. Geophys. Res. 85, 671 (1980).
9. Fitzgerald, T. Joseph, David J. Simons, and Morris B. Pongratz, Symposium on the Effects of the Ionosphere on Radiowave Systems, April 1981.
10. Linson, L. M. and J. B. Workman, J. Geophys. Res. 75, 3211 (1970).
11. Pongratz, M. B., 1980 COSPAR Symposium on Active Experiments to be published in Advances in Space Exploration, 1981.

STIMULATION OF ULF MAGNETIC PULSATIONS BY RAPID CHANGE IN IONOSPHERIC CONDUCTIVITY

P.B. Morris and T.J. Rosenberg
Institute for Physical Science and Technology
University of Maryland, College Park, MD 20742

and

L.J. Lanzerotti
Bell Laboratories
Murray Hill, NJ 07974

ABSTRACT

Observations are reported of ground level magnetic pulsations in the Pc3 and Pc4 frequency bands associated with sudden ionospheric disturbances (SID). The SID's, which arise from solar flare hard x-ray events, were observed as absorption events on dayside 30 MHz riometers and as phase anomalies on long-path VLF propagation signals. The magnetic activity was recorded with a 4-station meridional chain of magnetometers covering a range of magnetic latitudes from $L \cong 3.2$ to $L \cong 4.4$ and at a single site in the conjugate region. In two events described, several cycles of a highly monochromatic wave were excited coincident with the arrival of the solar x-ray flux at ionospheric heights. A sudden conductivity change due to the rapid increase in ionization caused by the enhanced x-radiation flux is thought to have stimulated hydromagnetic waves at the fundamental resonance frequency appropriate for the plasma conditions on these field lines. Theoretical considerations suggest that ULF waves may also be generated if the ionospheric conductivity is decreased. Thus, modification experiments which produce ionization density depletions (e.g., chemically-induced "holes") may also be effective in stimulating hydromagnetic waves in the magnetosphere.

INTRODUCTION

Ultra-low frequency (ULF) pulsations of the earth's magnetic field are a commonly observed phenomenon. Sensitive satellite and ground-based magnetometers have provided much detailed information on the properties of pulsations in different frequency bands and as a function of magnetospheric location and geomagnetic disturbance level. The occurrence of geomagnetic pulsations has been attributed to eigenoscillations of the geomagnetic field, instabilities and wave-particle interactions, and oscillations of magnetosphere-ionosphere current systems (e.g., see reviews by Lanzerotti and Southwood, 1979; and Rostoker, 1980). Regardless of the mechanism, however, their origin is considered to result from solar wind-magnetopause interactions or from internal magnetospheric processes influenced by the general level of geomagnetic activity.

In this paper we present evidence of the excitation of Pc3 (15-45 sec) and Pc4 (45-150 sec) pulsations in the dayside magnetosphere coincident with the arrival of solar x rays in the ionosphere. Data are presented for two events in 1980 from many cases that have been observed. We believe this is the first identification of hydromagnetic wave generation not associated with the solar wind or magnetic storms. A possible mechanism is suggested to explain the phenomena.

OBSERVATIONS

Figure 1 illustrates x-ray and ionization-related data for the solar flare event that occurred near 2030 UT on April 26, 1980. The top two panels show the counting rate response of two channels (12-20 keV and 37-70 keV) of the ISEE-3 x-ray detector (courtesy of Dr. S. Kane, University of California, Berkeley). The event began shortly after 2030 UT in the lower energy channel and returned to the background level about 7½ minutes later. The impulsive character of the hard x-rays near the peak of the event is particularly evident in the 37-70 keV channel. In this channel, the duration of the x-ray enhancement is about 1 minute and consists of two bursts separated by ~ 40 seconds in the interval 2031-2032 UT.

The response of the ionosphere to the enhanced x-ray flux is shown in the third panel of Figure 1 by the sudden phase anomaly (SPA) in the 10.2 kHz Omega navigation signal from Hawaii to Washington, D.C. and in the bottom panel by the absorption of 30 MHz galactic radio noise recorded by a riometer located at Durham, New Hampshire. The onset of ionospheric response was essentially coincident with the time of peak x-ray flux. These SPA and absorption profiles, exhibiting rapid excursion (~ 1-3 minutes) to maximum disturbance, followed by a slow (~ 15-30 minutes), exponential like recovery, are characteristic of solar flare x-ray events. The rise and decay times depend on the time profile and spectral intensity of the x-ray (and ultraviolet) flux, the solar zenith angle, and the response time of the ionosphere.

The ground level magnetic variations that accompanied the ionospheric event are shown in Figure 2 for the four stations of a meridional chain operated in New Hampshire and Quebec, Canada and for a station in the Antarctic conjugate region. Figure 3 and Table I give the location of these sites, their geographic and geomagnetic coordinates, and L values. Co-located with the magnetometers at each station are 30 MHz riometers with broadbeam (60°) zenith-centered antenna patterns. The riometer data for Durham and Pittsburg, the lowest latitude sites, are shown in the bottom left and right panels, respectively, in Figure 2. Similar absorption events occurred also at the higher latitude sites. The riometer at Siple, however, did not record an absorption event.

As shown in Figure 2, an overall increase (decrease) in the magnetic field H-component (D-component) at all northern sites accompanied the ionosphere conductivity enhancement represented by the onset of riometer absorption at ~ 2031 UT (the local time of these measurements was 1531). At Siple, a similar variation occurred in the H-component, but the D-component displayed opposite behavior to that in the north. Superimposed on the gross variations are several cycles of an ~ 90-sec period oscillation. These long-period waves are quite clear in the H-component at the two lower latitude sites, but are less distinct at the higher latitudes. At Siple, two large, distinct oscillations occurred in the D-component. Lower amplitude, higher frequency oscillations (such as, for example, those at ~ 40-sec period between 2038 and 2040 UT) were present throughout the hour around the time of the x-ray event, but were not affected by the event.

To investigate the polarization and spectral properties of the low frequency waves, the magnetometer data of Figure 2 were bandpass-filtered for the period range 60-120 seconds. The filtered data are presented in Figure 4 and show clearly the excitation of the ~ 90-sec period oscillations at all sites in the interval from 2031 to 2036 UT. Thereafter, the wave amplitudes decrease, the period lengthens, and in general the character of the variations becomes less coherent over the chain. No comparable oscillations in this period band were evident prior to 2030 UT. Note that at Siple the oscillations (dashed curves) are out of phase with the northern hemisphere components. Also, the Siple waveforms (and polarization, see below) resemble more closely those of the lower latitude stations.

The polarizations in the H-D plane of the long-period magnetic variations of Figure 4 are shown for each station in Figure 5. These hodograms show that the oscillations in the interval 2032-2035 UT are left-hand polarized (counter-clockwise looking down along the field line) at the two lower latitude stations. At the higher latitudes, the polarizations, although somewhat mixed, can be characterized as right-handed. Note that the polarization at Siple is left-handed, the same as that at the lower latitude stations in the north. Thus, in terms of the generated wave, Siple appears most nearly conjugate to a station (PB, for example) at L ~ 3.5 in the north even though field line models place its nominal conjugate point (see Figure 3) further poleward. The apparent change in polarization between L ~ 3.5 and L ~ 4.0 suggests that the excited wave resonance could be located in this latitude range (see, e.g., Lanzerotti et al., 1974).

Another example of hydromagnetic waves stimulated by the conductivity change arising from solar flare x-rays was found for the event shown in Figure 6. The x-ray and ionization-related data are presented in the same format as in Figure 1. In this event, which occurred on April 15, 1980, there is some precursor x-ray activity before the main burst at ~ 1509 UT. The main burst lasts ~ 5 minutes and exhibits more structure at the higher energies. As in the first event described, the onset of ionosphere response (bottom two panels) was essentially coincident with the time of peak x-ray flux. The spikey variations on the riometer trace, caused by solar radio noise bursts and propagated interference, are not related to conductivity change in the ionosphere.

Figure 7, similar in format to Figure 2 (except that La Tuque and Siple data were not yet available), shows the magnetic and riometer activity which accompanied this event. A gross variation in the magnetic field amplitude at the time of absorption onset is indicated by the steep fall (rise) in the H-component (D-component) from 1510-1511 UT. The principal oscillation excited at the two low latitude sites appears in both the H-and D-components with a period of ~ 90-100 seconds. However, at the highest latitude (GV), the dominant period in the H-component (~ 40-50 seconds) was approximately half the period of the D-component variation at this latitude and that in all components at the lower latitude stations. The excitation of the higher frequency wave at the higher latitude may indicate a plasma density gradient within the range L ~ 3.5 to L ~ 4.4, as suggested by

contemporary hydromagnetic wave theory (e.g., Southwood, 1974; Chen and Hasegawa, 1974; Lanzerotti et al., 1974). Polarization analysis of bandpass-filtered data, which might confirm this, has not yet been completed for this event.

DISCUSSION

The observations in conjugate areas of oscillations stimulated by solar x ray-induced increases of ionospheric conductivity at one end strongly suggest that the magnetic field variations can be considered hydromagnetic waves excited in the earth's magnetosphere. Experimental and theoretical results have shown that such waves can be interpreted as standing transverse Alfvén waves (i.e., field line resonances) on geomagnetic flux tubes.

The periods of eigenoscillations of geomagnetic field lines have been computed in the dipole approximation by Orr and Matthew [1971] for magnetospheric plasma density conditions typical of the L shell range of interest here. The eigenperiods of the fundamental oscillations are in the Pc4 (45-150 sec) and Pc3 (15-45 sec) bands for field lines located, respectively, inside and outside the plasmapause. The L shell location of the dayside plasmapause, at any given time, is determined, statistically, by the average level of geomagnetic activity during the preceding 9-12 hours.

Recall now the observation in Figure 7 that the frequency of the H-component oscillation at Girardville (at $L \sim 4.4$, the highest latitude station) was about twice that of the H-component at lower latitude ($L \leq 3.5$). The average planetary Kp index for the 12 hours preceding the event of April 15 was about 3+, a moderately disturbed condition, which would place the dayside plasmapause at $L \sim 4$ at the time of the event (1010 LT). Thus, the presence of more than one wave within this L range is consistent with the simultaneous excitation of a Pc3 oscillation outside the plasmapause and a Pc4 oscillation inside. In contrast, the geomagnetic field was relatively quiet (average $Kp \sim 1$) for the event of April 26. For this condition the location of the dayside plasmapause is expected to be poleward of the highest latitude station which is consistent with the excitation of only a Pc4 wave within the L range observed. Note, however, that for both events the wave amplitudes and polarizations observed at ground stations are influenced by the proximity of the station to the resonant field lines (Chen and Hasegawa, 1974) as well as by atmospheric screening effects (Hughes and Southwood, 1976).

The mechanism by which these eigenoscillations of the field line are excited must be related to the sudden, large increase in conductivity in the ionosphere D- and E-regions produced by the solar flare x-ray (and ultraviolet) flux. Detailed model calculations based on realistic configurations of ionospheric electric fields and current flow need to be performed which incorporate rapid, large-scale, height-dependent changes in conductivity. If we assume, for simplicity, that no change occurs in the ambient (horizontal) ionospheric electric fields on the time scale of the impulsive phase of a solar x-ray event ($\lesssim$ 5 minutes), then the conductivity change should lead to an impulsive change in horizontal current flow and thence perhaps to a transverse magnetic perturbation sufficient to excite resonance oscillations. Excitation of hydromagnetic field line resonances might also be expected to accompany a rapid decrease of ionospheric conductivity such as could be produced by appropriate chemical releases in the ionosphere.

CONCLUSION

In summary, the data presented in this paper show that oscillations of the geomagnetic field in the Pc3 and Pc4 bands can be excited by rapid increases of ionospheric conductivity produced by solar x-ray events. The observation at conjugate ends of a field line of oscillations stimulated by a conductivity increase at only one end provides strong support for interpreting the magnetic variations as hydromagnetic field line resonances. The periods of the observed oscillations are consistent with those predicted for the fundamental eigenoscillations of the dipole geomagnetic field in the L range 3.2 to 4.4 for the plasma density conditions expected to prevail. This is the first definite confirmation of a mechanism for the generation of magnetospheric hydromagnetic waves that is not associated with solar wind or magnetic storm effects.

ACKNOWLEDGEMENTS

This material is based upon work supported by the Division of Atmospheric Sciences, National Science Foundation under Grant ATM-8006496 and by the Division of Polar Programs under Grant DPP-7925014.

REFERENCES

Chen, L. and A. Hasegawa, A theory of long-period magnetic pulsations, 1, Steady state excitation of field line resonance, J. Geophys. Res., 79, 1034, 1974.

Hughes, W. J. and D.J. Southwood, The screening of micropulsation signals by the atmosphere and ionosphere, J. Geophys. Res., 81, 3234, 1976.

Lanzerotti, L.J. and D.J. Southwood, Hydromagnetic waves, Chapter III.1.3 in Solar System Plasma Physics, Vol. III, North-Holland Publ. Co., 1979.

Lanzerotti, L.J., H. Fukunishi and L. Chen, ULF pulsation evidence of the plasmapause, 3. Interpretation of polarization and spectral amplitude studies of Pc3 and Pc4 pulsations near L = 4, J. Geophys. Res., 79, 4648, 1974.

Orr, D. and J.A.D. Matthew, The variation of geomagnetic micropulsation periods with latitude and the plasmapause, Planet. Space Sci., 19, 897, 1971.

Rostoker, G., Recent advances in the study of electric and magnetic fields in the ionosphere and magnetosphere, J. Geomag. Geolectr., 32, 431, 1980.

TABLE I: COORDINATES OF STATIONS

Station		Geographic latitude	Geographic longitude	Geomagnetic latitude	Geomagnetic longitude	Approx. L value
Siple	(SI)	76.00°S	84.00°W	64.76°S	8.02°W	4.0
Girardville	(GV)	49.00°N	72.53°W	60.42°N	3.68°W	4.4
La Tuque	(LQ)	47.68°N	72.58°W	58.72°N	2.43°W	4.0
Pittsburg	(PB)	45.03°N	71.44°W	56.85°N	2.30°W	3.5
Durham	(DU)	43.10°N	70.90°W	54.43°N	1.43°W	3.2

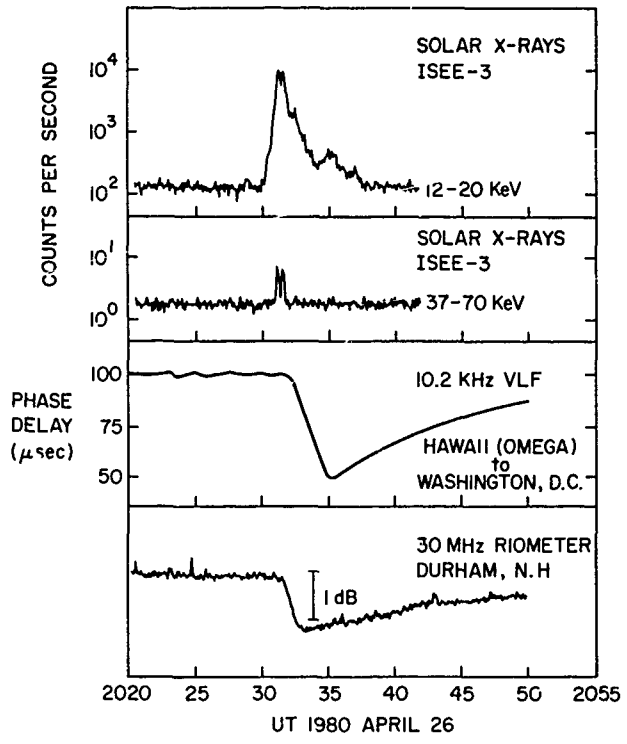


Fig. 1. Solar x-ray data in two energy channels (upper two panels) from the ISEE-3 satellite for the flare event at 2030 UT on April 26, 1980 (courtesy Dr. S. Kane, University of California, Berkeley). The response of portions of the dayside ionosphere is indicated by a sudden phase anomaly on long-path VLF propagation (third panel) and enhanced absorption at 30 MHz (bottom panel).

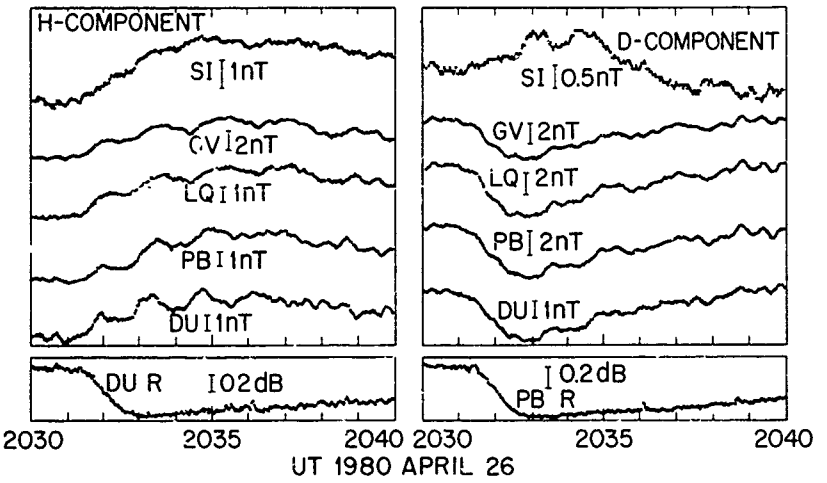


Fig. 2. Surface magnetic field variations in the horizontal plane at 4 locations along a meridional chain in the northern hemisphere (GV, LQ, PB, and DU in order of decreasing latitude) and at a station (SI) in the Antarctic conjugate region during the solar x-ray event of Figure 1 (the stations are identified in Table I and their locations shown in Figure 3). The H-component (D-component) lies along the magnetic north-south (east-west) direction. 30 MHz riometer data are also shown for the two lowest latitude sites.

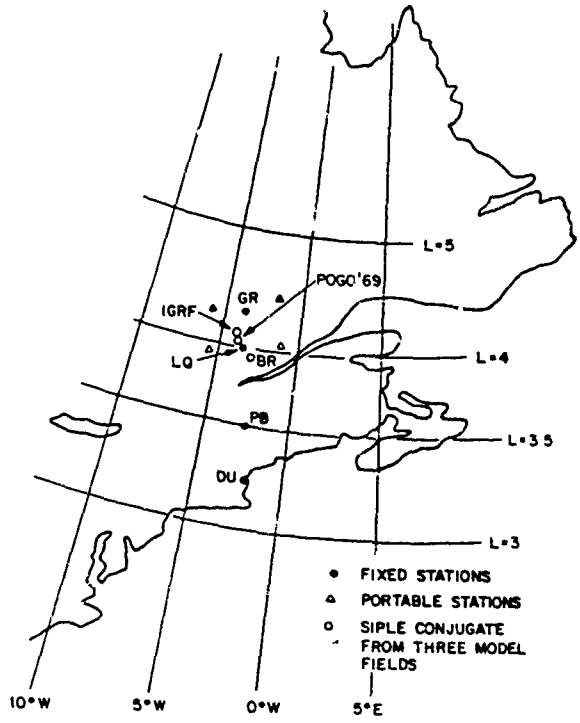


Fig. 3. Locations in a magnetic coordinate frame of the northern hemisphere magnetometer/riometer sites (solid circles) and projections of the Siple, Antarctica conjugate point (open circles) according to three geomagnetic field models. Note that the station labeled GR is the station GV elsewhere in this paper (see Table I for all identifications). The triangles should be ignored.

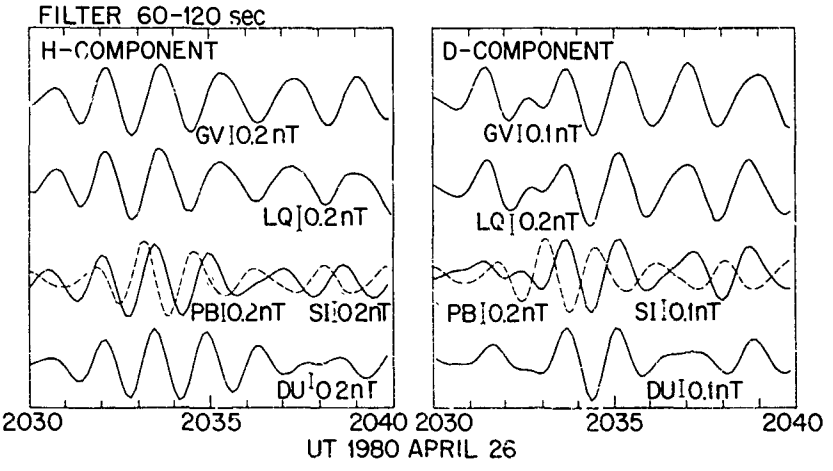


Fig. 4. The magnetic field data of Figure 2 after application of a digital band-pass filter with cutoffs at 60 and 120 seconds. The SI data (dashed curves) are superimposed on the curves for PB which, apart from the phase shift, they most closely resemble.

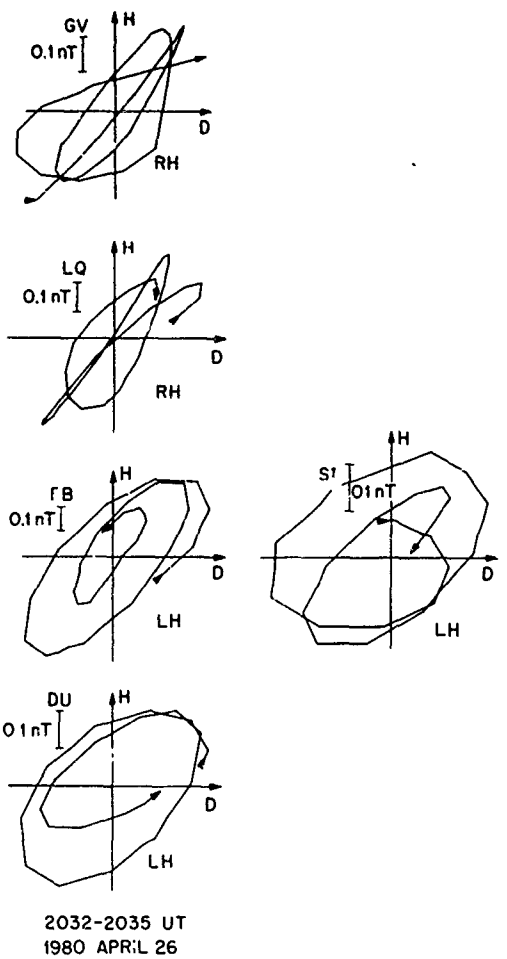


Fig. 5. Hodograms of the magnetic field oscillations of Fig. 4 in the interval 2032-2035 UT showing the sense of polarization (LH or RH) in the horizontal (H-D) plane. LH is lefthand polarization (counter-clockwise looking down along the field line).

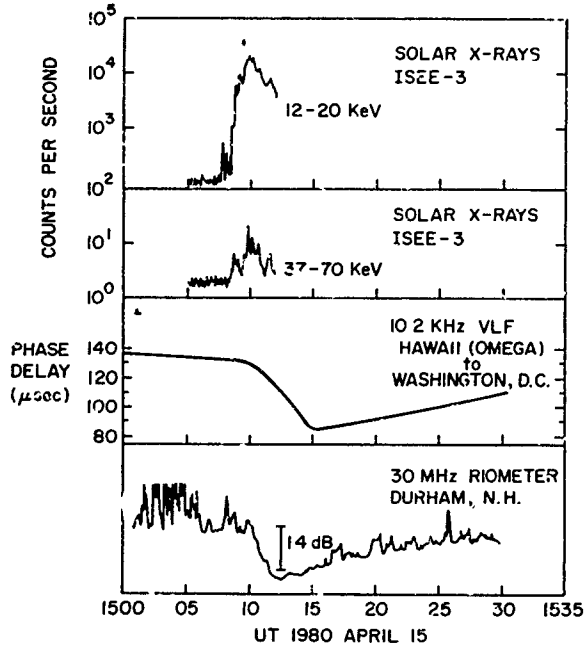


Fig. 6. Solar x-ray and ionization-related data for the flare event at 1510 UT on April 15, 1980, in the same format as Figure 1. The spikey variations in the Durham riometer record (bottom panel) are caused by solar radio noise bursts and propagated interference.

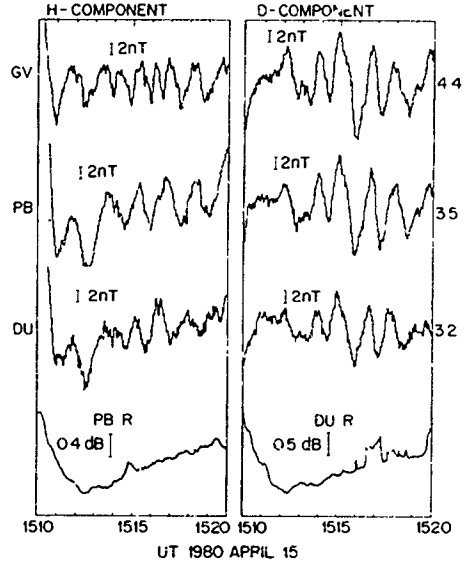


Fig. 7. Magnetic field variations in the horizontal plane and riometer data during the solar x-ray event of Figure 6. Data from LQ and SI were not yet available.

ENHANCEMENT OF RADIO PROPAGATION WITH REACTIVE GAS RELEASES INTO THE IONOSPHERE

Paul A. Bernhardt
University of California, Los Alamos National Laboratory
Los Alamos, New Mexico 87545

ABSTRACT

The propagation of radiowaves can be either adversely or favorably affected by the release of chemically reactive vapors into the F-region ionosphere. Vapors, such as H_2O, H_2, and CO_2, rapidly react with the O^+ ion in the upper atmosphere, producing polyatomic ions such as H_2O^+, H_3O^+, OH^+, and O_2^+. These ions quickly recombine with electrons leaving neutral species. The enhanced F-region neutralization can produce large ionospheric holes which refract radiowaves. Depending on the ray trajectories, the ionospheric hole may produce focusing, fading, or defocusing. Artificially produced plasma gradients may be used to increase the maximum usable frequency (MUF) for systems using sub-ionospheric radio propagation. Besides producing large ionospheric holes, reactive vapor releases can smooth ionospheric irregularities. Irregularities can produce phase and amplitude scintillations in trans-ionospheric signals. Plasma instabilities which are responsible for the growth of F-region irregularities are damped by the enhanced ion-electron recombination caused by the vapor release. For example, both gravitationally driven (Rayleigh-Taylor) instabilities near the geomagnetic equator and current driven convective instabilities in the polar regions can be inhibited by reactive vapor injections into the F-region.

INTRODUCTION

Over the past twenty years, there have been a number of observations of ionospheric depletions produced by chemical releases in the ionosphere. The first observations were of large scale reductions in plasma concentrations coincident with the firings of rockets in the F-region (Booker, 1961; Jackson et al., 1962). Probably the largest depletion was produced during the launch of SKYLAB I by a Saturn V (Mendillo et al., 1975). In 1977, Los Alamos conducted a rocket experiment (code named LAGOPEDO) dedicated to the formation of ionospheric holes (Pongratz et al., 1978). In September 1979, a compaign was conducted to observe ionospheric holes produced by the launch of the HEAO-C satellite. A summary of the results is presented by Mendillo et al. (1980). In the last quarter of 1983, during the Spacelab 2 Mission, the Space Shuttle is scheduled to fire its engines in the ionosphere over several ground observation sites. These engine burns are dedicated ionospheric modification.

One objective of the chemical release experiments is to change the radio propagation characteristics of the ionospheric medium. These changes can be due 1) to large-scale (10 km or greater) plasma gradients produced by the direct effects of the chemical releases or 2) to small-scale irregularities (less than 10 km) produced by instabilities triggered by the chemical releases. The effects of the large-scale gradients have been described by Bernhardt (1981) and are summarized as follows:

VLF (3-30 kHz) waves propagate either via the Earth-ionosphere (EI) waveguide or via field aligned ducts in the earth's magnetosphere. The top edge of the EI waveguide is the D or E region which is relatively insensitive to the effects of chemical release. The F-region, however, is very sensitive to the release of chemicals such as H_2O, H_2, CO_2, NO_2, etc. Modification of the F-region will change the coupling between EI waveguide and ducts in the magnetosphere. Defocusing by the ionospheric hole produced by 100 Kg of H_2 released at 300 km altitude will block VLF transmissions though the ionosphere for a radius of 150 km around the point of release. As the ionosphere recovers from the effects of chemical releases, plasma flows out of the overlying magnetosphere into the ionosphere. This can produce a field aligned column of depleted plasma extending into the magnetosphere capable of guiding VLF (whistler-mode) waves. Thus, the effects of chemical releases initially inhibit the coupling of VLF waves into the magnetopshere, while, at later times, ducts may be formed which can guide these waves to the conjugate hemisphere.

Refraction by the plasma gradients in an ionospheric hole can effect propagation of HF (3-30 mHz) radio signals. Depending on the geometry of the transmitter, receiver, and ionospheric hole,

the HF signals can experience focusing, defocusing, and/or multi-path fading. Under certain conditions, the maximum usable frequency can be increased. The ionospheric hole acts like a convergent lens for waves propagating from satellites to ground-based receivers. The focusing can be as much as 40 dB for HF waves. At VHF (30-300 MHz), the focusing is only a few dB.

The rest of this paper describes previously unpublished research on the stability of ionospheric holes. Instabilities can produce plasma irregularities which affect transionospheric, propagating signals. The effects of chemical releases on both the inhibition and the triggering of these instabilities is considered.

THEORY

The injection of chemicals which react with the O^+ ion in the ionosphere has been used to produce large plasma holes in the medium. The edges of the ionospheric holes are regions of steep plasma gradients. These regions may be unstable to the influence of external energy source such as neutral winds, electric currents and gravity. The chemical reactions which produce the plasma gradients can also have a stabilizing effect on the plasma, inhibiting the growth of plasma irregularities.

In this paper, we present the linear theory of fluid instabilities associated with reactive chemical releases in the ionosphere. The three instabilities investigated are 1) the wind-driven gradient drift instability, 2) the current convective instability, and 3) the gravational Rayleigh-Taylor instability. The linear growth rate for the gradient drift instability will be derived in detail. The growth rates for the other two instabilities will be stated without proof.

A nonlinear set of equations is derived for ionospheric holes produced by CO_2 releases in the ionosphere. The ionospheric reactions involve the ambient N_2 and O_2 molecules as well as the injected CO_2 molecules.

$O^+ + N_2 \rightarrow NO^+ + O$	rate: $\beta_1 = 1.3 \times 10^{-12}$ cm^3/sec
$O^+ + O_2 \rightarrow O_2^+ + O$	rate: $\beta_2 = 2.0 \times 10^{-11}$ cm^3/sec
$O^+ + CO_2 \rightarrow O_2^+ + CO$	rate: $\beta_3 = 1.2 \times 10^{-9}$ cm^3/sec
$NO^+ + e^- \rightarrow N + O$	rate: $\alpha_1 = 4.1 \times 10^{-7}$ cm^3/sec
$O_2^+ + e^- \rightarrow O + O$	rate: $\alpha_2 = 2.0 \times 10^{-7}$ cm^3/sec

The concentration of the molecular ions is denoted by $n_M \equiv [NO^+] + [O_2^+]$. The O^+ concentration is N_I and the electron concentration is $n_e = n_I + n_M$. The velocities along magnetic field lines are determined by plasma diffusion and by collisional coupling of the velocity component of the neutral wind parallel to the magnetic field. These velocities for the molecular ions, O^+ ions and electron are $\bar{v}_{M\parallel}$, $\bar{v}_{I\parallel}$, and $\bar{v}_{e\parallel}$, respectively. The continuity equations for the ions and electrons are

$$\frac{\partial n_e}{\partial t} + \nabla_\perp \cdot (n_e \bar{v}_{e\perp}) + \nabla_\parallel \cdot (n_e \bar{v}_{e\parallel}) = -\alpha n_e n_M \tag{1}$$

$$\frac{\partial n_M}{\partial t} + \nabla_\perp \cdot (n_M \bar{v}_{i\perp}) + \nabla_\parallel \cdot (n_M \bar{v}_{M\parallel}) = -\alpha n_e n_M + \nu_R n_I \tag{2}$$

$$\frac{\partial n_I}{\partial t} + \nabla_\perp \cdot (n_I \bar{v}_{i\perp}) + \nabla_\parallel \cdot (n_I \bar{v}_{I\parallel}) = -\nu_R n_I \tag{3}$$

where $\nabla_\perp$ and $\nabla_\parallel$ are the gradients perpendicular and parallel to the magnetic field, $\alpha = 2 \times 10^{-7}$ cm^3/sec is the ion-electron recombination rate, $\nu_R = \beta_1[n_2] + \beta_2[O_2] + \beta_3[CO_2]$ is the ion-molecule reaction rate, $\bar{v}_{e\perp}$ and $\bar{v}_{i\perp}$ are velocities perpendicular to the magnetic field and t is time. One of the equations is redundant because $n_e = n_M + n_I$. The perpendicular velocities are derived from the equations of motion (see Ossakow et al., 1978a).

$$\bar{v}_{e\perp} = \frac{-\nabla\phi \times \bar{B}}{B^2} \tag{4}$$

$$\frac{\partial \bar{v}_{i\perp}}{\partial t} + \bar{v}_{i\perp} \cdot \nabla \bar{v}_{i\perp} = \frac{e}{r_i}(\bar{v}_{i\perp} \times \bar{B} - \nabla\phi) - \nu_{in} \bar{v}_{i\perp} \tag{5}$$

where ϕ is an electrostatic potential, $\bar{B}$ is the magnetic flux vector and ν_{in} is the ion-neutral collision frequency. The ion masses m_I and m_M are assumed to both equal m_i. The parallel velocities are assumed to be unaffected by the instability.

For the growth of the gradient drift instability, we assume that field aligned currents are zero (i.e., $J_{\parallel} = 0$) and that charge is conserved

$$\nabla_{\perp}\cdot\bar{J}_{\perp} = \nabla_{\perp}\cdot[n_e(\bar{v}_{e\perp} - \bar{v}_{i\perp})] = 0 \tag{6}$$

The ion velocities parallel to the magnetic field are given by the coupled equations (see Banks and Kockarts, 1973)

$$\frac{kT_M}{n_M m_M}\frac{\partial n_m}{\partial z} + \frac{kT_e}{n_e m_M}\frac{\partial n_e}{\partial z} + g_{\parallel} = -(\nu_{in} + \nu_{MI})\, v_{M\parallel} + \nu_{MI} v_{I\parallel} \tag{7}$$

$$\frac{kT_I}{n_I m_I}\frac{\partial n_I}{\partial z} + \frac{kT_e}{n_e m_I} + \frac{\partial n_e}{\partial z} + g_{\parallel} = -(\nu_{in} + \nu_{IM})\, v_{I\parallel} + \nu_{IM} v_{M\parallel} \tag{8}$$

where k is Boltzmann's constant, T_M and T_I are ion temperatures, T_e is the electron temperature, $g_{\parallel}$ is the gravitational acceleration projected along the magnetic field axis(z). The ion-ion collision frequencies are

$$\nu_{IM} = 1.26 \times 10^{-12}\, n_M(m^{-3})\ s^{-1} \qquad \nu_{MI} = 1.26 \times 10^{-12}\, n_I(m^{-3})\ s^{-1}$$

for O_2^+, O^+ momentum transfer collisions. The electron velocity is found from flux conservation (i.e., $n_e v_e = n_M v_M + n_I v_I$).

The coupled set of nonlinear equations for the variables n_e, n_m, $\bar{v}_e$, $\bar{v}_i$ and ϕ are

$$\frac{\partial n_e}{\partial t} + \nabla_{\perp}\cdot\left[-\left(\frac{1}{B^2}\right)\nabla\phi \times \bar{B}\right] + \nabla_{\parallel}\cdot(n_e\bar{v}_{e\parallel}) = -\alpha n_e n_M \tag{9}$$

$$\frac{\partial n_m}{\partial t} + \nabla_{\perp}\cdot[n_M\bar{v}_i] + \nabla_{\parallel}\cdot(n_M\bar{v}_{M\parallel}) = -\alpha n_e n_M + n_R(n_e - n_M) \tag{10}$$

$$\frac{\partial \bar{v}_{i\perp}}{\partial t} + \bar{v}_{i\perp}\cdot\nabla\bar{v}_{i\perp} + \frac{e}{m_i}\,(\bar{v}_{i\perp} \times \bar{B} - \nabla\phi) - \nu_{in}\bar{v}_{i\perp} \tag{11}$$

$$\nabla_{\perp}\cdot[n_e(\bar{v}_{e\perp} - \bar{v}_{i\perp}] = 0 \tag{12}$$

The (x,y,z) coordinate system is chosen such that z is along the magnetic field axis (i.e., $\nabla_{\parallel} = \partial/\partial z\ \bar{a}_z$), and the y-axis is in the direction of the external neutral wind (i.e., $\bar{U}_n = U_n\,\bar{a}_y$). This wind produces ion motion in the x-direction.

The variables are written as perturbations of the ambient quantities.

$$n_e = n_{eo} + n_{e1}e^{i(kx-\omega t)} \qquad \bar{v}_{i\perp} = \bar{v}_{io\perp} + \bar{v}_{i1\perp}e^{i(kx-\omega t)}$$

$$n_M = n_{Mo} + n_{M1}e^{i(kx-\omega t)} \qquad \phi = \phi_o(x) + \phi_1 e^{i(kx-\omega t)}$$

where Fourier analysis in x and t has been applied. The unperturbed ion velocity is $v_{iox} = +U_n(\nu_{in}/\Omega)$ where Ω is the ion gyro frequency.

With these substitutions, the linearized equations become

$$\left(Z + \frac{\Omega^2}{Z}\right) v_{i1x} + \frac{e}{m}\, ik\phi_1 = 0 \tag{13}$$

$$(-n_{eo}ik + \frac{\Omega}{Z}\frac{\partial n_{eo}}{\partial y})\, v_{i1x} + \frac{1}{B}\, ik\, \frac{\partial n_{eo}}{\partial y}\, \phi_1 - v_{iox} ik n_{e1} = 0 \tag{14}$$

$$\frac{ik}{B}\frac{\partial n_{eo}}{\partial y}\phi_1 + (-i\omega + \alpha n_{Mo})\, n_{ei} + \alpha n_{eo} n_{M1} = 0 \tag{15}$$

$$\frac{-\partial n_{Mo}}{\partial y}\frac{\Omega}{Z} v_{i1x} + (\alpha n_{Mo} - \nu_R)\, n_{e1} + (-i\omega + \alpha n_{eo} + \nu_R)\, n_{M1} = 0 \tag{16}$$

where $Z = -i\omega + ik v_{iox} + \nu_{in}$, the effects of the field-aligned velocities for the diffusing electrons and molecular ions are neglected. These ations can be solved yielding a quartic equation for ω. The order of the equation can be reduced by limiting the solutions to cases where $\Omega >> \nu_{in} >> -i\omega + ik v_{iox}$. This is easily justified for ionospheric holes at 300 km altitude where $\nu_{in}/\Omega \approx 0.002$ and the ion inertia can be ignored. These assumptions lead to a quadratic equation which has the solution

$$\gamma = -i\omega = \left[\frac{S}{n_{eo}}\frac{\partial n_{eo}}{\partial y} - \alpha(n_{eo}+n_{Mo}) - \nu_R \pm \sqrt{\left(\frac{S}{n_{eo}}\frac{\partial n_{eo}}{\partial y} - \alpha(n_{eo}+n_{Mo}) + \nu_R\right)^2 + 4S\alpha\frac{\partial(n_{eo}-n_{Mo})}{\partial y}}\right]/2$$

where $S = -U_n$ and γ is the instability growth rate. (17)

The parameter S is used in this equation because it can be replaced by driving terms for other instabilities. A similar treatment for the gravitational Rayleigh-Taylor instability yields an identical expressions for the growth rate with $S = g/\nu_{in}$ where g is the gravitational acceleration in the negative y direction and ν_{in} is the ion-neutral collision frequency. This derivation assumes that B is horizontal in the z-direction. [The derivation is an extension of the one given by Ossakow et al., 1979].

Finally, the current convective instability (see Ossakow and Chaturvedi (1979), Chaturvedi and Ossakow (1979) and Huba and Ossakow (1980)) can be influenced by reactive gas releases. This instability is characterized by streaming electrons with velocity V_d. These electron currents can occur in regions of diffuse aurora. Current convective instabilities grow according to Eq. (17) with

$$S = V_d/2\ \mathrm{SIGN}\left(\frac{\partial n_{eo}}{\partial y}\right)\Big/\left(1 + \frac{\Omega_e \nu_i}{\nu_e \Omega_i}\right)^{1/2} \tag{17a}$$

where SIGN denotes the sign, Ω_e and Ω_i are electron and ion gyrofrequencies, ν_e and ν_i are electron and ion collision frequencies. Nonlinear stabilization (Chaturvedi and Ossakaw, 1979) and magnetic shears (Huba and Ossakow, 1980) have not been included in the linear analysis. All growth rates have been maximized with respect to the wavenumber (k).

Two limits occur in the modified ionosphere. In the center of an ionospheric hole created by a large CO_2 release, most of the O^+ ions have been converted to O_2^+ ions. In this region, $n_M = n_e$ because $\nu_R >> \alpha N_e$ and the instability growth rate simplifies to

$$\gamma = \frac{S}{n_{eo}}\frac{\partial n_{eo}}{\partial y} - 2\alpha n_{eo} \quad \text{or} \quad \gamma = -\nu_R \tag{18}$$

As expected, in regions where the plasma is predominantly composed of polyatomic ions, the damping of instabilities is controlled by recombination of these ions with electrons.

In the unperturbed ionosphere or at the outer edges of the ionospheric hole, the plasma is mainly composed of O^+ ions. Here, $n_M >> n_e$ because $\nu_R << \alpha n_e$ and, consequently, the growth rate reduces to

$$\gamma = \frac{S}{n_{eo}} \frac{\partial n_{eo}}{\partial y} - \nu_R \quad \text{or} \quad \gamma = -\alpha n_e \tag{19}$$

In this outer region, the damping is controlled by ion-molecule reaction rates. The form of Equation (19) has been used by Ossakow et al., (1979) to describe the linear growth of Rayleigh-Taylor instabilities in the natural, unmodified equatorial ionosphere. Only the growth rates given by the general Equation (17) are applicable throughout the modified ionosphere.

The stability of holes created specific releases of CO_2 is calculated. The nonuniform expansion of injected gases is governed by diffusion. An analytic expression for the diffusive flow of vapors from point releases at t = 0 has been derived by Bernhardt [1976].

$$n_1(a,b,c) = \frac{N_o}{(4\pi D_o t)^{3/2}} \exp\left[-c\left(\frac{3}{4H_a} + \frac{1}{2H_i}\right) - \frac{H_a^2[1-\exp[-c/(2H_a)]]^2}{D_o t} - \frac{(a^2+b^2)\exp[-c/(2H_a)]}{4D_o t} - \left(\frac{1}{H_a} - \frac{1}{H_i}\right)^2 \frac{D_o t \exp[c/2H_a)]}{4}\right] \tag{20}$$

where n_1 is the CO_2 concentration, N_o is the number of vapor molecules releases, D_c is the diffusion coefficient at the point of release, H_a is the atmospheric scale height, $H_i = kT_n/m_n g$ is the injected gas scale height, T_n is the neutral temperature, m_n is the CO_2 molecular mass, the coordinates a and b are in the horizontal plane and c is the vertical axis. This expression has been shown to be accurate for describing the first 10 minutes of a release into the ionosphere. All calculations will be for releases at 300 km altitude where

$$[N_2] = 1.3 \times 10^{14}\ m^{-3} \quad [O_2] = 4.7 \times 10^{12}\ m^{-3} \quad [O] = 6.17 \times 10^{14}\ m^{-3}$$

$$D_o = 2.21 \times 10^6\ m^3/sec \qquad H_a = 5 \times 10^4\ m \qquad H_i = 1.91 \times 10^4\ m$$

$$T_n = 1000^o\ K \qquad \nu_i/\Omega_i = 1.9 \times 10^{-3} \qquad \nu_e/\Omega_e = 7.2 \times 10^{-5}$$

The gas injection produces temporal variations in the reaction rate ν_R. The zero order variations in electron concentration (n_{eo}), molecular ion concentration (n_{MO}) and O^+ ion concentration (n_{IO}) can be formed by solving equations (1) through (3) using the field-aligned velocities given in (7) and (8). A number of models have been constructed to perform these calculations at the equator (Anderson and Bernhardt, 1978) and at mid-latitudes (Bernhardt et al., 1975; Bernhardt, 1976; Zinn and Sutherland, 1980, and Bernhardt et al., 1980). For this study, however, we decided to use a simplification suggested by Mendillo and Forbes (1978). The transport terms $\nabla \cdot (nv)$ are ignored and the simplified equations for production and loss are solved. This computation is valid before the diffusive recovery rate, $\nabla_{II} \cdot (n_I \bar{v}_{II})/n_I$ exceeds the reaction time constant ν_R. The diffusion velocities are computed by substituting the plasma concentrations determined from Equations (21) through (23) into Equations (7) and (8). These velocities, coupled with the plasma concentrations, yield the diffusion recovery rate.

The stability of the ionospheric holes is determined by substituting the computed plasma gradients into Equation (17). At the time when the diffusive recovery rate equals the chemical loss rate, the maximum growth rate is examined. If the diffusive recovery rate is much larger than the instability growth rate, the ionospheric hole will "fill-in" before the instability can grow. If, however, the instability growth rate is larger than the diffusive time constant, the hole will be unstable long enough for field aligned irregularties to form.

GRADIENT DRIFT INSTABILITY

The formation of field aligned irregularities in ionized barium clouds has been attributed to the E x B gradient drift instability (Linson and Workman, 1970). One might expect that ionospheric holes are similarly unstable to the gradient drift mechanism. Experimental searches for irregularities in ionospheric holes have been made. In 1977 ionospheric holes were created by releasing CO_2 and H_2O from rockets during Project Lagopedo. Measurements with a 50 MHz radar did not show irregularity formation (Chesnut, private communication, 1978). VHF signals propagating from satellites through the ionospheric holes have not shown evidence of scintillations associated with irregularities. Similarly, measurements of holes produced by rocket exhaust (see Mendillo, 1975 and Mendillo et al., 1980) have not shown any irregularity formation. We seek a theoretical explanation for the lack of irregularities.

The results for the gradient drift instability ($S = -U_n$) are summarized in Table I. The numerical simulations use neutral wind velocities of 50, 100 and 200 m/sec perpendicular to the magnetic field lines in the +y direction. Releases of 20 and 200 kg are considered. The field lines are assumed to be horizontal so that the coordinate z is identical to the coordinate c in Eq. (20).

Table I

Rates for gradient drift instability.

Mass of Co_2 Released (kg)	Neutral Wind Velocity U_n (m/s)	Time for Transition Between Chemical and Diffusive Processes (s)	Diffusive Recovery Rate (s^{-1})	Maximum Instability Growth Rate γ (s^{-1})	Location of Maximum Growth y (km)
20	50	80	2.5×10^{-3}	$-1. \times 10^{-6}$	-10
20	100	80	2.2×10^{-3}	2.9×10^{-3}	-10
20	200	70	2.1×10^{-3}	7.7×10^{-3}	- 9
200	50	250	2.9×10^{-3}	7.2×10^{-4}	-24
200	100	180	3.6×10^{-3}	4.5×10^{-3}	-20
200	200	120	3.8×10^{-3}	1.4×10^{-2}	-16

The effects of the neutral winds are twofold. First, the neutral wind drives the instability according to Eq. (17). Second, the wind convects the neutral gases away from the back side of the ionospheric hole, causing the gradients on the up-wind side of the hole to be steeper than the gradients on the down wind side of the hole. This is fortunate since Eq. (17) implies that positive growth rates can exist only on the up-wind side of the hole.

The amount of gas released affects the time when the transition between chemical production of the hole and the diffusive recovery takes place. This amount also affects the location, relative to the release point, where maximum growth will occur. The maximum growth rate, however, is only slightly affected by the amount released.

Table I demonstrates that 50 m/s winds are not sufficient to stimulate gradient drift instabilities in ionospheric holes produced by CO_2 releases at 300 km altitude. A 100 m/s wind produces instability growth rate on the order of the recovery time due to field-aligned diffusion into the ionospheric hole. Neutral wind velocities of 100 m/s or larger are required to initiate gradient drift instabilities in the hole.

Figure 1 illustrates the result of releasing 200 kg of CO_2 at 300 m altitude into a horizontally stratified ionosphere. The neutral wind velocity is 200 m/s along the ordinate (y-axis) of the figures. The abscissa is time in minutes.

The dashed line in the figures is the transition time between chemical and diffusive processes. The neutral wind blows the CO_2 away from the release point, causing the backside of the hole to be steepened relative to the front side (Fig. 1a). The concentration of the molecular ions is skewed toward the front of the hole (Fig. 1b). Contours of $-\nu_R$ (Fig. 1c) illustrate the movement of the neutral cloud and the dissipation of the enhanced reactions by diffusive expansion of the neutrals. Finally, Fig. 1d illustrates the instability growth rate variation in space and time. In the region $y > -10$ km, the gradient drift instability is strongly dampled by chemical reactions and positive plasma gradients. Two minutes after the release, the region $-45 \text{ km} \leq y \leq 10$ km is unstable to the gradient drift mechanism.

From this study, we conclude the neutral winds larger than 200 m/s are required to trigger the gradient drift instability in ionospheric holes at 300 km altitude. The rarity of such high velocity winds explains the lack of plasma irregularities in the observations.

CURRENT CONVECTIVE INSTABILITY

The National Research Council of Canada and the Space Science Office at Los Alamos have conducted and are currently planning chemical modification experiments in the auroral ionosphere. These experiments, named Waterhole, will use CO_2 and H_2O to reduce the electron concentration in regions where large, field-aligned electric currents are present. While the primary goal of the experiments are to investigate the effects of the plasma reduction on the electron precipitation, the generation of plasma irregularities via the current convective instability is also of interest.

The results for the current convective instability (e.g., Eq. 17a) are summarized in Table II. The numerical simulations use $V_d = 500$ m/s corresponding to an electron current density if 8 μA/m for $n_e = 1 \times 10^{11}$ m^{-3}. At 300 km altitude, the expression $(1+\Omega_e\nu_i/\nu_e\Omega_i)^{1/2}$ has a numerical value of 5.2. The field lines are assumed to be vertical so that the z axis and the c axis are identical.

Table II

Rates for Current Convective Instability

Mass of CO_2 Released (kg)	Electron Drift V_d (m/s)	Time for Transition Between Chemical and Diffusive Processes (s)	Diffusive-Recovery Rate (s^{-1})	Maximum Growth Rate for Instability γ (s^{-1})	Location of Maximum Growth y (km)
20	500	90	2.45×10^{-3}	1×10^{-4}	$\pm$ 10
200	500	300	2.9×10^{-3}	2×10^{-4}	$\mp$ 28

Figure 2 illustrates the effects of a 20 kg CO_2 release. The ordinate of the figure is in the horizontal plane perpendicular to the magnetic field at 300 m altitude. The abscissa is time in minutes. Since no neutral wind is used in this calculation, the plasma concentrations are cylindrically symmetric (Figs. 1a and 2b). The chemical damping due to reaction between CO_2 and O^+ is reduced by diffusion of the CO_2 (Fig. 2c). At the time when diffusive recovery equals $-\nu_R$ (vertical dashed lines in Fig. 2), the instability growth rate has just become positive (Fig. 2d). Because the instability growth rate is much smaller than the rate of diffusive recovery, we conclude that irregularities will not form for either the 20 and 200 kg CO_2 release. Electron velocities substantially larger than 500 m/s are required to stimulate the instability.

GRAVITATIONAL RAYLEIGH-TAYLOR INSTABILITY

There have been two previous studies concerning the effects of reactive-gas releases on the equatorial ionosphere. The first (Anderson and Bernhardt, 1978) is a model simulation of ionospheric hole formation considering chemical processes, field-aligned diffusion and plasma convection across magnetic field lines due to external electric fields. The second study (Ossakow et al., 1978b) did not consider the above processes but did use an ionospheric hole as a seed for self consistent calculations of plasma transport across magnetic field lines due to the Rayleigh-Taylor instability. Complete modeling of equatorial instabilities triggered by reactive chemical releases requires an amalgamation of te above two studies. Both of these studies indicate that gas releases can stimulate the Rayleigh-Taylor instability in the equatorial ionosphere.

In this paper, we investigate the stabilizing effects of CO_2 releases in the bottom side equatorial ionosphere. The y axis [see Eq. (13) through (16)] is vertical. At 300 km altitude, the plasma concentration is 10^{11} electrons/cm^3, the electron gradient ($\partial n_e/\partial y$) is 1.1×10^7 m^{-4}, and the ambient growth rate is 1.6×10^{-3} s^{-1}. The gas releases take place between 200 and 400 km altitude. The grow rate at 300 km altitude is recorded as a function of 1) the altitude of the release, and 2) the time after release.

Figure 3 illustrates the effects of 20 kg CO_2 releases. The vertical dashed line indicates the time of transition between chemical and diffusive processes. The ordinate in the figures give the locations of the CO_2 releases. The isopynic contours of electron and molecular ion concentration are not symmetric about 300 km altitude because of the increase in the neutral diffusion coefficient with altitude (Figs. 3a and 3b). The O^+-molecule reaction rate is largest around release altitudes of 300 km. The reaction rate damping causes the region to be stabilized. For releases at 300 km, the bottom side ionosphere remains stable for over 1.5 minutes.

Figure 4 illustrates the effects of a 200 kg release of CO_2. For releases at 300 km, the O^+ ions are almost completely converted to molecular ions (Figs. 4a and b). Due to enhanced chemical reactions (Fig. 4c), the region at 300 km altitude is stabilized (i.e., $\gamma<0$) for releases between 260 and 400 km altitude. Releases substantially above 300 km stabilize the region because of negative gradients ($\partial n_e/\partial z$) at the bottom of the hole. The stabilization lasts for more than 5 minutes.

While previous works have concluded that reactive gas release can trigger equatorial instabilities, we propose that these releases can produce localized regions of stabilization. The size and duration of the stable regions increase for larger amounts of release material.

ACKNOWLEDGMENT

This work was supported by the Department of Energy and the Defense Nuclear Agency.

REFERENCES

Anderson, D. A. and P. A. Bernhardt, Modeling the Effects of an H_2 Gas Release on the Equatorial Ionosphere, J. Geophys. Res., 83, 4777, 1978.

Bernhardt, P. A., K. M. Price and A.V. da Rosa, Ionospheric Mocification by Rocket Effluerts, Argonne National Laboratory report ANL/EES-TM-100, June 1, 1980.

Bernhardt, P. A., G. G. Park, and P. M. Banks, Depletion of the F2 Region Ionosphere and the Protonsphere by the Release of Molecular Hydrogen, Geophys. Res. Let., 2, 341, 1975.

Bernhardt, P. A., The Response of the Ionosphere to the Injection of Chemically Reactive Vapors, Tech. Rept. No. 17 SU-SEL-75-009, Stanford University, 1976.

Bernhardt, P. A., Changes in the Electromagnetic Properties of the Upper Atmosphere Due to Rocket Effluents, J. Spacecraft and Rockets, in press, 1981.

Booker, H. G., A Local Reduction of F-Region Ionization Due to M ssile Transit, J. Geophys. Res., 66, 1073-1079, 1961.

Chaturvedi, P. K., and S. L. Ossakow, Nonlinear stabilization of the current convective instability in the diffuse aurora, Geophys. Res. Lett., 6, 957, 1979.

Huba, J. D. and S. L. Ossakow, Influence of Magnetic Shear on the Current Convective Instability in the Diffuse Aurora, J. Geophys Res., 85, 6874, 1980.

Jackson, J. E., H. A. Whale, and S. J. Bauer, Local Ionospheric Disturbanc Created by a Burning Rocket, J. Geophys. Res., 67, 2059-2061, 1962.

Linson, L. M., and J. B. Workman, Formation of Striations in Ionospheric Plasma Clouds, J. Geophys. Res., 75, 3211, 1970.

Mendillo, M., Hawkins, G. S. and Klobuchar, J. A., A Sudden Vanishing of the Ionospheric F-Region Due to the Launch of Skylab, J. Geophys. Res., 80, 2217, 1975.

Mendillo, M. and Forbes, J. M., Artificially Created Holes in the Ionosphere, J. Geophys. Res., 83, 151-162, 1978.

Mendillo, M., Rote, D. and Bernhardt, P., Preliminary report on the HEAO-Hole in the Ionosphere, EOS, 61, 529, 1980.

Ossakow, S. L., P. K. Charurvedi, and J. B. Workman, High-Altitude Limit of the Gradient Drift Instability, J. Geophys. Res., 83, 2691, 1978a.

Ossakow, S. L., S. T. Zalesak, and B. E. McDonald, Ionospheric Modification: An Initial Report of Artificially Created Equatorial Spread-F, Geophys. Res. Lett., 8, 591, 1978b.

Ossakow, S. L., and P. K. Chaturvedi, Current Convective Instability in the Diffuse Aurora, Geophys. Res. Lett., 6, 332, 1979.

Ossakow, S. L., S. T. Zalesak, B. E. McDonald, P. K. Chaturvedi, Nonlinear Equatorial Spread F; Dependence on Altitude of the F Peak and Bottomside Background Electron Density Gradient Scale Length, J. Geophys. Res., 84, 17, 1979.

Pongratz, M. B., G. M. Smith, C. D. Sutherland, and J. Zinn, LAGOPEDO - Two F-Region Ionospheric Depletion Experiments, Effects of the Ionosphere on Space and Terrestrial Systems, p. 438, January 24-26, 1978.

Zinn, J. and C. D. Sutherland, Effects of Rocket Exhaust Products in the Thermosphere and Ionosphere, Space Solar Power Review, 1, 109, 1980.

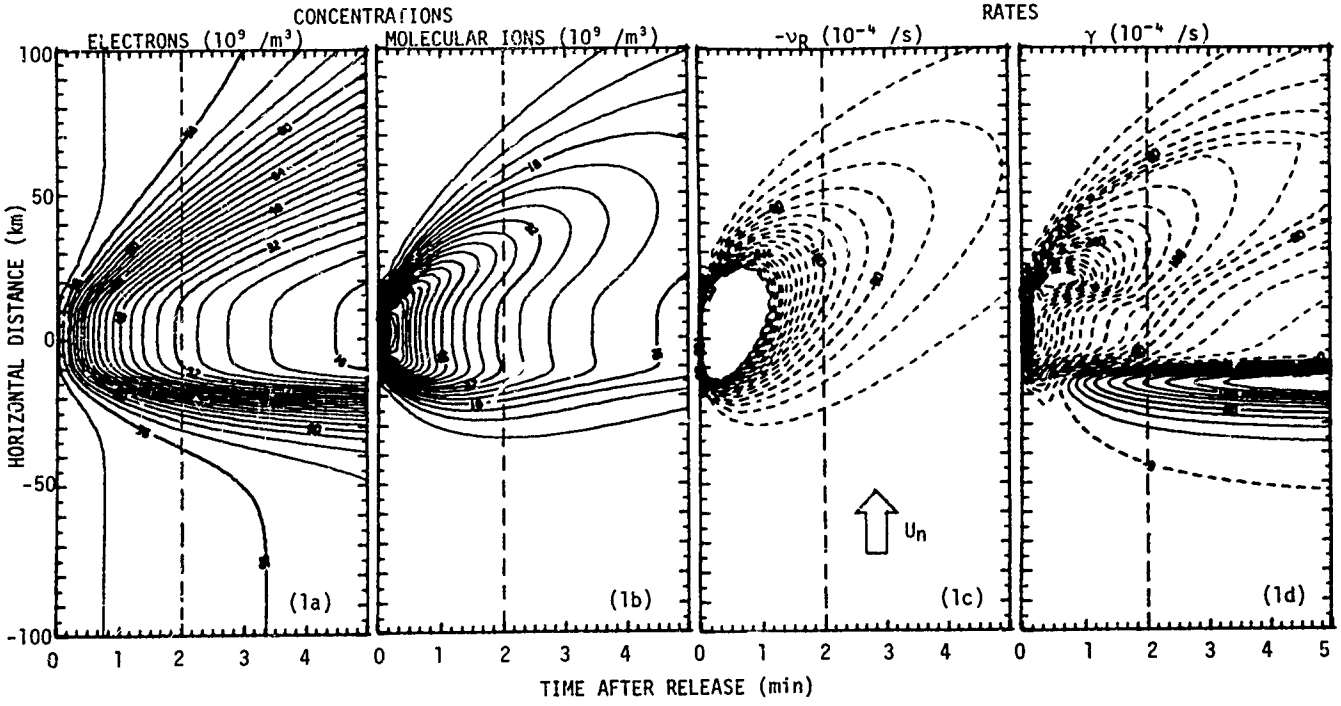


Figure 1. Gradient-drift-instability calculations for an ionspheric hole produced by 200 kg of CO_2. The neutral wind velocity (U_n) is 200 m/s. The dashed line marks the time of transition between chemical and diffusive processes.

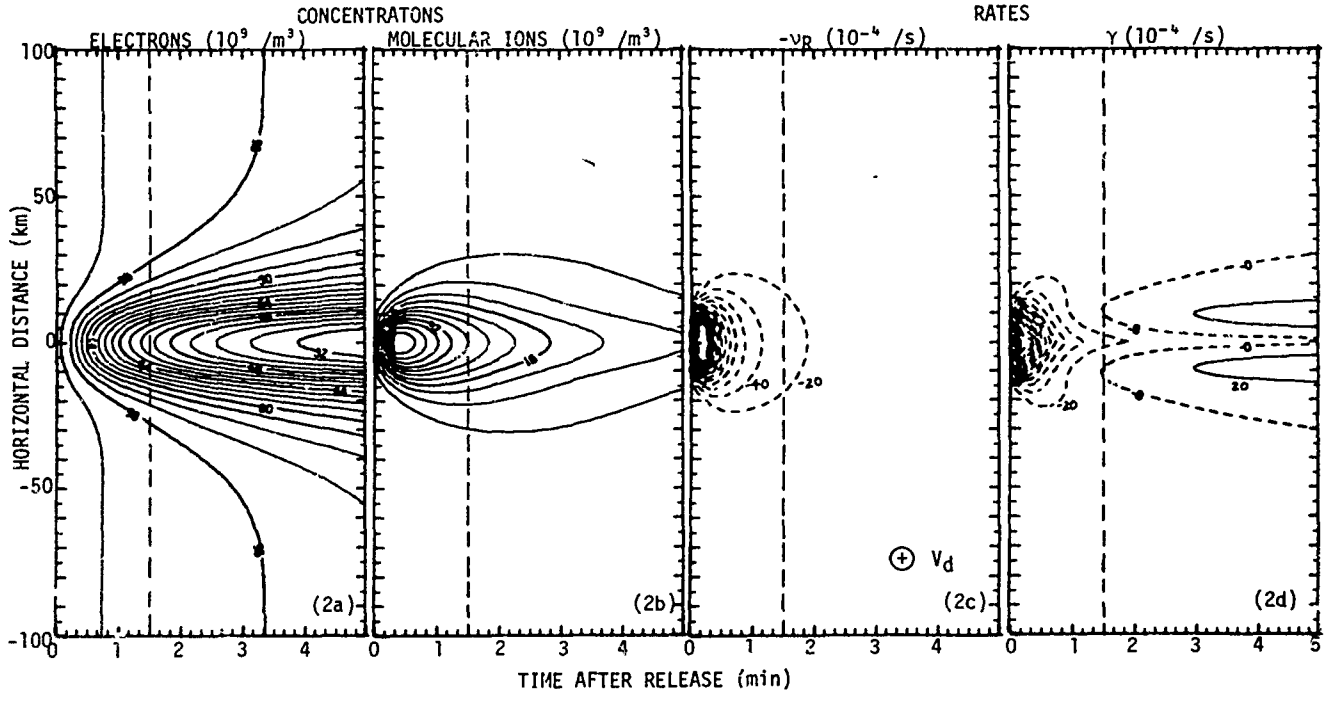


Figure 2. Current-convective-instability calculations for an ionospheric hole produced by 20 kg of CO_2. The electron drift velocity (V_d) is 500 m/s. The dashed line marks the time of transition between chemical and diffusive processes.

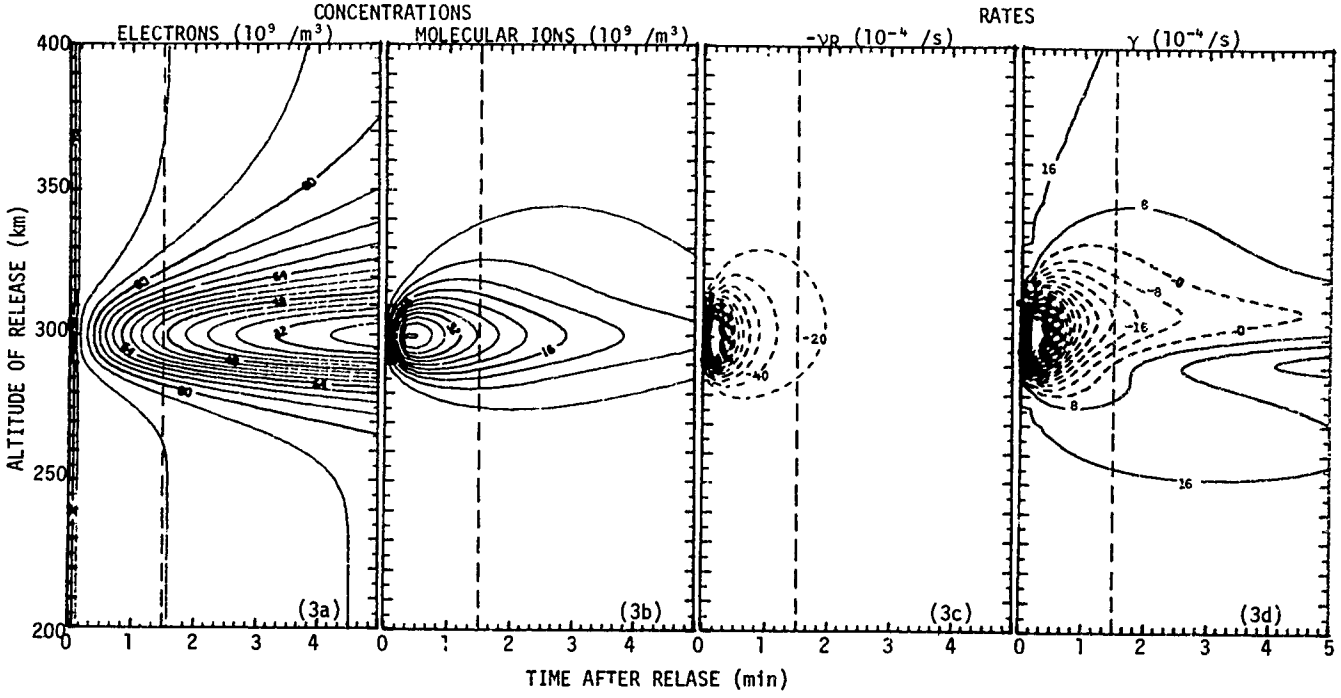


Figure 3. Rayleigh-Taylor-instability calculations at 300 km altitude for ionospheric holes produced with 20 kg of CO_2. The dashed line marks the time of transition between chemical and diffusive processes.

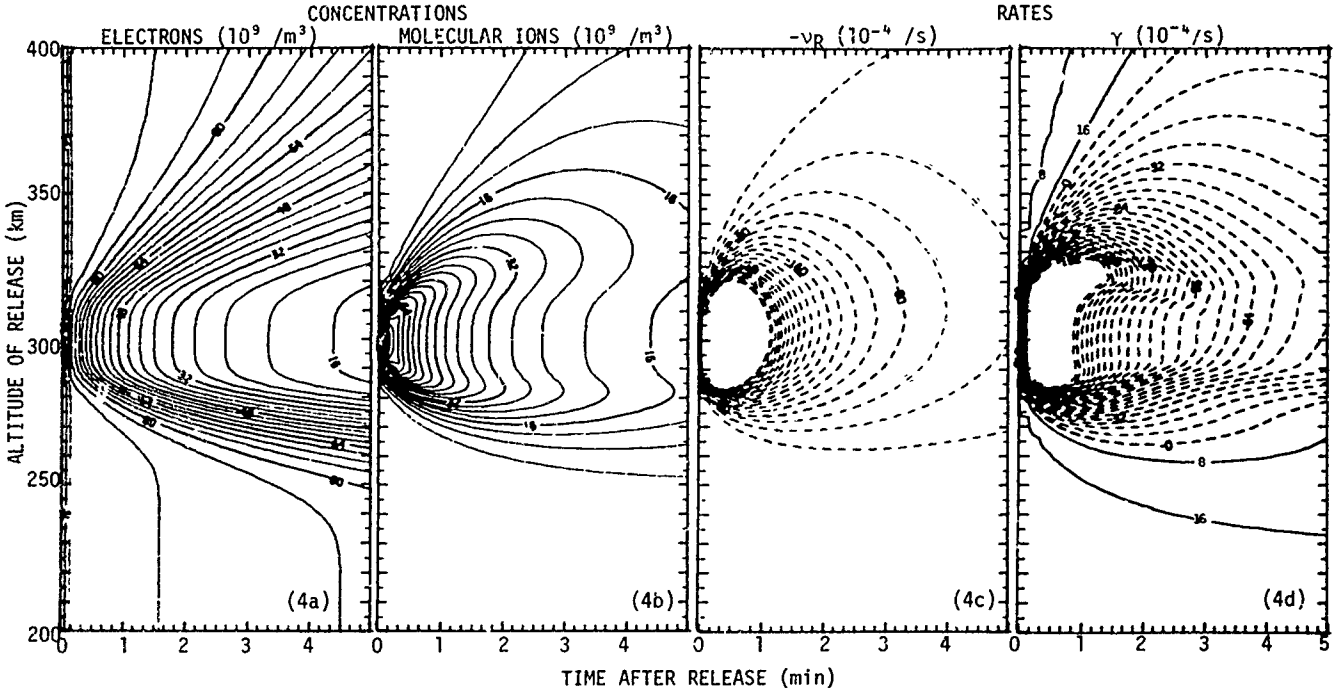


Figure 4. Rayleigh-Taylor-instability calculations at 300 km altitude for ionospheric holes produced with 200 kg of CO_2. The transition between chemical and diffusive processes occurs 5 minutes after the release.

RECENT DEVELOPMENTS IN ARTIFICIAL IONOSPHERIC HEATING

Charles M. Rush
National Telecommunications and Information Administration
Institute for Telecommunication Sciences
Boulder, Colorado 80303

and

Kenneth C. Allen
Boulder Experimental Studies Institute
Boulder, Colorado 80303

ABSTRACT

Intentional modification or heating of the ionosphere by high-powered radio waves continues to be a subject involving much theoretical and experimental interest. The ionospheric changes that result from high-powered HF radio waves permit detailed investigation of the physical processes occurring in a weakly ionized, unbounded plasma that is subjected to intense electromagnetic energy. In recent years, substantial emphasis has been placed upon the study of numerous plasma instabilities that are invoked in order to explain the experimental observations. In addition, there is a renewed interest in determining the impact of ionospheric heating on the performance of telecommunication systems operating in an ionosphere that is subjected to the passage of intense electromagnetic energy. A number of experimental and theoretical studies have been undertaken with the view toward isolating the magnitude of the changes in system performance that can be attributed to ionospheric heating effects. The most recent of these experiments will be reviewed. and the results discussed in terms of current theoretical explanation.

INTRODUCTION

Intentional modification or heating of the ionosphere by high-powered radio waves continues to draw the interest of numerous ionospheric workers. Fejer (1979) and Walker (1979) have recently discussed the latest theoretical work pertaining to the types of interactions between waves in the ionospheric plasma that result in parametric instabilities. Parametric instabilities such as stimulated Brillouin scattering, stimulated Raman scattering, and the parametric decay instability are known to play key roles in the modification of the ionosphere by high-powered HF radio waves that are reflected from the ionosphere. Parametric instabilities are generated as the result of nonlinear interactions between the heater wave and the ionospheric plasma. The nonlinear process responsible for most parametric instabilities in the ionosphere is the ponderomotive force that arises from consideration of nonlinear terms in the equation governing the motion of electrons in a magnetoionic medium.

In addition to instabilities derived from ponderomotive forces, other plasma instabilities associated with ionospheric heating have received attention, particularly the self-focusing instability. This instability results from the heating of the pump wave with a forward-scattered wave. A temperature perturbation is created that leads to a change in the local electron density and index of refraction. The instability produces large-scale plasma striations, and the modified refractive index diverts (or focuses) the pump wave into troughs of modulated density. Duncan and Behnke (1978) have observed self-focusing of radio waves in overdense (heater wave reflected from the ionosphere) heating experiments and found that the kilometer scale sizes associated with the instability are consistent with theory. Perkins and Goldman (1981) have developed equations for the threshold for the onset of the self-focusing instability that are applicable to heating with waves that pass through the ionosphere (underdense heating).

Concurrently with the theoretical and experimental efforts directed toward understanding and isolating plasma instabilities in an artificially modified ionosphere, renewed efforts have been undertaken to determine the impact of ionospheric heating on the performance of telecommunication systems operating in an ionosphere that is subjected to the passage of intense electromagnetic energy. The ionospheric heating facility located at Platteville, Colorado, operating essentially between 5 and 10

MHz, is the focus for such activities in the western hemisphere. The review given by Utlaut (1975) provides an excellent account of the understanding gained in the late 1960's and early 1970's from telecommunications related studies that were undertaken with the operation of the Platteville Facility.

In this paper, the results of studies undertaken in the past two years to determine ionospheric modification impacts on specific telecommunication systems are briefly reviewed. These studies were directed toward (1) assessing the impact of the Satellite Power System (Solar Power Satellite) on ionospheric-dependent telecommunication systems and (2) the extent to which ionospheric modification can be used in conjunction with long-distance HF propagation systems.

IONOSPHERIC MODIFICATION AND THE SATELLITE POWER SYSTEM (SPS)

The Satellite Power System (SPS) is one of a number of alternatives under study by the Department of Energy to meet the energy needs of the Nation in the year 2000. The SPS concept involves one or more satellites in geostationary orbit collecting radiant solar energy and beaming the converted power at 2.45 GHz to the earth's surface. Between 5 and 10 GW of power would be beamed through the ionosphere continuously to a receiving location about 10 km in diameter. Because of the large amounts of power involved, concern was raised as to whether the ionosphere would be modified by the passage of the SPS power beam. More to the point, was the concern about what impact such modification would have upon the performance of ionospheric-dependent telecommunication systems.

At the 2.45-GHz frequency, the heating that the SPS power beam would provide to the lower ionosphere is believed to be that arising from ohmic interactions between the power beam and the electrons, ions, and neutral particles that make up the ambient ionosphere. Under conditions of ohmic heating, the resulting power flux at a microwave frequency can be related to the power flux at another frequency through the relationship (Gordon and Duncan (1978)):

$$\frac{P_{SPS}}{f_{SPS}^2} = \frac{P_{HF}}{f_{HF}^2} \tag{1}$$

where f_{SPS} and P_{SPS} are the SPS microwave frequency and power density and f_{HF} and P_{HF} are another frequency and the power density. It follows from Eq. 1 that heating the ionosphere using radio waves at a lower frequency than that of the SPS requires a smaller power density to achieve an SPS-comparable effect. Provided the frequency is higher than the plasma frequency, the heating is accomplished by radio waves that pass through the ionosphere (the underdense case), and high-powered, high-frequency waves can be used to simulate SPS heating.

At F region heights, self-focusing effects are believed to be particularly important. The threshold for the onset of thermal self-focusing is believed to be proportional to the cube of the wave frequency (Perkins and Goldman (1981)). Thus, the rate at which energy is imparted into the self-focusing instability can be expressed as:

$$\frac{P_{SPS}}{f_{SPS}^3} = \frac{P_{HF}}{f_{HF}^3} \tag{2}$$

Equations 1 and 2 indicate that the amount of SPS microwave energy that would heat the ionosphere and generate thermal self-focusing instabilities can be realistically simulated using much lower frequencies and power densities, provided that the lower frequencies are higher than the plasma frequencies.

The validity of Eqs. 1 and 2 is crucial to the ground-based simulations of the SPS operation. The results obtained by heating the ionosphere with high-frequency waves must be extrapolated to a frequency of nearly 1000 times greater to arrive at the SPS operational frequency. It is possible that some instabilities in the ionosphere that would result from the passage of the SPS power beam cannot be simulated using the lower-frequency, ground-based, ionospheric-heating facilities. However, the current understanding of the processes that are anticipated to occur in the SPS environment indicates that the ohmic heating ($1/f^2$) and thermal self-focusing instability ($1/f^3$) scaling laws are reasonable assumptions.

Experiments were performed using the Platteville Facility to determine the degree to which ionospheric changes induced by ohmic heating due to SPS operation would affect telecommunications systems. The Platteville Facility can provide SPS-comparable ohmic heating to the lower ionosphere by exploiting the $1/f^2$ scaling law. Telecommunications systems whose signals are reflected and controlled by the lower ionosphere were investigated. The current Platteville ionospheric heating facility is essentially the same as that described in Carroll et al. (1974). It provides SPS-comparable power

density at 5 MHz to an area of the lower ionosphere that is 30 km in diameter at an altitude of 75 km and 40 km in diameter at 100 km. This area is three to four times larger than that anticipated for the SPS microwave beam as it passes through the ionosphere. The telecommunications systems chosen for investigation were representative of those operating in the very-low-frequency (VLF, 3 kHz-30 kHz), low-frequency (LF, 30 kHz-300 kHz), and medium-frequency (MF, 300 kHz-3 MHz) portions of the electromagnetic spectrum.

The relationship of the propagation paths of the telecommunication systems to the area of the ionosphere modified by the Platteville Facility is shown in Figure 1. During the course of the experiment, over 40 hours of recording of VLF, LF, and MF signals was made while the facility was heating the ionosphere. Figure 2 shows the amplitude of the OMEGA signal transmitted from Hawaii and received at Brush, Colorado, during times when the Platteville Facility was operating (hatched area on abscissa) and when it was not. The results of these and other experiments indicate that no adverse impact on the performance of VLF, LF, and MF systems would ensue from the operation of an SPS with a maximum power density of 23 mW/cm^2 --the current system design. Details of these experiments and the results derived from them have been given by Rush et al. (1981).

Experiments to determine whether thermal self-focusing effects could be produced in underdense plasma were also undertaken using the Platteville Facility. Thermal self-focusing can create striations or irregularities in the ionospheric electron density. These irregularities, if they occur, could scatter radio waves in the high-frequency (HF), very-high-frequency (VHF), and ultra-high-frequency (UHF) portion of the spectrum. The performance of telecommunications systems operating in these bands could be degraded due to interference from scattered signals. The scattered signal could travel over great distances and interfere with systems operating far from the SPS power beam. It was pointed out earlier that the threshold for the onset of the instability varied as $1/f^3$. The Platteville Facility can provide more than five times SPS-equivalent power density to the ionosphere at 300 km for a 10-MHz scaling to the SPS operational scenario ($1/f^3$ scaling). At 5 MHz, the SPS-equivalent power density is more than 50 times the SPS operational power density of 23 mW/cm^2. For this reason, measurements were made to determine if self focusing instabilities could be generated using underdense radio waves and what effect these instabilities would have on specific telecommunication systems.

The effects of thermal self-focusing are anticipated to be most pronounced in the F-region. The instabilities could lead to striations in the electron density resulting from electrons aligning along the geomagnetic field lines. The experimental arrangement, therefore, emphasized use of telecommunications systems operating at frequencies that are sensitive to the electron distribution in the F-region. The systems utilized were satellite-to-ground and satellite-to-aircraft transmissions operating in the VHF (30-300 MHz) portion of the spectrum. Such transmissions are rather sensitive to irregularities in the electron density structure along the satellite-to-observer radio path.

Basu et al. (1980) have described the results of these experiments that were conducted by measuring LES-8 satellite signals on an aircraft configured to move with respect to Platteville as shown in Figure 3 and also on the ground. Figure 4 shows that the LES-8 signal displays scintillation as it passes through the ionoshere heated by Platteville in the underdense mode and received in an aircraft. The results of these experiments clearly show effects of ionospheric irregularities on the LES-8 satellite signals. These preliminary results, however, cannot be interpreted at this time for the SPS reference system. The $1/f^3$ scaling law is yet to be validated, but if correct, indicates that about 90 mW/cm^2 of SPS comparable power density were delivered to the F-region for these experiments. Also, the experiments maximized the effect of the irregularities because the satellite signal was observed along the earth's magnetic field lines.

In addition to the telecommunication studies described above, studies were undertaken to determine, in a more quantitative manner, the degree to which the ionosphere above Platteville was modified during the heating experiments. Figure 5 shows the time history of the amplitude and phase of a 3.4 MHz ionosonde signal during the time the Platteville Facility was turned on. It can be seen that obvious changes in the amplitude and phase of the 3.4 MHz signal are associated with heater turn-on. Recently, Meltz et al. (1981), using a D region ionospheric model based on the work of Tomko et al. (1980), have shown that the changes seen in Figure 5 are consistent with our current theoretical understanding of the response of the D and E region to high-powered HF heating.

IONOSPHERIC MODIFICATION AND LONG DISTANCE HF PROPAGATION

Propagation of HF radio waves over long distances is of concern in a number of applications. Since the late 1920's, it has been appreciated that HF radio waves can travel to distances beyond that associated with conventional ionospheric reflections without returning to the ground. It is of interest to determine whether modifications to the ionosphere induced by high-powered HF heating could launch long-distance, propagated, or ducted waves or could scatter energy trapped in ionospheric ducts. Gurevich and Tsedilina (1976) have calculated the regions on the ground where transmitters and receivers should be located in order to use artificially induced irregularities for coupling into ionospheric ducts.

In the spring and summer of 1979, an experiment was undertaken to determine if irregularities produced by the Platteville Facility could scatter radio energy that was trapped in ionospheric ducts. A transmitter was located in Australia transmitting signals toward the Platteville Facility. A receiver was located at Alamosa, Colorado, about 200 miles south of Platteville. Figure 6 shows a long-distance, oblique ionogram recorded at Alamosa on June 5, 1979. The various modes can be discerned as darker images among the large amounts of clutter. During this time, the Platteville Facility was not operating. Figure 7 shows an ionogram obtained during a time on June 5, 1979, when the Platteville Facility was operating. In addition to the modes seen on Figure 6, a clear indication of signals is seen to the right of the conventional modes. The source of these signals is attributed to energy scattered by the irregularities created by the Platteville Facility.

In order to interpret the results shown in Figure 6 and 7, it is necessary to understand the geometry of the experiment Figure 8 shows the relationship of Alamosa to the Platteville Facility. Radio waves transmitted from Australia can be received at Alamosa after 3, 4, or 5 reflections from the ionosphere. These signals are observed at all times. Signals that display a scatter-type behavior at Alamosa were observed only when Platteville was operating. In order that signals from Australia be scattered to Alamosa by the iregularities above Platteville, the Australian signals had to have been propagating at the height of the irregularities above Platteville. The times for which scattering by Platteville irregularities showed the largest effects corresponded to local dusk at the Australian transmitter. It is well-known (Toman, 1979, for example) that the gradients in the ionosphere around sunset are situated in such a fashion that propagation of radio waves into the nightside favors ducting or trapping of HF signals.

The results of the long distance propagation experiments are just now becoming available. Further work is required in order to demonstrate the applicability of using artificially-induced ionospheric irregularities to scatter HF ducted signals to the surface of the earth and to ascertain the scatter cross-section of the irregularities as a function of frequency.

CONCLUSIONS

Modification of the ionosphere by high-powered HF waves continues to draw the interest of numerous ionospheric workers. In addition to studies directed at understanding and observing the effects of the interaction between the heater wave and the ionosphere, studies directed toward assessing the impact of ionospheric modification on telecommunication systems performance are being actively pursued. These studies will continue to provide information about how telecommunication systems respond to changes in ionospheric electron density and temperature that result from intentional modification of the ionosphere.

REFERENCES

Basu, S., S. Basu, A. L. Johnson, J. A. Klobuchar, and C. M. Rush, Preliminary results of scintillation measurements associated with ionosphere heating and possible implications for the solar power satellite, Geophys. Res. Letters, 7, 609-612, 1980.

Carroll, J. C., E. J. Violette, and W F. Utlaut, The Platteville High Power Facility, Radio Sciences, 9, 889-894, 1974.

Duncan, L. M. and R. A. Behnke, Observations of self-focussing electromagnetic waves in the ionosphere, Phys. Rev. Letters, 41, 998-10001, 1978.

Fejer, J. A., Ionospheric modification and parametric instabilities, Rev. Geophys. Space Phys., 17, 135-153, 1979.

Gordon, W. E. and L. M. Duncan, Ionosphere/Microwave Beam Interaction Study, William Marsh Rice University Report DRL T01349, Houston, TX.

Gurevich, A. V. and Ye. Ye. Tsedilina, Trapping of radiation in the ionospheric duct during scattering on artificial inhomogeneities, Geomag. Aeron., 15, 713-715, 1976.

Meltz. G., C. M. Rush, and E. J. Violette, Simulation of D and E region nigh-power microwave heating with HF ionospheric modification experiments, National Radio Science Meeting, 12-16 January 1981 (Abstract).

Perkins, F. W. and M. V. Goldman, Self-focussing of r.c waves in an underdense plasma, J. Geophys. Res., 861, 600-608, 1981.

Rush, C. M., E. J. Violette, R. H. Espeland, J. C. Carroll, and K. C. Allen, Performance of VLF, LF, and MF telecommunication systems in a simulated satellite power system environment ascertained by experimental means, Radio Science, 16, 219-234, 1981.

Toman, K., High-frequency ionospheric ducting--A review, Radio Science, 14, 447-453, 1979.

Tomko, A. A., A. J. Ferraro, H. S. Lee, and A. P. Mitra, A theoretical model of D-region ion chemistry modifications during high power radio wave heating, J. Atmos. Terrest. Phys., 42, 275-285, 1980.

Utlaut, W. F., Ionospheric modification induced by high-powered HF transmitters: A potential for extended range VHF-UHF communications and plasma physics research, Proc. of IEEE, 63, 1022-1043, 1975.

Walker, J. C. G., Active experimentation with the ionospheric plasma, Rev. Geophys. Space Phys., 17, 534-544, 1979.

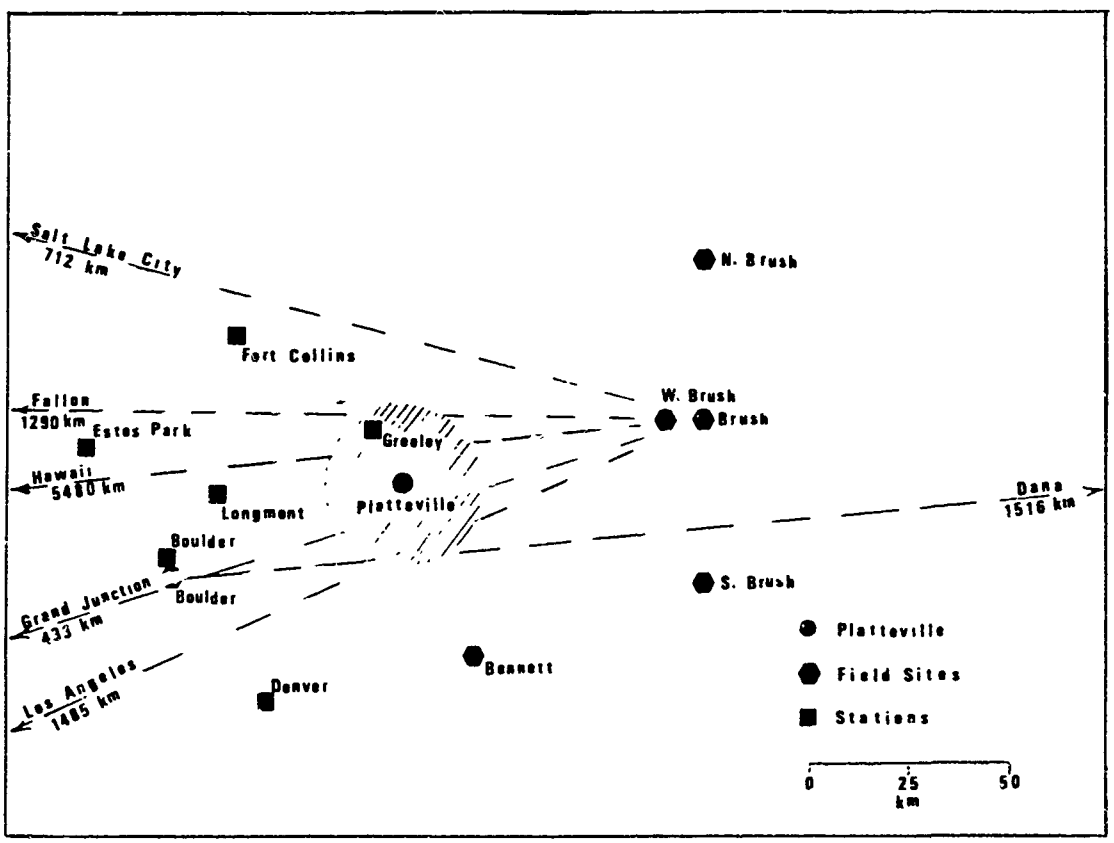


Figure 1. Relationship of propagation paths to the region of the ionosphere at 100 km heated by the Platteville Facility.

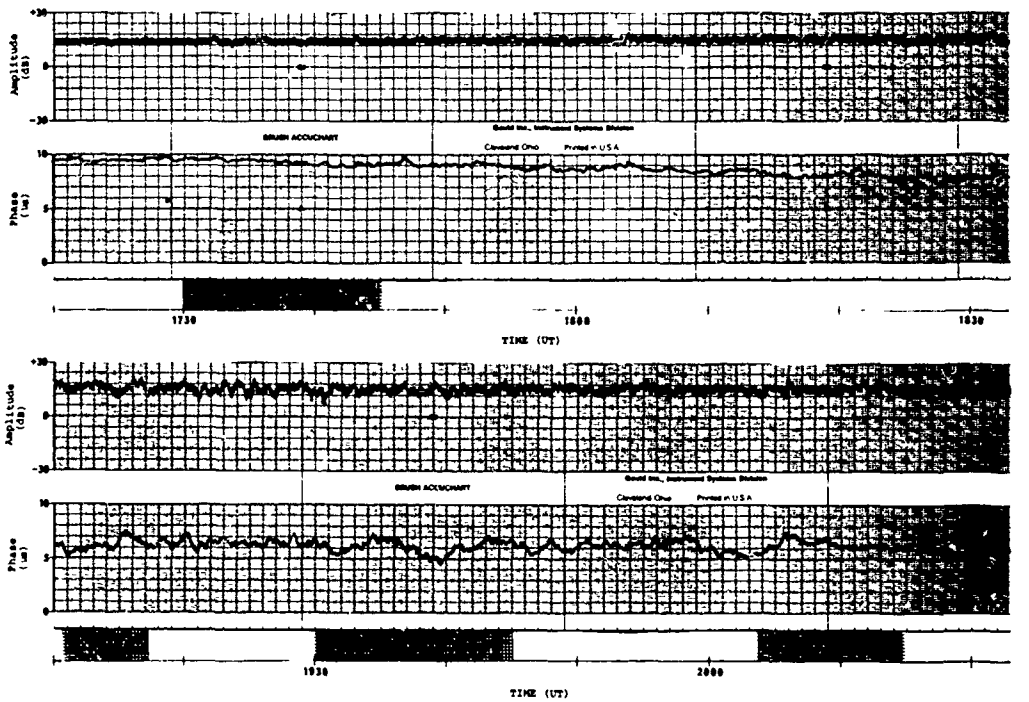


Figure 2. Amplitude and phase of the OMEGA signal transmitted from Hawaii to Brush, Colorado, during times when the Platteville Facility was operating (hatched area) and when it was not operating on September 26, 1979.

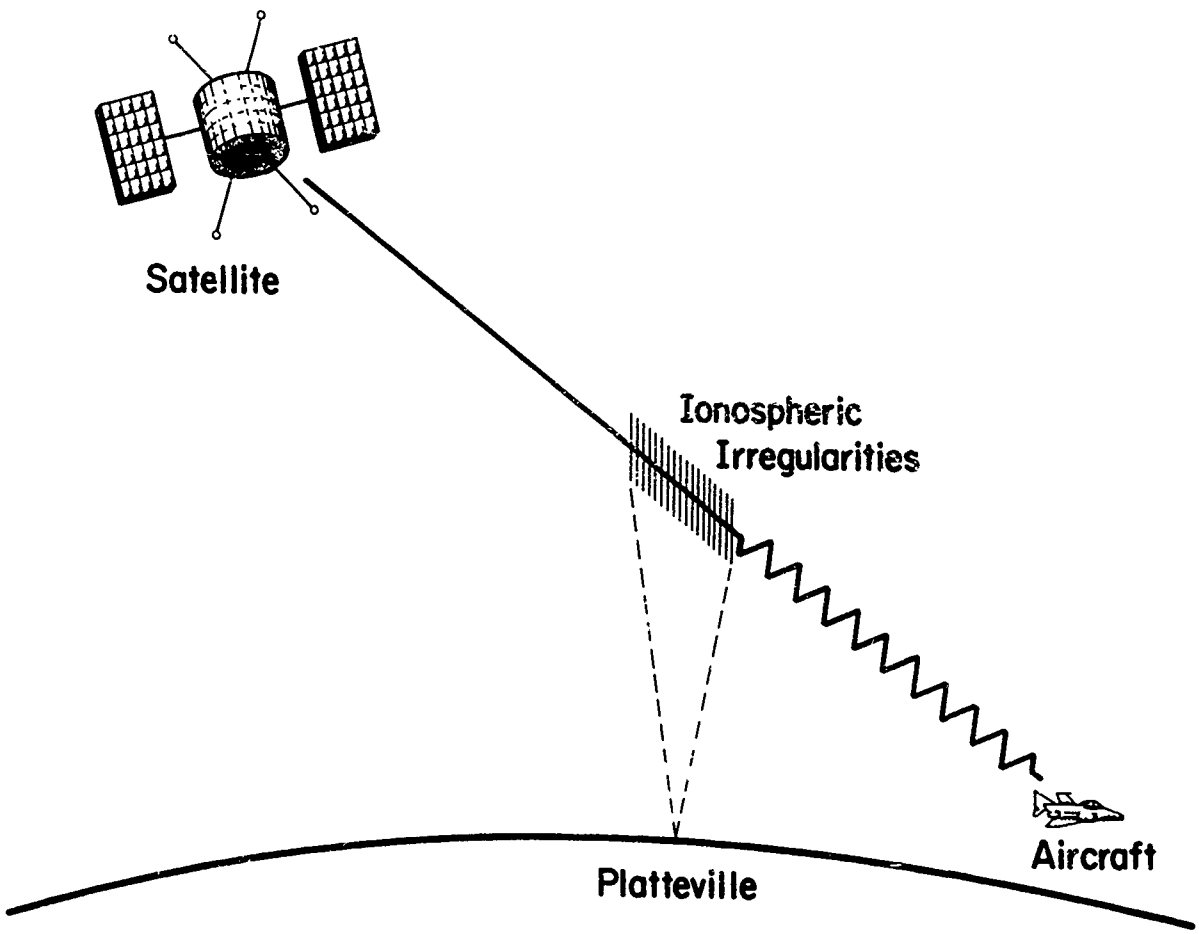


Figure 3. Geometry of a radio propagation path from LES-8 satellite to an Air Force Avionics Laboratory (AFAL) aircraft.

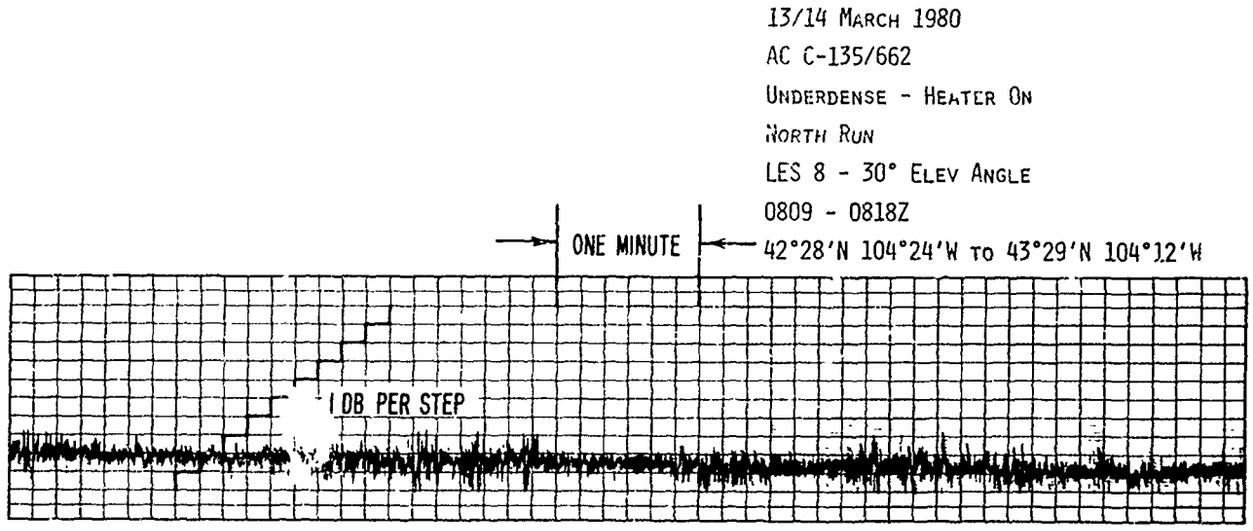


Figure 4. LES-8 satellite signal received on AFAL aircraft during a period of underdense heating.

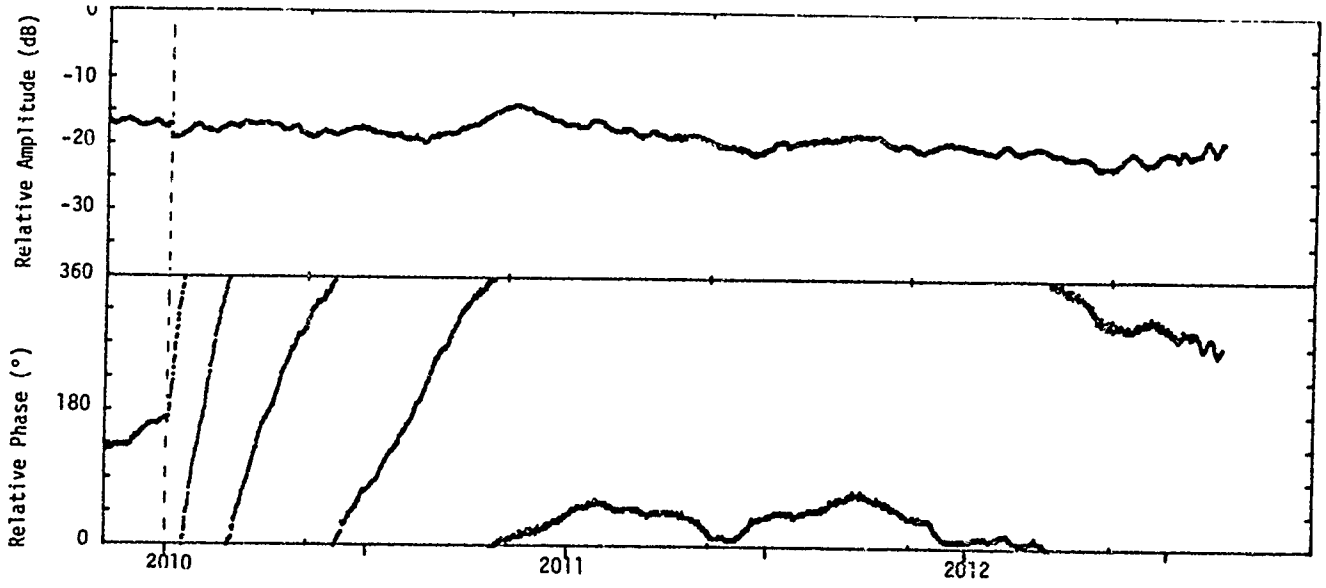


Figure 5. Amplitude and phase of a 3.4 MHz ionosonde signal recorded near Platteville, Colorado, on March 6, 1980.

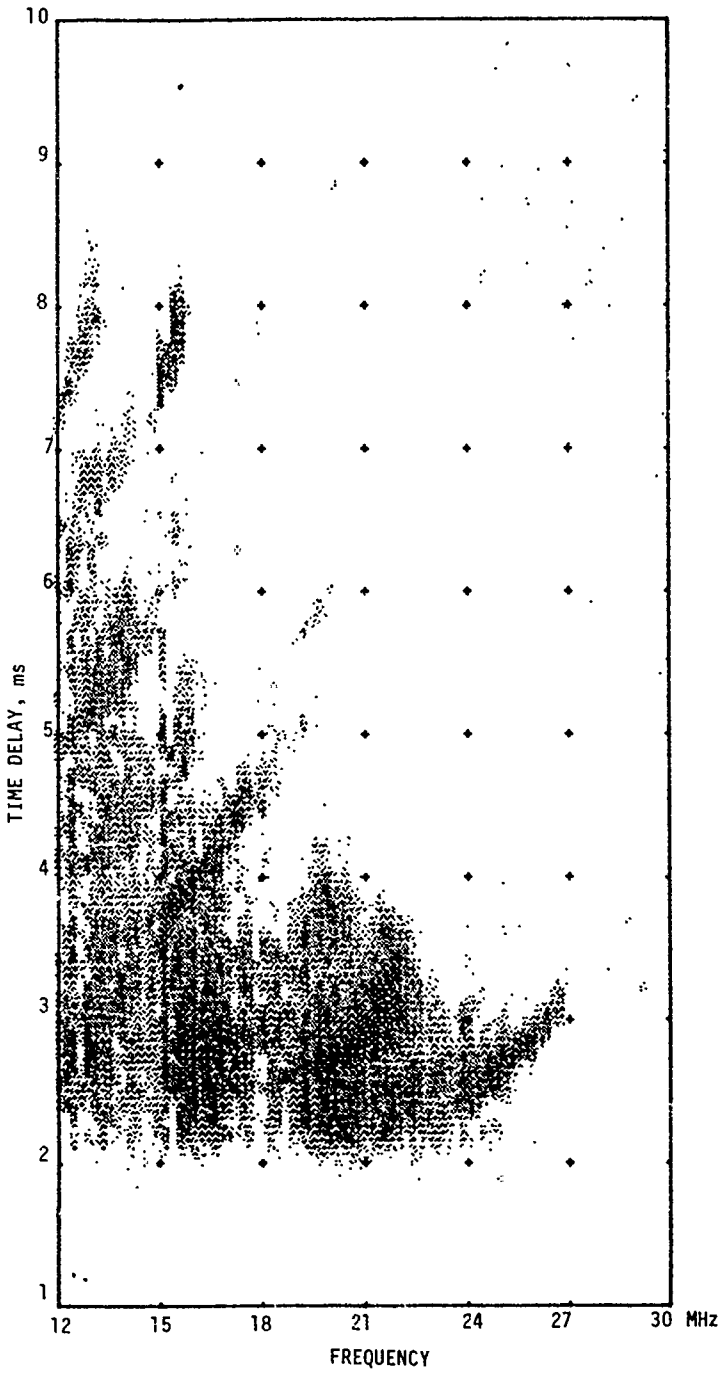


Figure 6. Australia to Alamosa, Colorado, oblique ionogram recorded on June 5, 1979, at 0643 UT, during a time when the Platteville Facility was OFF.

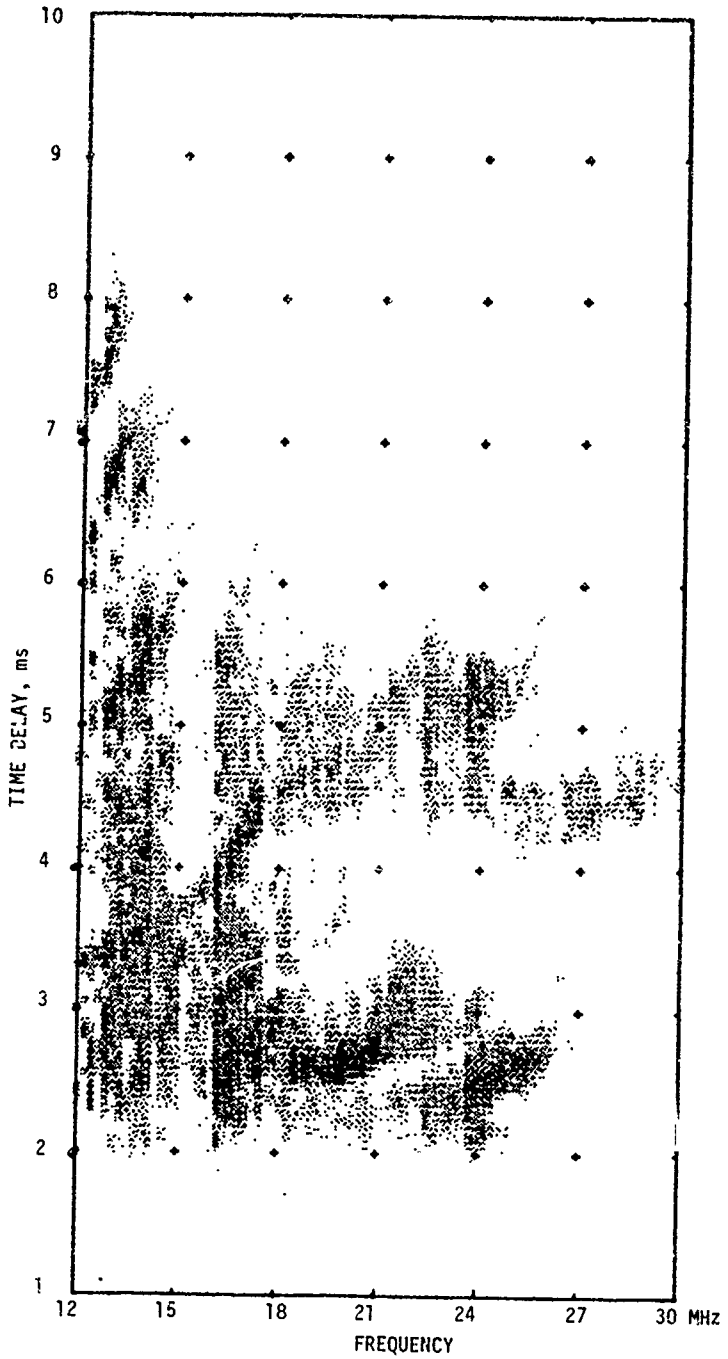


Figure 7. Australia to Alamosa, Colorado, oblique ionogram recorded on June 5, 1979, at 0650 UT during a time when the Platteville Facility was ON.

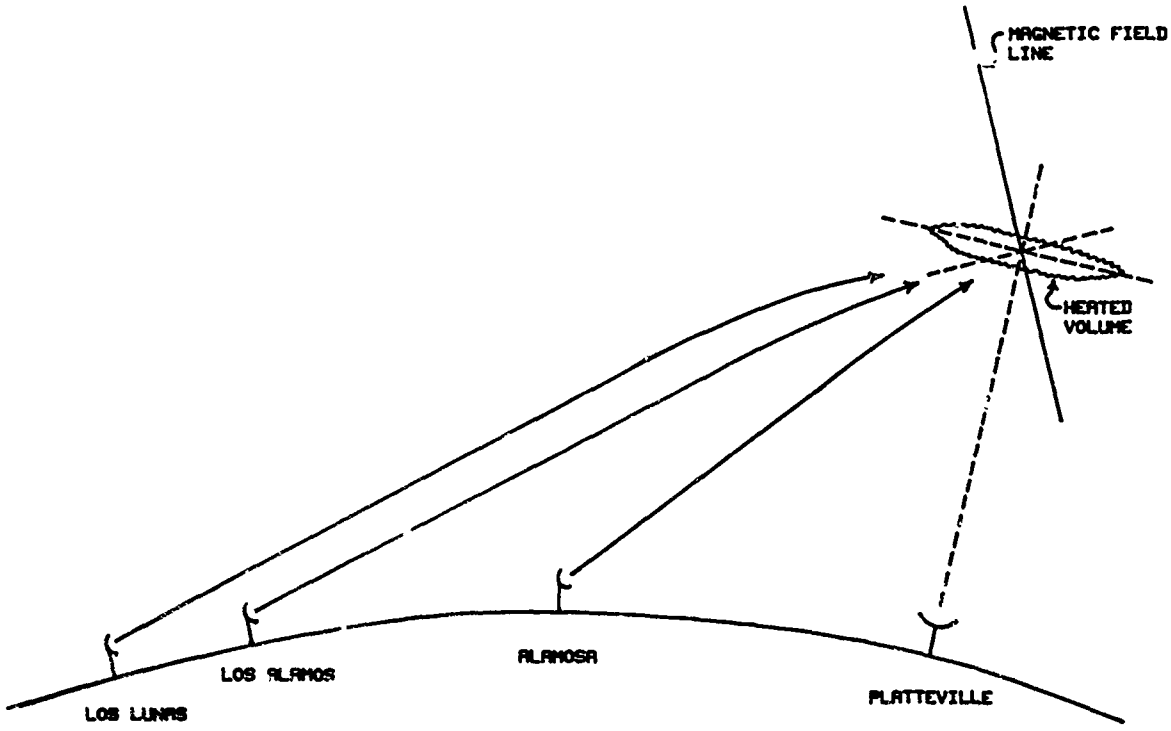


Figure 8. Location of field sites relative to the ionosphere modified by the Platteville Facility.

IONOSPHERIC MODIFICATION BY HIGH-POWER RADIO WAVES

Lewis M. Duncan
University of California, Los Alamos National Laboratory
Los Alamos, New Mexico 87545

ABSTRACT

Powerful, high-frequency radio waves have been used to temporarily modify the ionosphere. Thermal and parametric interactions have led to a diverse range of observed phenomena, including generation of density striations and artificial spread-F, enhancements of electron plasma waves, production of extrathermal electron fluxes and enhanced airglow, modification of the D-region temperature and densities, wideband signal attenuation, and self-focusing and scattering of the electromagnetic waves. The physics of ionospheric modification by high-power radio waves is reviewed in the context of our current theoretical understanding; disturbance gereration mechanisms are qualitatively described. In addition, results of recent experiments are summarized in which ionospheric irregularities are generated and their evolution and decay processes investigated in detail. The effects and potential controlled applications of these HF ionospheric modifications for various RF systems studies are discussed. The C^3I scientific community provides an important motivation for these ionospheric modification studies; their increased interaction and active participation in experimental design and interpretation is encouraged.

INTRODUCTION

The ionosphere is commonly defined as that part of the earth's upper atmosphere where free electrons exist in sufficient numbers so as to affect radio wave propagation. Numerous telecommunications systems rely on this partially ionized plasma through ionospheric reflections or transionospheric propagation as part of their communications signal path. The majority of studies of the effects of the ionosphere on communications systems address specifically the influence of ionospheric disturbances on radio wave propagation. However, the propagation of electromagnetic radiation through a plasma is inherently a nonlinear process. Not only does the plasma through its index of refraction and collisional damping affect the propagating radio wave, but also the radio wave through ohmic heating and the ponderomotive force may influence the behavior of the plasma. This study discusses ionospheric modifications and accompanying RF systems effects which may be generated by high-power radio wave propagation through the upper atmosphere.

Within the past decade it has become technologically feasible to construct ground-based high-frequency (HF) transmitting systems capable of delivering RF energy to the ionospheric plasma with power densities sufficient to alter the ionospheric electron thermal budget and local plasma characteristics, driving a wide variety of plasma instabilities and nonlinear wave propagation effects. Some of the primary and secondary manifestations of such energy deposition in the ionosphere and upper atmosphere are noted schematically in Figure 1 (from Carlson and Duncan, 1977). Results of experiments performed using such high-power HF facilities have by now had important applications to a number of areas of plasma physics, telecommunications science, and basic ionospheric research. Such "ionospheric modification" or "heating" facilities have been operated in the United States (at Arecibo, Puerto Rico and Platteville, Colorado), the USSR (at Gor'kii and Moscow) and most recently in Europe (at Tromso, Norway).

HF ionospheric modification research has many potential applications to studies of ionospheric effects on radio wave systems. High-power radio waves can be used to produce controlled ionospheric plasma environments to investigate the temporal and spatial evolution of induced ionospheric disturbances. Recent work in this area has concentrated on the dynamics and dissipation properties of ionospheric field-aligned density irregularities and artificial spread-F plasma striations. In addition, induced ionospheric irregularities or changes in the ionospheric thermal and IR backgrounds may perhaps be used to provide countermeasures to radars and satellite-based surveillance systems, to create new propagation paths not naturally present for over-the-horizon (OTH) radars or

command-control-communication-intelligence (C^3I) systems, or to serve as the basis for mitigation techniques against the effects of other ionospheric disturbances, either natural or man-made. Through frequency-scaled experiments, HF ionospheric research can be used to determine thresholds for high-power RF systems to excite nonlinear ionospheric interactions, and their associated propagation degradation and plasma turbulence effects. Furthermore, multi-frequency HF ionospheric heating can be applied to the development of new techniques for generating wide-coverage low-frequency communications signals.

As an active experimental procedure, HF ionospheric modification is a complementary technique to other current experimental approaches used in C^3I research, such as those employing chemical releases or studying large-scale equatorial spread-F disturbances. The HF experimental philosophy is usually one of driving the ionospheric plasma, through wave self-action and parametric wave-plasma interaction effects, to the desired experimental initial conditions, removing the external RF driving force, and releasing the ionosphere to relax back to its ambient state through natural plasma processes. This kind of approach is particularly well-suited to the study of late-time structural phenomenology such as striation drift and decay.

HF ionospheric modification experiments can have several advantages over other experimental strategies for the investigation of ionospheric disturbances and their dynamic evolution. Because Arecibo and Platteville both represent geomagnetic mid-latitude locations, the ionosphere is usually calm and stable, an ideal plasma laboratory-without-walls. HF/ionosphere interaction physics is relatively well understood; we can produce ionospheric disturbances such as plasma striations on command in a known and reproducible manner. This control allows us to study disturbance formation, development, drift and dissipation under a variety of natural ionospheric conditions. In this respect, the disturbances are evolving in a natural O^+ plasma, with on-going recombination and (daytime) photodissociation, similar to other ionospheric scenarios of interest such as the high-altitude nuclear environment. No major natural instabilities are required to generate the HF-induced ionospheric disturbances, and experiments are not restricted to any particular time of day. As a result, ionospheric structures such as striations can be studied as a single phenomenon, not as the consequence of some complex, interactive system of events. In addition, in this controlled experimental environment programs can easily be designed and carried out to investigate the actual effects of ionospheric disturbances on specific C^3I systems.

THEORY OF NONLINEAR PHENOMENA

This section of the paper describes qualitatively the principal nonlinear phenomena that can be excited by high-power radio wave propagation through the ionosphere. Detailed quantitative analyses of the described phenomena can be found in Gurevich (1978) and Fejer (1979).

Change in the Absorption and Modulation of the Wave. The collisional heating and cooling processes of the ionospheric plasma are all dependent on the electron temperature. Heating of the electrons in the field of high-power radio waves produces a change in the collision frequencies of the electrons with the ions and with the background neutral molecules and atoms, affecting the radio wave absorption. As the electrons gain energy from the radio waves and from solar UV radiation, they also lose energy by collisions with atoms and molecules of the background gas. In the collision-dominated lower ionosphere, thermal conduction is not an important cooling mechanism and the collision frequency increases with increasing electron temperature. The wave absorption therefore strongly increases with increasing wave power, limiting the penetration of high-power radio waves into the interior of the plasma. The field of the wave reflected from the ionosphere may even decrease with increasing incident radiation power.

The most effective kinds of electron energy transfer collisions in the lower ionosphere are inelastic interactions with O_2 and N_2, producing rotational and vibrational excitation, and collisions with atomic oxygen, producing excitation of hyperfine levels of the ^3P ground state. For sufficiently strong radiation, the rate of heating may increase much faster than the normal cooling interactions, initiating a rapid increase in the electron temperature that continues until compensating processes set in that limit the temperature rise. The net result is a phenomenon originally described as an electron thermal runaway (Holway and Meltz, 1973). We now understand that the compensating processes that saturate the heating develop quickly enough to preclude an actual runaway in electron temperature, although significantly enhanced electron heating can occur (Perkins and Roble, 1978; Duncan and Zinn, 1978).

In the upper atmosphere thermal conduction is the principal cooling process, rapidly diffusing excess heat along the geomagnetic field lines. In addition, at these altitudes wave absorption decreases with increasing wave power. As a result, the ionospheric plasma becomes more and more transparent to higher power radio waves, and much longer propagation paths become possible.

Electron heating can affect the electron-ion recombination rates, changing the local ionospheric density. Differential heating and associated thermal gradients can cause redistributions of the

plasma. Heating-induced changes in the local ionospheric conductivities can be used to generate LF and VLF radio waves emitted from the heated ionospheric regions. As a result of the nonlinear dependence of wave absorption on the incident field amplitude, strong amplitude-modulated radio waves may suffer appreciable self-distortion of their modulation as they propagate through the ionosphere. Analogous nonlinear distortions can affect the waveform of the envelope of strong radio wave pulses.

Change of Wave Refraction and Self-Focusing. Natural density fluctuations in the ionosphere cause small variations in the index of refraction of the plasma, resulting in a slight focusing and defocusing of an electromagnetic wave propagating through the medium. The electric field intensity increases as the incident wave refracts into regions of comparatively underdense plasma. Ohmic heating and the electric-field ponderomotive force (a radiation pressure) then drive plasma from these focused regions, further amplifying the initial perturbation. This self-focusing process continues until hydrodynamic equilibrium is reached, creating field-aligned striations within the plasma. This process is illustrated schematically in Figure 2.

As a result of the change of the refractive index of the wave in the plasma, the beam trajectory is distorted and the wave's direction and propagation path are shifted. For radio waves reflecting in the ionosphere, if thermally-induced imbalances of ionization are not significant, the wave's reflection height shifts upwards. Sufficiently powerful and narrow beams may then penetrate through critical density layers in the ionosphere.

Even a weak plasma nonlinearity causes radio wave propagation trajectories to bend noticeably. This leads to focusing and filamentation of the incident radio wave beam, to the onset of wave field intensity oscillations, and to stratification of the ionospheric plasma density. The resulting large-scale field-aligned ionospheric irregularities can impose severe scattering, scintillation, and distortion effects on other radio waves over a wide frequency range which may propagate through this modified region.

Wave Interaction. The nonlinearity of radio wave propagation through the ionosphere also manifests itself as a violation of the principle of superposition of waves. When a high-power radio wave propagates through a plasma, it changes the absorption and refraction properties of the medium not only for itself, but also for all other radio waves passing through the same region. The radio waves thus interact in the plasma. In particular, if the high-power radio wave is amplitude-modulated, then the resulting changes of absorption can cause this modulation to be transferred to other waves passing through the same region of the ionosphere. This phenomenon of cross modulation is of practical importance for radio wave systems operating at medium wavelengths and for short pulse wave interaction studies of the ionosphere. Absorption changes induced by strong radio waves can become extremely large, and in practice can enable a strong radio wave to totally suppress or distort beyond informational use other RF signals propagating through the perturbed region.

Nonlinear interaction between several high-power radio waves can also be used to generate new radio waves at combination and difference frequencies, to drive resonant plasma oscillations unstable, or to produce new LF and VLF radio waves by low-frequency modulation of a single high-power wave. In addition, various normal components of the radio wave polarization nonlinearly interact with one another in a magnetoactive plasma. This can lead to a nonlinear rotation of the polarization ellipse and to self-modulation effects. The nonlinear interaction of radio waves in the ionosphere can also generate through nonlinear wave transformation processes other types of electromagnetic and electrostatic waves, such as plasma waves, whistlers, ion-sound waves, magnetohydrodynamic waves, or acoustic waves.

Breakdown Ionization. In the fields of very powerful radio waves, free electrons may become accelerated to energies sufficient to impact ionize molecules and atoms in the ionosphere during collisions. This electric breakdown of the neutral gas results in a rapid and large increase in the plasma electron number density. The nonlinearity associated with this process also leads to a very fast increase in the wave absorption, and to a rapid saturation of the wave field. The effect is thus self-quenching. An equally strong increase occurs in the absorption of other waves passing through the ionization region; the high-power radio wave effectively attenuates other RF signals propagating through this medium. In addition, the artificial ionization regions can be used to scatter or reflect VHF and UHF radio waves which normally pass through the ionosphere unimpeded, potentially impacting numerous telecommunications systems relying on transionospheric propagation.

Instability Excitation. High-power radio waves increase the electron temperature and change the electron and ion densities in the modified ionosphere. Artificial inhomogeneities can easily become unstable, such as with respect to flute, drift and ion-cyclotron waves. Ionospheric modification in the field of high-power radio waves can excite new plasma oscillations and enhance existing oscillations, forming an electrostatically turbulent plasma environment. This strongly influences the conditions of radio wave propagation in the modified region of the ionosphere, causing additional anomalous signal absorption and scattering.

Resonant instabilities can be excited in regions where the radio wave frequency is close to some natural oscillation frequency of the plasma. In these regions, the electromagnetic wave may excite natural plasma oscillations through linear wave transformation. This process leads to the development of small-scale field-aligned plasma striations and the effective absorption of radio waves by the plasma. The resulting small-scale plasma irregularities also effectively scatter VHF and UHF radio waves over a wide frequency band.

Another important nonlinear phenomenon that occurs in the plasma resonance region is the excitation of parametric instabilities. Parametric wave-plasma interactions can generate enhanced plasma and ion-acoustic waves in the field of a powerful radio wave. In addition, parametric instabilities can lead to effective nonlinear generation of noise and to absorption of the incident radiation by the plasma. Parametric interactions are usually accompanied by pulsations of the wave reflected from the ionosphere and by the generation of energetic electrons (1-20 eV) accelerated as a result of Landau damping of the enhanced plasma waves. Interactions between parametrically-enhanced electron plasma waves may also lead to the formation of short-scale plasma striations capable of coherently scattering VHF and UHF radio waves.

In addition to their local ionospheric effects, nonlinear wave-plasma interactions may be responsible for a number of new phenomena outside of the modified region. Electrons accelerated in the perturbed F-region ionosphere and high-power low-frequency radiation produced by wave modulation or wave-wave interaction processes both may have a strong influence on the behavior of the magnetospheric plasma. Wave energy absorbed by the ionsopheric free electrons ultimately may be transferred through collisions to the neutral molecules and atoms of the upper atmosphere. This process can result in noticeable heating of the neutral atmsophere, airglow both in the visible and infrared, composition changes, and possibly the launching of acoustic-gravity waves and their associated travelling ionospheric disturbances. Each of these perturbations, through their interactions with other radio waves. ionospheric fields and currents, fast electrons, and magnetospheric whistlers, may themselves give rise to additional nonlinear phenomena.

EXPERIMENTAL HISTORY

The initial high-power high-frequency ionospheric modification facilities were constructed in anticipation of rather gentle enhancements of the ionospheric free electron temperature, with associated plasma redistribution. However, in addition to the expected increases of the electron gas temperature, a rich spectrum of plasma instabilities and nonlinear wave propagation effects were also excited. The most prominent nonlinear phenomenon detected in these early experiments was associated with HF-induced parametric instabilities. producing enhanced plasma waves, anomalous RF absorption, accelerated energetic electrons, and airglow excitation (Carlson and Duncan, 1977). These studies benefited greatly from the active interest of many leading plasma physicists [e.g., Perkins and Kaw, 1971; Valeo et al., 1972; Fejer and Lear, 1972; Harker, 1972; DuBois and Goldman, 1972; Rosenbluth, 1972]. Reviews of results obtained in these early experiments can be found in special issues of Radio Science (9, 11, 1974) and Izv. Vussh. Uchebn. Raved. Radiofiz. (18, 9, 1975; 20, 11, 1977).

The parametric instability mechanism entails a pump or driving field whose energy cascades into plasma oscillations at two lower natural resonant frequencies in the plasma. In this case, the high-power HF electromagnetic radiation provides the initial driving field, and the longitudinal electrostatic electron plasma wave and the ion-acoustic wave (parametric decay instability) or zero-frequency ion mode (two-stream or purely growing instability) represent the parametrically enhanced oscillations. Radar measurements of the spectra of these enhanced plasma waves (Kantor, 1974; Showen, 1975; Duncan, 1977) investigated the enhanced plasma wave structure, including additional features apparently associated with the saturation of the parametrically enhanced oscillations. These structures were satisfactorily explained in terms of a saturation mechanism based on secondary parametric decay interactions, with the enhanced electrostatic plasma oscillations acting as new pump waves (Perkins et al., 1974; Fejer and Kuo, 1973). The instability threshold and saturation spectrum is affected by the local plasma collision frequency, density gradient, geomagnetic orientation, and amplification of the incident radio wave through focusing (self-action or pre-existing ducts) and Airy-structure swelling near the reflection altitude. Figure 3 presents an ambient ionospheric electron density profile as measured with the Arecibo incoherent backscatter radar, and a corresponding profile measured with the added influence of high-power HF radiation, showing a narrow region of parametrically-excited electrostatic turbulence near the HF reflection height.

Pulsed radar studies of the ionosphere often detect multiple or spread return echoes. Such echoes from the natural ionosphere have been used historically to identify and characterize an F-region phenomenon called spread-F, and have been attributed to electron density irregularities. The physics behind the generation and maintenance of the irregularities is still uncertain. Morphologically, naturally occurring irregularities are seen in the F-region ionosphere with fractional density changes as large as 10^{-2} to 10^{-3} in only a few kilometers. Satellite-borne probes on Atmospheric Explorer have measured scale sizes from several kilometers to as small as 60 m,

limited by the telemetry rate on that information channel; coherent radar studies, particularly involving the 50-MHz radar at Jicamarca, Peru, have measured spread-F echoes from structure with 3-meter scale size. A condition indistinguishable on an ionogram from naturally-occurring spread-F can be generated reproducibly by illuminating the ionosphere with intense HF radio waves (Figure 4; Utlaut and Violette, 1972). Experimental HF radar studies (Figure 5; Thome and Perkins, 1974), radio scintillation measurements (Figure 6; Gel'mantsev et al., 1976; Rufenasch, 1973), AE satellite measurements (Figure 7, J. P. McClure, 1977; Carlson and Duncan, 1977), and interaction region striation mapping (Figure 8; Duncan and Behnke, 1978) have presented observational data to support interpretation of these large-scale spread-F effects as due to a thermal self-focusing instability (Perkins and Valeo, 1974; Perkins and Goldman, 1980). Additional studies have confirmed that these HF irregularities apparently form with power spectral densities similar to natural spread-F, barium cloud striations, and irregularity structure found in the high-altitude nuclear environment (Figure 9, Ganguly, 1980). Ray tracing simulations of observations have suggested scale sizes perpendicular to B of a few kilometers, growth times of a few minutes, and saturation fractional plasma concentration changes of the order of 10^{-1} (Allen et al., 1974), in general agreement with the experimental observations. A different approach considering stimulated Brillouin forward scattering (decay of the electromagnetic wave into another electromagnetic wave and an ion sound wave) has also been developed (Cragin and Fejer, 1974), with collisional heating dominating the ponderomotive force and generating field-aligned irregularities with perpendicular scale sizes on the order of 500 meters. Although the initial approach differs, the physics of these mechanisms is quite similar.

VHF and UHF coherent radar observations of aspect-sensitive echoes from field-aligned irregularities have demonstrated the generation and decay of short-scale plasma striations on time scales on the order of 10 ms. The short-scale structures are believed to be excited by secondary thermal processes associated with the parametrically-enhanced electron plasma waves (Perkins, 1974; Vaskov and Gurevich, 1975; Cragin et al., 1977; Lee and Fejer, 1978). Experimental observations of this coherent scatter have been detected for radar frequencies from 50 to 400 MHz (Figure 10; Minkoff, 1974). The observed short-scale irregularities exhibit strong temporal and spatial variabilities (Figure 11; Frey, 1980). The role of short-scale field-aligned plasma striations in scattering of the incident HF radiation, producing an overshooting of the HF induced parametric effects, and as a source of anomalous signal absorption in the HF interaction region, is a current subject of study.

As we can see from the preceding descriptions, nonlinear phenomena accompanying ionospheric modification by high-power radio waves constitute a rather diverse and extensive class of effects. This is due both to the inherent variety of nonlinear effects in plasmas and to the great differences in physical conditions found in the ionosphere as a function of time and altitude. Generally, the predominant effects in the lower ionosphere are associated with nonlinear changes in wave absorption, while upper ionospheric phenomena are more usually associated with changes in wave refraction and wave-plasma instabilities.

IONOSPHERIC IRREGULARITIES

The study of plasma striation dynamics using ionospheric modification by high-power radio waves is just in its infancy. Previous observations have demonstrated that ionospheric irregularities can be formed over a wide range of ionospheric conditions in a controlled experimental environment. A comprehensive array of ionospheric diagnostics have been fielded to investigate the associated plasma behavior. The principal experimental results are:

(1) Large-scale ionospheric irregularities form preferentially with 500 m to 1 km dominant scale sizes.
(2) The irregularities consist of density fluctuations of $\Delta n/n \approx 2\text{-}10\%$.
(3) These large-scale structures decay on time-scales of several tens of minutes to hours, but also show significant early dissipation effects on a time scale of less than a few minutes.
(4) Small-scale field-aligned striations form with scale sizes of approximately one-third meter to several meters under specific ionospheric modification conditions and within a narrow altitude region.
(5) The small-scale structure decays on time-scales of less than 1 s.
(6) Both large- and small-scale irregularities drift together with apparently the background wind velocities.

Our investigation of HF-induced ionospheric irregularities suffers most from a lack of interaction with the theoretical scientific community studying irregularity stability and evolution. We solicit your advice and assistance in pursuing this relatively underdeveloped experimental capability. The observations to date address a number of unresolved phenomenological issues. These outstanding questions include:

(1) Why do large-scale irregularities form preferentially with 500 m to 1 km scale sizes?
(2) Why do these large-scale structures decay with such long time constants, and what is the dissipation mechanism?

(3) What determines the geometry and spatial extent of the large-scale ionospheric irregularities?
(4) By what processes do small-scale field-aligned striations form and decay?
(5) What secondary instabilities are likely to develop during the evolution of the irregularities?
(6) How do ionospheric disturbances drift?

In addition to the above physics questions, we can begin to address the problems of designing definitive experimental programs to determine the size and duration of ionospheric disturbance impacts on generic C^3I systems. The potential applications of HF ionospheric modification research to controlled experimental investigations of RF system performance in disturbed ionospheric conditions could be vigorously pursued if encouraged by the C^3I scientific community.

REFERENCES

Allen, E. M., G. D. Thome, and P. B. Rao, HF Phased Array Observations of Heater-Induced Spread-F, Radio Sci., 9, 905, 1974.

Carlson, H. C. and L. M. Duncan, HF Excited Instabilities in Space Plasmas, Radio Sci, 12, 1001, 1977.

Cragin, B. L. and J. A. Fejer, Generation of Large-Scale Field-Aligned Irregularities in Ionospheric Modification Experiments, Radio Sci, 9, 1071, 1974.

Cragin, B. L., J. A. Fejer, E. Leer, Generation of Artificial Spread-F by a Collisionally Coupled Purely Growing Parametric Instability, Radio Sci., 12, 273, 1977.

DuBois, D. F. and M. V. Goldman, Spectrum and Anomalous Resistivity for the Saturated Parametric Instability, Phys. Res. Lett., 28, 218, 1972.

Duncan, L. M., Experimental Saturation Spectrum of Parametric Instabilities, Ph.D. dissertation, Rice University, Houston, TX, 1977.

Duncan, L. M. and R. A. Behnke, Observations of Self-Focusing Electromagnetic Waves in the Ionosphere, Phys. Rev. Lett., 41, 998, 1978.

Duncan, L. M. and J. Zinn, Ionosphere-Microwave Interactions for Solar Power Satellites, Los Alamos National Laboratory report LA-UR-78-758, March 1978.

Fejer, J. A. and E. Leer, Excitation of Parametric Instabilities by Radio Waves in the Ionosphere, Radio Sci., 7, 481, 1972.

Fejer, J. A. and Y. Y. Kuo, Structure in the Nonlinear Saturation Spectrum of Parametric Instabilities, Phys. Fluids, 16. 1490, 1973.

Fejer, J. A., Ionospheric Modifications and Parametric Instabilities, Rev. Geophys. and Space Physics, 17, 135, 1979.

Frey, A., Artificially Produced Field-Aligned Short-Scale Striations, M.S. thesis, Rice University, Houston, TX, 1980.

Ganguly, S., Short-Term Variability in the HF-Enhanced Plasma Line, J. Geophys. Res., 85, 1783, 1980.

Gurevich, A. V., Nonlinear Phenomena in the Ionosphere, Physics and Chemistry in Space, Vol. 10, Springer, Berlin-New York, 1978.

Getmantsev, G. G., L. M. Erukhimov, E. E. Mityakova, N. A. Mityakov, N. M. Prytkov, V. O. Rapoport, and V. A. Cherepovitskii, Certain Results of Investigating Ionospheric Irregularities Induced by Powerful Radio Emission Using Baseground Reception of Satellite Signals, Izv. Vussh. Uchebn. Raved. Radiofiz., 19, 505, 1976.

Harker, K. J., Induced Enhancement of the Plasma Line in the Backscatter Spectrum by Ionospheric Heating, J. Geophys. Res., 77, 6904, 1972.

Holway, L. H. and G. Meltz, Heating of the Lower Ionosphere by Powerful Radio Waves, J. Geophys. Res., 78, 8402, 1973.

Kantor, I. J., High Frequency Induced Enhancements of the Incoherent Scatter Spectrum at Arecibo, 2, J. Geophys. Res. 79, 199, 1974.

Lee, M.-C. and J. A. Fejer, Theory of Short-Scale Field-Aligned Density Striations Due to Ionospheric Heating, Radio Sci., 13, 893, 1978.

McClure, J. P., private communications, 1977.

Minkoff, J., Radio Frequency Scattering from a Heated Ionospheric Volume, 3, Cross-Section Calculations, Radio Sci., 9, 997, 1974.

Perkins, F. W. and P. K. Kaw, On the Role of Plasma Instabilities in Ionospheric Heating by Radio Waves, J. Geophys. Res., 76, 282, 1971.

Perkins, F. W. and E. J. Valeo, Thermal Self-Focusing of Electromagnetic Waves in Plasmas, Phys. Rev. Lett., 32, 1234, 1974.

Perkins, F. W., C. Oberman, E. J. Valeo, Parametric Instabilities and Ionospheric Modification, J. Geophys Res., 9, 1065, 1974.

Perkins, F. W. and R. G. Roble, Ionospheric Heating by Radiowaves: Predictions for Arecibo and the Satellite Power Station, J. Geophys. Res., 83, 1611, 1978.

Perkins, F. W. and M. V. Goldman, Self-Focusing of Radio Waves in an Underdense Ionosphere, J. Geophys. Res., 86, 600, 1981.

Rosenbluth, M. N., Parametric Instabilities in Inhonogeneous Media, Phys. Rev. Lett., 29, 565, 1972.

Rufenach, C. L., Radio Scintillation of Stellar Signals During Artificial Ionospheric Modification, J. Geophys. Res., 73, 5611, 1973.

Showen, R. L., Time Variations of HF Induced Plasma Waves, Ph.D. dissertation, Rice University, Houston, TX, 1975.

Thome, G. D. and F. W. Perkins, Production of Ionospheric Striations by Self-Focusing of Intense Radio Waves, Phys. Rev. Lett., 32, 1238, 1974.

Utlaut, W. F. and E. J. Violette, Further Observations of Ionospheric Modification by a High-Powered HF Transmitter, J. Geophys. Res., 77, 6804, 1972.

Valeo, E., C. Oberman, and F. W. Perkins, Saturation of the Decay Instability for Comparable Electronic and Ion Temperatures, Phys. Rev. Lett., 28, 340, 1972.

Vaskov, V. V. and A. V. Gurevich, Nonlinear Resonant Instability of a Plasma in the Field of an Ordinary Electromagnetic Wave, Sov. Phys. JETP, 42, 91, 1975.

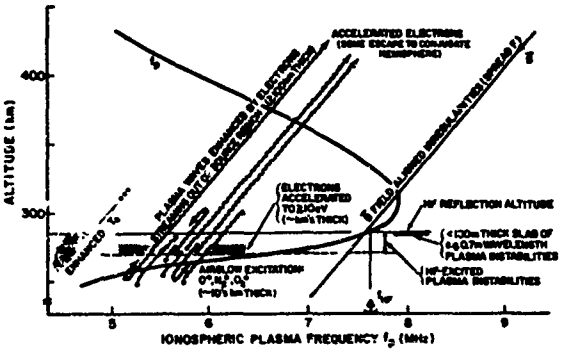


Figure 1. Effects produced by high-power HF ionospheric modification.

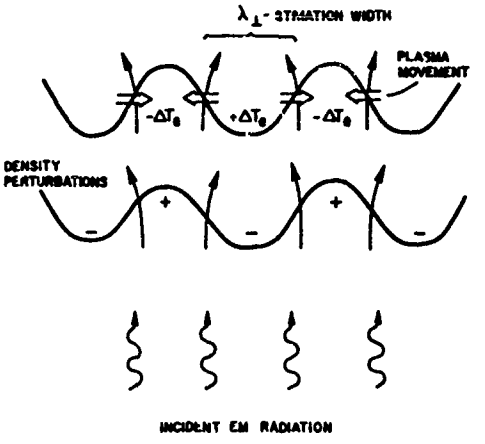


Figure 2. Schematic description of the thermal self-focusing processs. Increased electron heating in the focused regions produces a temperature gradient that drives plasma out of the region, further focusing the beam.

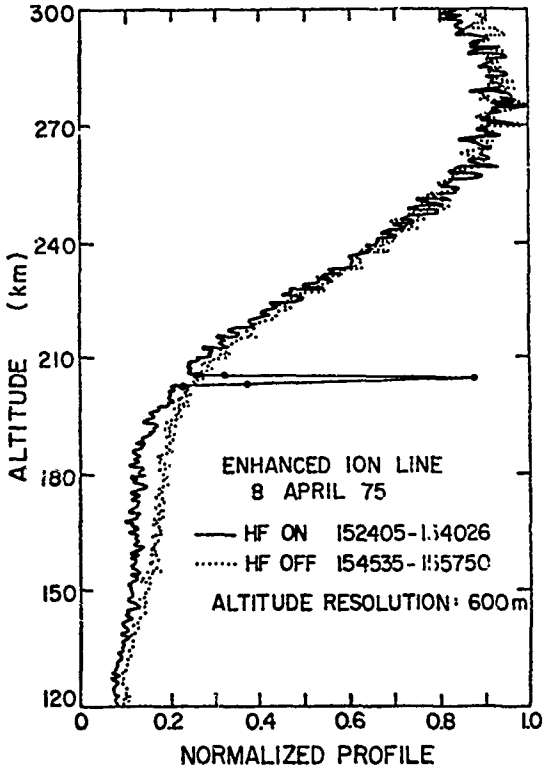


Figure 3. Arecibo ionospheric profiles measured by incoherent scatter radar, with and without HF ionospheric modification effects.

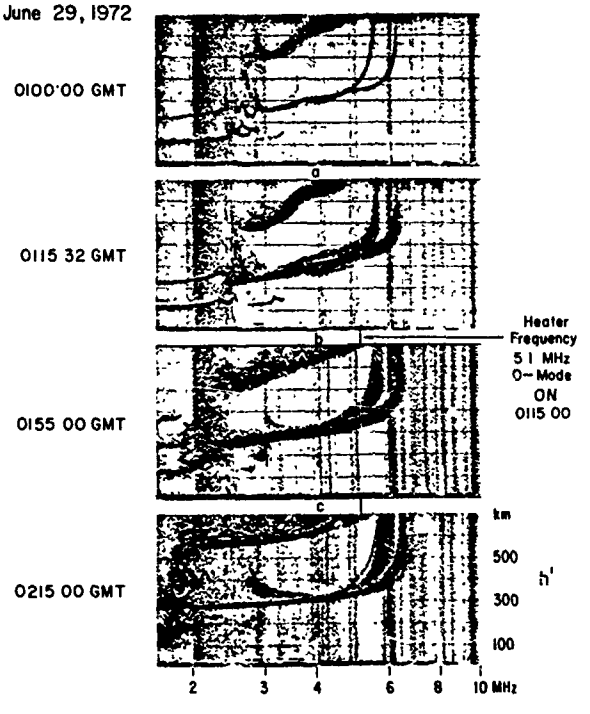


Figure 4. Ionograms illustrating artificial spread-F.

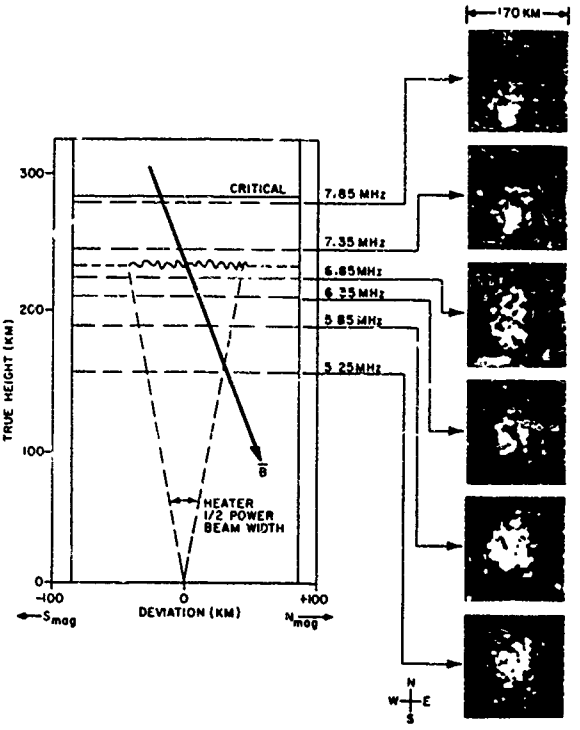


Figure 5. HF phased array frequency scan sky maps taken at six diagnostic frequencies.

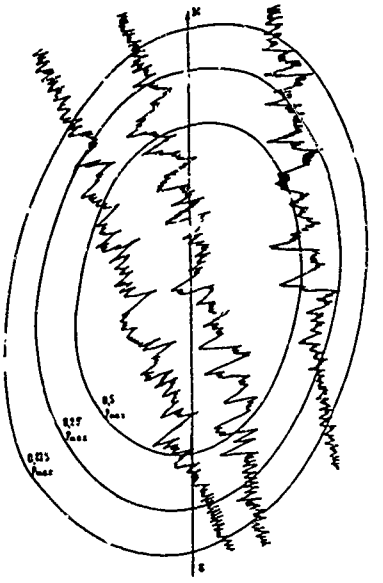


Figure 6. Satellite-to-ground signal fluctuations at 136-137 MHz for three flight paths through the modified ionosphere. The ovals estimate the heated region at 300 km.

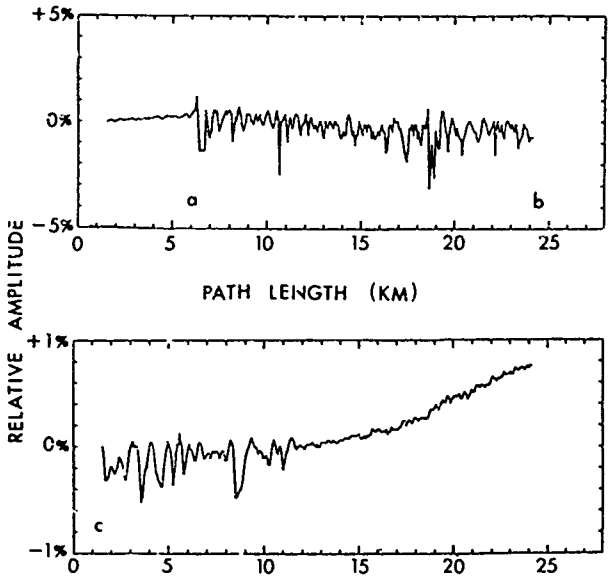
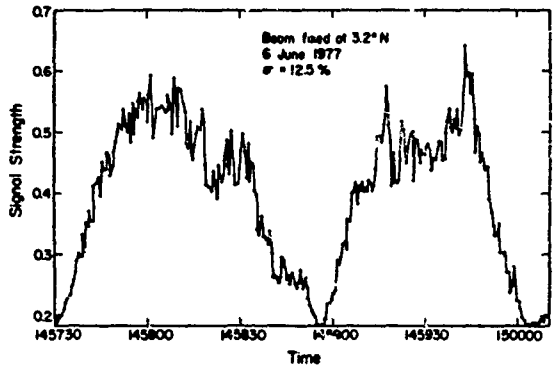


Figure 7. Atmospheric Explorer ion density fluctuation data for flight path through the modified region, (a) onset of instabilities (b) no data for next 23 km (c) data resumes.



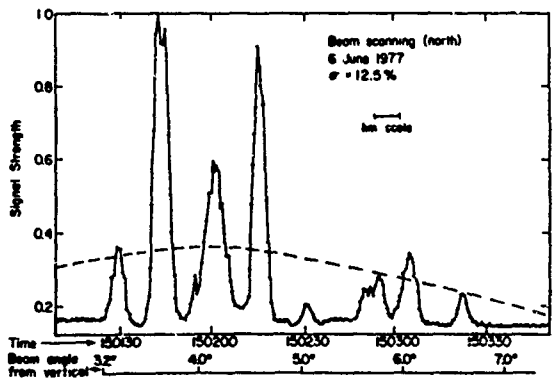


Figure 8. Experimental observations of wave self-focusing. The top figure shows the backscattered enhanced plasma line signal modulation induced by the natural drift of self-focusing striations through the fixed radar beam. The lower figure presents a series of striations observed from rapid scanning of the radar beam across the interaction region immediately after the drift measurements. The dashed curve estimates the unstriated beam.

Figure 10. Observed coherent backscatter coefficient for scatter from small-scale field-aligned plasma striations, as a function of frequency.

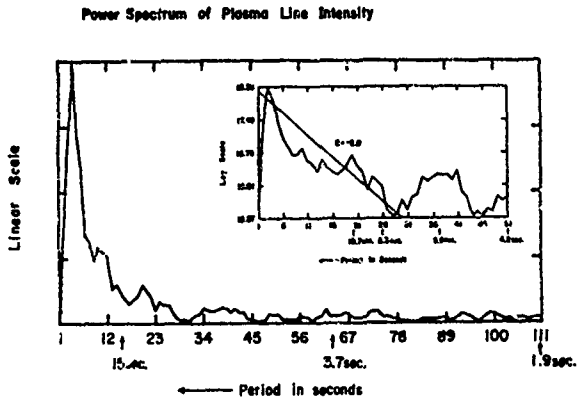


Figure 9. The power spectrum of fluctuations in the enhanced plasma line intensity, indicative of the associated electron density fluctuations.

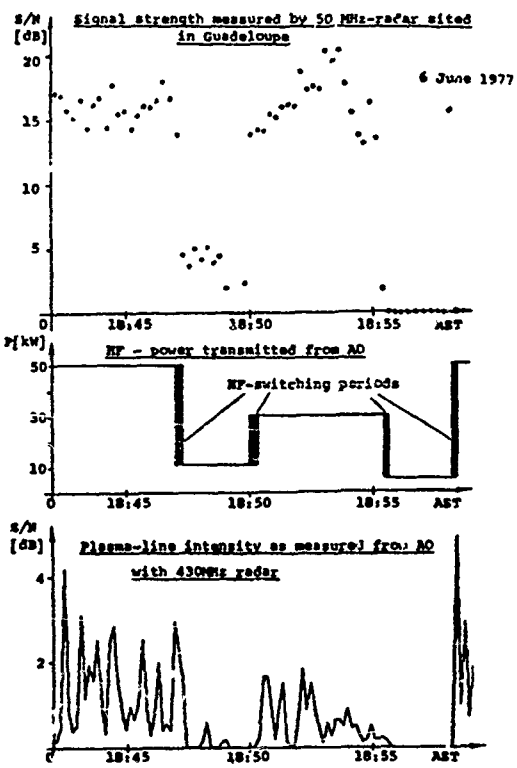


Figure 11. Comparison of 50-MHz radar coherent scatter signal strength with HF transmitted power and enhanced plasma line intensity.

ARTIFICIAL IRREGULARITIES GENERATED BY IONOSPHERIC HEATING AND THEIR EFFECTS ON TRANSIONOSPHERIC PROPAGATION

Santimay Basu and Sunanda Basu
Emmanuel College
Boston, MA 02115

S. Ganguly
Department of Space Physics
Rice University
Houston, TX 77001

J.A. Klobuchar
Air Force Geophysics Laboratory
Hanscom AFB, MA 01731

C.M. Rush
U.S. Department of Commerce
Institute for Telecommunication Sciences
Boulder, CO 80303

ABSTRACT

The results of 250 MHz scintillations observed during ionospheric heating in both the overdense (heater frequency below the critical frequency) and the underdense (heater frequency above the critical frequency) cases with the high-power high frequency transmitter at Platteville, Colorado are discussed. In the overdense case, strong irregularities are found to be excited promptly within a few seconds giving rise typically to scintillations in the range of 3-6 dB at 244 MHz. In the underdense case of heating, on the other hand, weak irregularities are usually excited after some delay and are found to cause 1-3 dB scintillations in the 244 MHz - 249 MHz frequency range, although one 10 dB scintillation event was encountered on a field aligned propagation path. The temporal structures of scintillations in the two cases of heating are found to be very different, with slow fadings dominating the scintillation structure in the underdense case. The spatial structures of irregularities generated in the two cases of heating are discussed from power spectral studies of scintillations and measurements of irregularity drift speed from spaced receiver scintillation observations.

Observations of radio star and satellite scintillations associated with ionospheric heating by the use of the newly constructed facility at Arecibo are also discussed. Radio star measurements were conducted at 50 and 430 MHz while geostationary satellite observations with three spaced receivers were made at 250 MHz. These preliminary measuremen s indicate discrete 'clumping' of irregularities near the center of the heated volume and a weak wavelength dependence of scintillations.

INTRODUCTION

It has been established that high power high frequency radio waves reflected from the ionosphere can not only introduce the expected modification of electron gas temperature and number density near the altitude of reflection (Gordon et al., 1971; Utlaut and Cohen, 1971), but can also give rise to a variety of physical phenomena related to non-linear plasma physics (for comprehensive reviews see Perkins et al., 1974; Carlson and Duncan, 1977; Fejer, 1979; Gurevich, 1978). Among the various manifestations of plasma instabilities induced by the heating experiments, the generation of artificial spread-F was one of the most striking and immediate experimental results (Utlaut et al., 1970; Utlaut and Violette, 1972; Wright, 1973). The generation mechanism of long wavelength (~1 km) field aligned irregularities remained obscure for quite sometime and is now attributed either to thermal self-focusing (Perkins and Valeo, 1974; Thome and Perkins, 1974) or to stimulated Brillouin scattering (Cragin and Fejer, 1974). The substantial level of the spectral intensity of km-scale

irregularities was subsequently demonstrated from the observations of scintillations of VHF/UHF signals received from radio stars and satellites through the artificially heated ionospheric volume (Rufenach, 1973; Pope and Fritz, 1974; Bowhill, 1974). Bowhill (1974) performed scintillation measurements with both geostationary and orbiting satellites and established the field-aligned nature of the irregularities, their transverse scale size and drift velocity. In all the above studies, the heater wave frequency was below the critical freuqency of the F region. Recently, Basu et al (1980) have performed scintillation measurements on the ground and on the moving platform of an aircraft with geostationary satellites when the heating transmitter was operated at a frequency both below and above the critical frequency of the F region.

In the overdense case of heating (heater frequency below the critical frequency), they noted a prompt excitation of the irregularities causing as large as 10 dB scintillations at 250 MHz. On the other hand, in the underdense case when the heater frequency was about 15% higher than the critical frequency, they observed a delayed and often sporadic onset of long period scintillations not generally exceeding 3 dB at 250 MHz. Perkins and Goldman (1981) have recently considered a theory of self-focusing instability in an underdense ionosphere and have predicted the generation of sheet-like irregularities aligned with the magnetic meridian that could have given rise to the scintillation effects discussed above.

In the present paper, we shall utilize our earlier geostationary satellite scintillation observations (Basu et al., 1980) performed in conjunction with ionospheric heating at Platteville for a study of the spectral characteristics of scintillations in the overdense and underdense cases of heating. We shall also discuss our recent results of radio star and satellite scintillation measurements in conjunction with ionospheric heating at the newly constructed facility at Arecibo.

EXPERIMENTAL DETAILS

The experimental details and the geometry of scintillation observations performed in conjunction with the heating transmitter at Platteville are outlined in Basu et al. (1980) and will not be repeated here. During December 1980, radio star and satellite scintillation measurements were conducted at Puerto Rico by using the newly constructed h.f. heating facility at Arecibo. Figure 1 shows the extent of the central heated region at 200 km altitude above the heating facility as limited by the estimated half power beam circle at 5 MHz (35 km E-W and 70 km N-S). Each of the four transmitters were operated at 75 kw power level during the period of observation. Ordinary mode heating was performed during the period of observation. The intersections of the ionospheric height of 200 km with the ray paths from the Roosevelt Roads Naval Station to the LES-9 satellite are shown in Figure 1. The subionospheric (200 km) tracks of several radio sources, namely Taurus, 3 C 166 and 3 C 210 as viewed by the 1000 ft radio telescope at Arecibo are shown in Figure 2. Radio star scintillation data were acquired during the meridian transit of these sources. The LES-9 satellite scintillation measurements were performed at 249 MHz with three spaced receiving systems. Scintillation measurements with Taurus were performed at 50 MHz with the 1000 ft reflector at Arecibo. The 50 MHz receiving system was kindly placed at our disposal by Dr. J. Röttger of Max Planck Institute, Lindau, West Germany. The other radio sources were observed with the 430 MHz receiving system of the Arecibo Observatory.

OBSERVATIONAL RESULTS

Figure 3a shows a 15 minute scintillation data segment that was obtained at Carpenter, Wyoming by the use of 249 MHz transmissions of LES-8 satellite on March 13, 1980. The heater at Platteville was operated at 9.9 MHz with ordinary mode polarization and the critical frequency of the F region was 10.3 MHz. The normalized second central moment (S_4) of intensity scintillations was 0.37. Figure 3b shows the corresponding scintillation spectrum. The high frequency roll off starts at a frequency of about 0.2 Hz and the slope of the roll off portion corresponds to a frequency 'f) dependence of $f^{-2.5}$. The observed slope is shallower as compared to the spectral slope of natural scintillations at midlatitudes. It should be noted that detectable spectral intensity is obtained at several Hz in the case of overdense heating.

In Figure 4a we show a sample of 249 MHz scintillation observed at Carpenter, Wyoming on March 13, 1980 during an underdense heating cycle. The heater frequency was 9.9 MHz during this cycle while the critical frequency of the F region was 7.9 MHz. The quasi-periodic fluctuations are reminiscent of naturally occurring scintillations caused by ionization gradients. Figure 4b shows the corresponding scintillation spectrum which indicates that the spectral power is concentrated below about 0.3 Hz with a very sharp high frequency roll-off.

Figure 5a illustrates another scintillation data sample acquired at Carpenter, Wyoming on March 13, 1980. The heater frequency was 9.9 MHz and the critical frequency was 8.5 MHz during this underdense heating period. Figure 5b shows the corresponding power spectrum. Spectral power is concentrated over the low frequency band and a very shallow slope ($f^{-1.5}$) is obtained over the roll-off portion. The comparison between the spectra obtained under two different kinds of heating

indicates that while spectral power exists up to several Hz in the overdense case, almost all the power is confined to frequencies below 1 Hz in the underdense case.

We shall now examine the results of satellite and radio star scintillation measurements that have been performed at Puerto Rico in conjunction with ionospheric heating (on December 22, 1980). Sustained heating cycles over several hours were maintained during this test in contrast to the 15-minute or shorter duration of heating cycles employed at Platteville.

Figure 6 shows 249 MHz scintillations observed at Roosevelt Roads station during an overdense heating cycle. The heater at Arecibo was operated at 5.1 MHz and radiated ordinary mode polarization. The ray path to the satellite intersected the western fringe of the half power beam circle as shown in Figure 1. Scintillation index of 2.5 dB corresponding to the S_4 index of about 0.15 was obtained.

Figure 7 shows scintillations observed with the Taurus radio source at 50 MHz during an overdense heating period. A peak to peak fluctuation of 3.7 dB corresponding to S_4 index of 0.21 is obtained.

Figure 8 shows scintillations observed with radio source 3 C 166 at 430 MHz during another overdense heating period. Maximum scintillation activity of about 0.8 dB was recorded. If we compare the scintillation levels obtained with Taurus and 3 C 166 at 50 MHz and 430 MHz respectively, a very weak frequency dependence ($f^{-.5}$) of scintillation is obtained. It should be emphasized, however, that the above comparison is not strictly valid since the measurements were not simultaneous. However, we make the assumption that steady state heating conditions were achieved and the background ionosphere did not change greatly during the fifty-minute delay between the two sets of observations.

Figure 9 shows the weak and slow variations of signal intensity of 3 C 210 at 430 MHz when the heater frequency was marginally higher than the critical frequency. Such variations of signal intensity were recorded briefly during underdense heating as was the case also at Platteville.

The interesting differences of the irregularity characteristics in the underdense and overdense cases of heating as discussed in this report needs further study. We plan to perform similar experments with ground based and airborne instruments at Arecibo later this year.

ACKNOWLEDGMENTS

The authors wish to thank J. Carroll and E.J. Violette for the operation of the Platteville facility and A. Veldhuis and D. Albino for the operation of the Arecibo heating facility. The assistance of H.D. Craft and the staff of the Arecibo Observatory are gratefully acknowledged. Special thanks are due to J.M. Goodman for providing necessary support at Roosevelt Roads Naval Station and to C. Ferioli for help with the observations.

The National Astronomy and Ionosphere Center is operated by Cornell University under contract with the National Science Foundation. The work at Emmanuel College was partially supported by NSF grant ATM78-25264 and AFGL contract F19628-81-K-0011.

REFERENCES

Basu, S., S. Basu, A.L. Johnson, J.A. Klobuchar and C.M. Rush, Preliminary results of scintillation measurements associated with ionosphere heating and possible implications for the solar power satellite, Geophys. Res. Lett., 7, 609, 1980.

Bowhill, S.A., Satellite transmission studies of spread-F produced by artificial heating of the ionosphere, Radio Sci., 9, 975, 1974.

Carlson, H.C. and L.M. Duncan, HF excited instabilities in space plasmas, Radio Sci., 12, 1001, 1977.

Cragin, B.L. and J.A. Fejer, Generation of large-scale field-aligned irregularities in ionospheric modification experiments, Radio Sci., 9, 1071, 1974.

Fejer, J.A., Ionospheric modification and parametric instabilities, Rev. Geophys. Space Phys., 17, 135, 1979.

Gurevich, A.V., "Nonlinear Phenomena in the Ionosphere", Springer Verlag, New York, 1978.

Gordon, W.E., H.C. Carlson and R.L. Showen, Ionospheric heating at Arecibo: First tests, J. Geophys. Res., 76, 7808, 1971.

Perkins, F.W., C. Oberman and E.J. Valeo Parametric instabilities and ionospheric modification, J. Geophys. Res., 79, 1478, 1974.

Perkins, F.W. and E.J. Valeo, Thermal self focusing of electromagnetic waves in plasma, Phys. Rev. Lett., 32, 1234, 1974.

Perkins, F.W. and M.V. Goldman, Self-focusing of radio waves in an underdense ionosphere, to be published in J. Geophys. Res., 1981.

Pope, J.H. and R.B. Fritz, Observations of artificially produced scintillations using satellite transmissions, J. Geophys. Res., 79, 1074, 1974.

Rufenach, C.L., Radio scintillation of stellar signals during artificial ionospheric modification, J. Geophys. Res., 78, 5611, 1973.

Thome, G.D. and F.W. Perkins, Production of ionospheric striations by self-focusing of intense radio waves, Phys. Rev. Lett., 32, 1238, 1974.

Utlaut, W.F., E.J. Violette and A.K. Paul, Some ionosonde observations of ionospheric modification by very high power high-frequency ground-based transmission, J. Geophys. Res., 25, 6429, 1970.

Utlaut, W.F. and Cohen, Modifying the ionosphere with intense radio waves, Science, 174, 245, 1971.

Utlaut, W.F. and E.J. Violette, Further observations of ionospheric modification by a high-powered HF transmitter, J. Geophys. Res., 77, 6804, 1972.

Wright, J.W., Kinesonde observations of ionosphere modification by intense electromagnetic fields from Platteville, Colorado, J. Geophys. Res., 78, 5622, 1973.

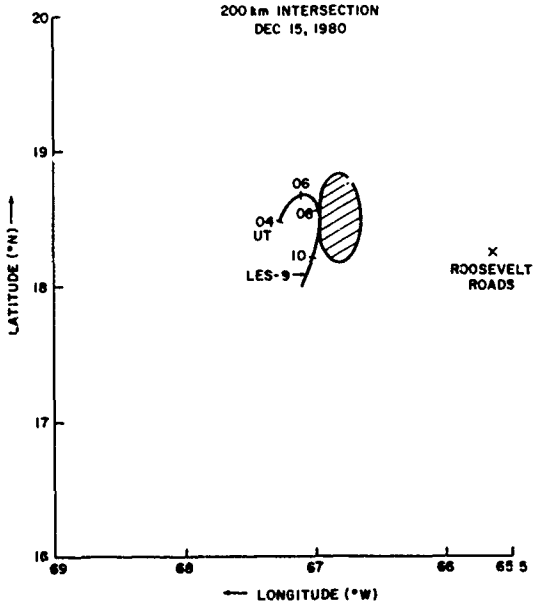


Figure 1. Geometry of observations of the LES-9 geostationary satellite from the Roosevelt Roads Naval Station in conjunction with the ionospheric heating at Arecibo in December, 1980.

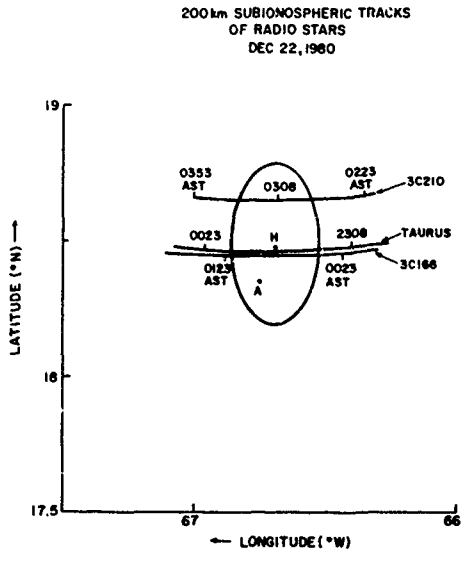
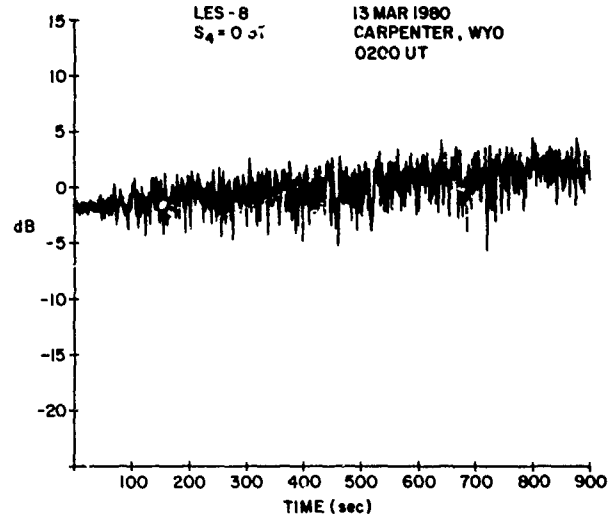


Figure 2. Subionospheric (200 km) tracks of Taurus, 3 C 166 and 3 C 210 as viewed by the Arecibo 1000 ft radio telescope (A) with the heater (H) on December 22, 1980.



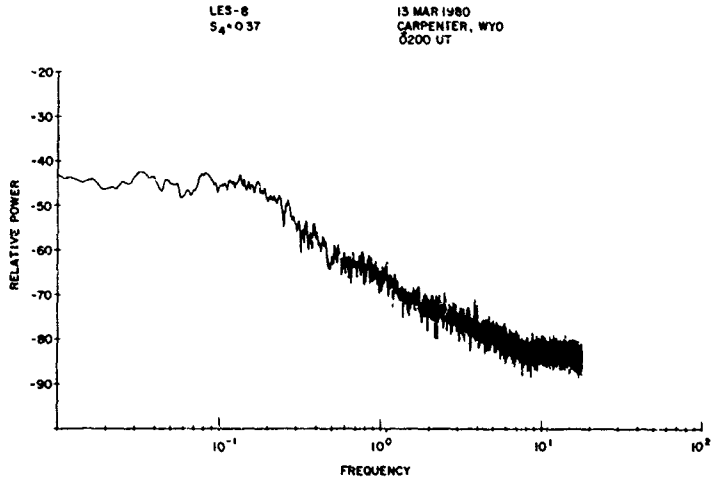
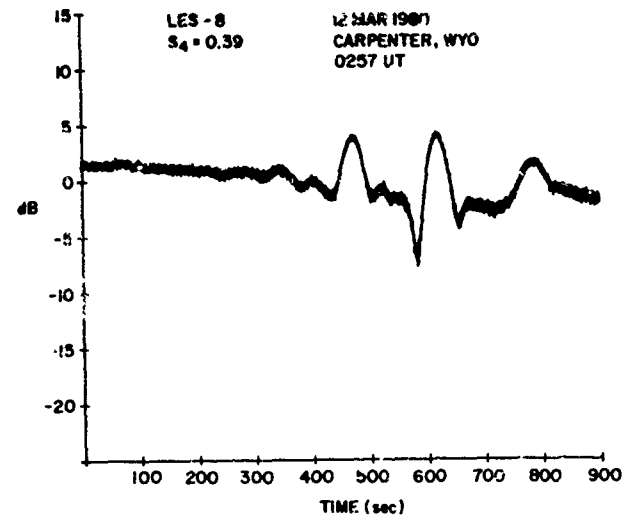


Figure 3a and b. A 15-min sample of scintillations from LES-8 at 249 MHz observed at Carpenter, Wyoming on March 13, 1980 during overdense heating and its corresponding spectrum



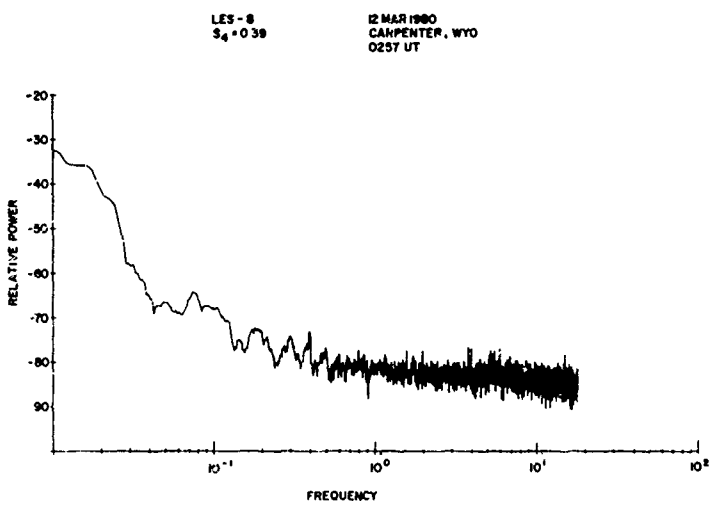
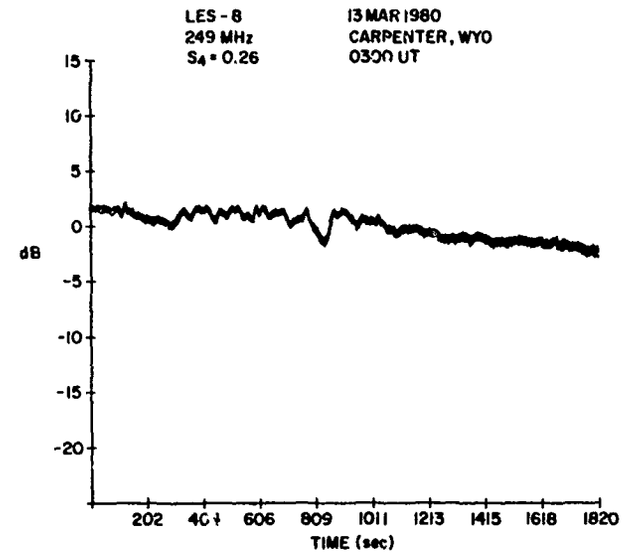


Figure 4a and b. A 15-min sample of scintillations from LES-8 at 249 MHz observed at Carpenter, Wyoming on March 13, 1980 during underdense heating and its corresponding spectrum



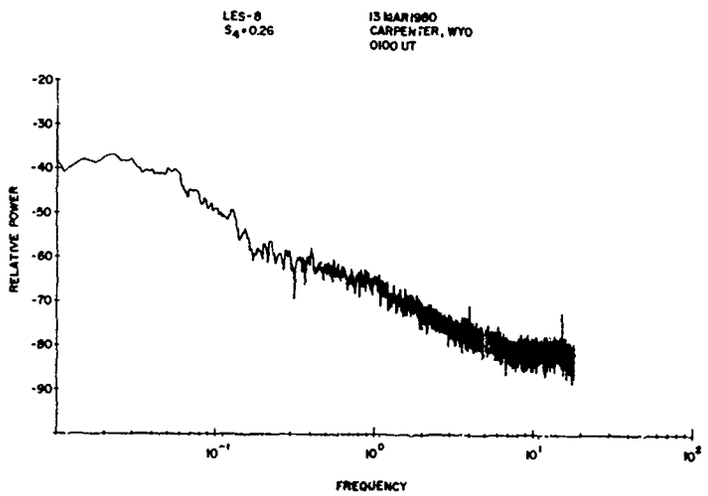


Figure 5a and b. A 15-min sample of scintillations from LES-8 at 249 MHz observed at Carpenter, Wyoming on March 13, 1980 during underdense heating and its corresponding spectrum

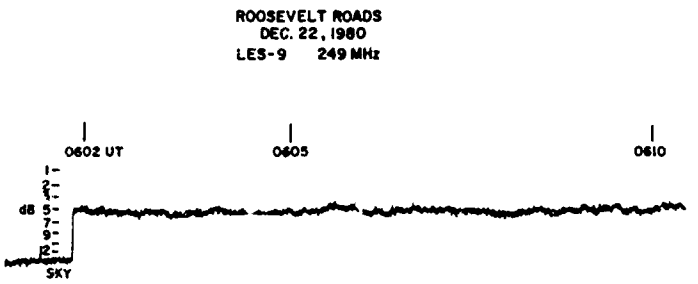


Figure 6. Scintillations from LES-9 at 249 MHz observed at Roosevelt Roads, P.R. on December 22, 1980 during overdense heating at Arecibo

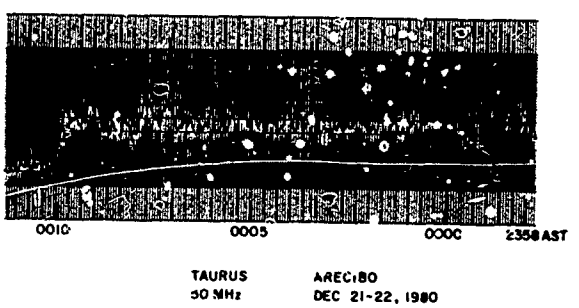


Figure 7. Scintillations at 50 MHz using Taurus radio source at Arecibo during overdense heating on December 22, 1980

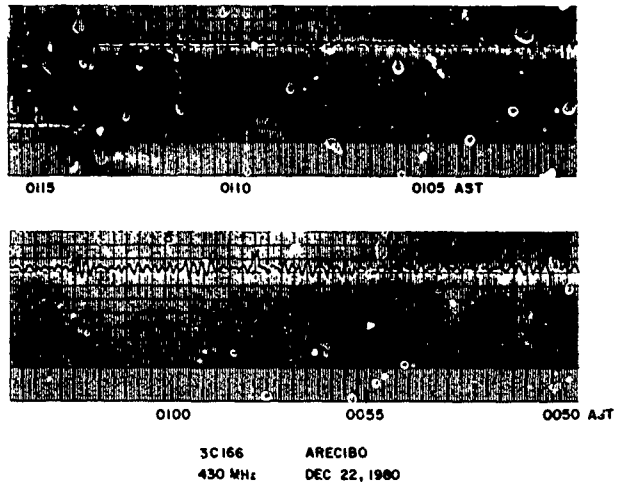


Figure 8. Scintillations at 430 MHz using 3 C 166 at Arecibo during overdense heating on December 22, 1980

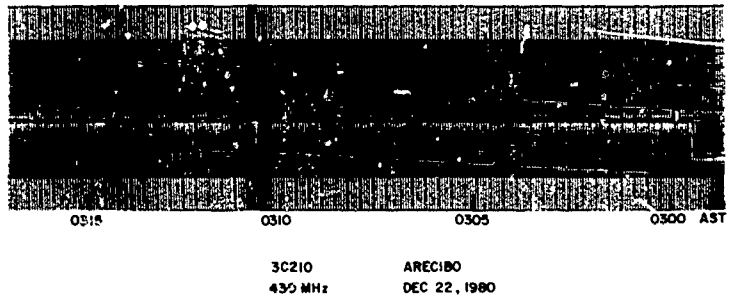


Figure 9. Scintillations at 430 MHz using 3 C 210 at Arecibo during marginally underdense heating on December 22, 1980

VERTICAL INCIDENCE PULSE ABSORPTION MEASUREMENTS DURING HIGH POWER RADIO WAVE HEATING OF THE D-REGION

A. A. Tomko, A. J. Ferraro and H. S. Lee
Department of Electrical Engineering
Ionosphere Research Laboratory
Pennsylvania State University
University Park, PA 16802

ABSTRACT

A vertical incidence pulse absorption experiment was conducted at the Arecibo HF heating facility in order to measure the change in absorption on a diagnostic wave arising from HF heating of the ionospheric plasma. For a heating frequency of 3.175 MHz, left circular polarization and an effective radiated power of 60 MW, a 9 db change on a 2.4 MHz diagnostic wave was observed. The absorption measurements were made after sunset and are consistent with predicted values for this time, based on plasma heating theory.

INTRODUCTION

Measurements of radio wave absorption by the vertical incidence pulse sounding technique (A1 method) have been used routinely by numerous experimenters for studies of the natural variations of the ionosphere (e.g. Gnanalingam and Kane, 1973, Ganguly, 1974). This technique has also been employed as a straight forward means of monitoring plasma modifications resulting from high power radio wave heating of the ionosphere. Measurements of this type were performed at the Platteville, Colorado high power HF transmitting facility during the early seventies, and were summarized by Utlaut and Violette (1974). The A1 absorption measurements were made on a 2.667 MHz diagnostic wave and showed about 6 dB increase in absorption over the two-way path from ground level to the diagnostic wave reflection height in the E-layer and back, as the plasma was heated at frequencies between 3 to 6 MHz and at an effective radiated power of about 100 MW. Diagnostic wave absorption was observed to increase promptly (< 40 ms) after turn on of the heating transmitter, but the time required for the absorption to return to its unheated level following a long period (10 min.) of CW heating was very slow (~ 10 min.). The prompt increase in absorption is attributed to the change in the electron-neutral collision frequency as the electrons are quickly heated by the high power wave field, while the long term recovery is most likely due to an electron density modification arising from the electron temperature dependence of various ion chemistry reaction rates (Meltz et. al., 1974; Tomko et. al. 1980a).

This paper will present the results of A1 absorption measurements made during the Fall, 1980 ionospheric modification program (Djuth, 1980) at the new high power HF transmitting facility of the National Astronomy and Ionosphere Center on the island of Puerto Rico. The A1 absorption measurements were made using a portable 2.4 MHz ionospheric sounding system developed at Penn State.

THE NAIC HEATING FACILITY

The NAIC HF heating facility is located on the northeast coast of Puerto Rico near the town of Islote, about 50 miles west of San Juan. The facility has four 200 KW transmitters which drive a 4 x 8 array of non-planar log periodic antennas operating in the backfire mode (Trask, 1979). The maximum directive gain of this array is expected to exceed 23 dBi over the 3 to 12 MHz range, giving an effective radiated power of over 160 MW. The half power beam width of the array is about 10 degrees.

THE 2.4 MHZ SOUNDING SYSTEM

The Penn State A1 sounding system was located 5 km southwest of the HF heating facility at Higuillales. This site was chosen because it is close enough to the heating facility that the vertical sounding intersects the main beam of the heater in the D-region, it is in mountainous terrain

thereby providing good shielding of the receiving equipment from the heater ground wave, and it is the only available site where direct communications with both the heating facility and the main NAIC facility, Arecibo Observatory, were available. A block diagram of the sounding system is given in Figure 1. The sounder consists of a 2.4 MHz transmitter with a peak power output of 5 KW which drives an inverted V dipole antenna. The transmitter is pulsed at a rate of 300 pulses/s and the pulse width is about 330 μs. The receiving system is a modified version of that developed by Sulzer (1973). The receiver has a gain of about 10^5 and a bandwidth of about 60 KHz (a compromise between narrow band for improved signal to noise ratio and a short rise time ($\sim$ 3 μs) for accurate location of the detected echo. The receiver contains automatic gain control circuitry developed by Sulzer (1979) to keep the detected receiver output constant for wave interaction measurements. The response time of the AGC system is about 0.05s. The AGC system generates a voltage which is proportional to the logarithm of the RF input voltage and thus serves as an indicator of diagnostic wave absorption. The AGC voltage was sampled using a 12 bit A/D converter. Data collection and experiment control were provided by a microcomputer and the digitized AGC voltages were stored on 5" floppy disks. The status of the heating transmitter (ON or OFF) was simultaneously sampled and recorded.

EXPERIMENTAL RESULTS

Table 1 gives a quick summary of the experimental conditions under which 2.4 MHz absorption was measured. On 8/29, all four heating transmitters were operational at 75KW on 3.175 MHz. Based on the model calculations of Trask (1979) the antenna gain at this frequency is about 23 dB (allowing 3 dB for losses). The effective radiated power for this day was thus estimated to be about 60 MW. On all of the other dates listed in Table 1, failures in one or more of the heating transmitters and their associated feedlines restricted the heating capability of the Islote HF facility to two transmitters operating at 75 KW or less and driving half of the 32 element array. The effective radiated power on these dates was $\lesssim$ 15 MW. At this reduced level of heating no correlated variation in Al absorption with heater status (i.e. OFF or ON) was observed. Accordingly, the data presented here are the 60 MW results from 8/29.

Figures 2 and 3 illustrate the observed variation in Al absorption at 2.4 MHz due to CW heating at 3.175 MHz and 60 MW ERP. The 2.4 MHz signal level at the input to the receiver is plotted in decibels relative to $10mV_{pp}$ as a function of time. Typical strong daytime echos from an unheated ionosphere are about $10mV_{pp}$. The data plotted in these figures has been digitally low pass filtered to a bandwidth of 50 mHz to minimize noise and short term variability of the data and thus give a relatively smooth plot. The dark bars at the bottom of each figure indicate periods when the heating transmitter was on while the white areas indicate periods of heater off.

Figure 2 presents a 23 minute segment of data at the start of the observation period on 8/29. The HF heating transmitter was on at the start of this period and the diagnostic signal level was relatively strong. Sunset occurred shortly after the start of the experiment and is readily apparent as a sharp drop in signal strength starting at about 1814 LT. A similar drop in signal level is observed at sunset during unheated conditions. The heating transmitter was turned off at 1820 for a three-minute period and a rapid increase in signal level was observed. The average signal enhancement during the off period was 8.4 $\pm$ 0.5 dB.

During the period 1856 to 1929 LT the HF heating transmitter was turned off and on at one minute intervals. Figure 3 shows the variation in 2.4 MHz absorption over a segment of this period. The average increase in signal level during the heater off periods was about 9dB.

COMPARISON WITH HF HEATING THEORY

Theoretical estimates of the dependence of 2.4 MHz absorption on HF heating power and frequency have recently been given by Tomko et. al. (1980b). These estimates were based on classical magnetoionic theory and current knowledge of electron energy loss rates. The approach taken was to obtain simultaneous solutions of the coupled differential equations governing power flow and electron energy in order to determine the change in electron temperature due to various levels of HF heating. The resultant electron temperature distributions were then used to compute the change in 2.4 MHz absorption over the two way path from ground to the wave reflection height and back again. The calculations of Tomko et. al. (1980b) were for daytime conditions corresponding to a solar zenith angle of 30°. In this work, similar calculations have been performed for electron density models corresponding to solar zenith angles of 30, 60, 70 and 80°. These models are given in Table 2, and the resultant electron temperature distribution and change in 2.4 MHz absorption due to HF heating at

3.175 MHz is given in Table 3. The electron temperature distribution is given in terms of the ratio of the heated electron temperature to its ambient unheated value. Two sets of values are given: one for 60 MW and another (in parentheses) for 120 MW of heating. The change in 2.4 MHz absorption due to these two levels of heating, the diagnostic wave reflection height and the height of the peak in the heated electron temperature distribution are also given in Table 3. The numerical model used in these calculations covered the 50 to 120 km region and could not be used much beyond $\chi = 80^{o}$ because the reflection height moves progressively upwards as the solar zenith angle increases. Nevertheless, it is apparent from the results given in Table 3 that the change in Al absorption due to HF heating should increase with solar zenith angle, as the upper ionosphere is more effectively heated. The limited experimental data available would seem to support these theoretical predictions since the sunset change in absorption at Puerto Rico given here ($\sim$ 9dB) is larger than that observed at Platteville during the daytime ($\sim$ 6dB) despite the fact that the Platteville heating power may have been larger (100 MW compared to 60 MW at Islote).

ACKNOWLEDGEMENTS

The authors wish to thank the staff of Arecibo Observatory for their assistance in the experimental phase of this research. Arecibo Observatory is operated by Cornell University under a contract with the National Science Foundation.

This study was supported by the National Science Foundation under Grant No. ATM 78-19964.

REFERENCES

Djuth, F. T., Preliminary heating results for September, 1980, Aerospace Corporation, Los Angeles, CA, November, 1980

Gnanalingam, S. and J.A. Kane, A study of electron density profiles in relation to ionization sources and ground based radio wave absorption measurements, Report No. X-625-73-355, Goddard Space Flight Center, 1973.

Ganguly, S., Nighttime and sunrise variations of absorption, Radio Science, 9, 1101-1108, 1974.

Meltz, G., L.H. Holway and N.M. Tomljanovich, Ionospheric heating by powerful radio waves, Radio Science, 9, 1049-1063, 1974.

Sulzer, M.P., A portable receiving system for use in the wave interaction experiments at Arecibo Observatory, PSU-IRL-IR-30, The Pennsylvania State University, 1973.

Sulzer, M.P., The detection of atmospheric gravity waves by means of the wave interaction experiment, PSU-IRL-SCI-461, The Pennsylvania State University, 1979.

Tomko, A.A., A.J. Ferraro, H.S. Lee and A.P. Mitra, A theoretical model of D-region chemistry modifications during high power radio wave heating, J. Atmos. Terr. Phys., 42, 275-285, 1980a.

Tomko, A.A., A.J. Ferraro and H.S. Lee, D-region absorption effects during high-power radio wave heating, Radio Sci., 15, 675-682, 1980b.

Trask, C., A high-gain vertical beam antenna array using orthogonal non-planar log-periodic structures in the backfire configuration, M.S. Thesis, The Pennsylvania State University, 1979.

Utlaut, W.F. and E.J. Violette, A summary of vertical incidence radio observations of ionospheric modification, Radio Science, 9, 895-903, 1974.

TABLE 1 SUMMARY OF ABSORPTION MEASUREMENTS

DATE	LOCAL TIME	SOLAR ZENITH ANGLE	HEATING FREQUENCY	HEATING WAVE POLARIZATION	HEATING TRANSMITTER POWER	EFFECTIVE RADIATED POWER (ESTIMATED)	AGC SAMPLE SPACING
8/29	18:05-19:35	86-106°	3.175 MHz	L-CR	75 KW x 4	60 MW	0.48s
8/30	16:10-19:00	60-100°	3.175 MHz	Varied	<75 KW x 2	<15 MW	0.48s
8/31	7:46-10:11	66-29°	3.175 MHz	LN	<75 KW x 2	<15 MW	0.48s
9/4	7:12-10:46	72-25°	3.175 MHz	L-CR	75 KW x 2	15 MW	0.22s

NOTE: L-CR = left circular polarization

LN = linear polarization

TABLE 2 ELECTRON DENSITY MODELS

h (km)	$\chi_s = 30^\circ$ n_e	$\chi_s = 60^\circ$ n_e	$\chi_s = 70^\circ$ n_e	$\chi_s = 80^\circ$ n_e
50	5.65×10^0	5.34×10^0	3.28×10^0	2.27×10^0
60	5.75×10^1	3.70×10^1	2.36×10^1	1.03×10^1
70	5.82×10^2	2.36×10^2	9.85×10^1	3.70×10^1
80	1.53×10^3	1.17×10^3	5.63×10^2	2.83×10^2
90	1.48×10^4	4.55×10^3	1.69×10^3	1.16×10^3
100	1.76×10^5	8.45×10^4	4.71×10^4	1.26×10^4
110	1.91×10^5	1.46×10^5	1.42×10^5	8.15×10^4

NOTE: χ_s = solar zenith angle. Electron densities in cm^{-3}.

TABLE 3 ELECTRON TEMPERATURE AND ABSORPTION MODELS

h (km)	$\chi_s = 30^\circ$ T_e/T_n	$\chi_s = 60^\circ$ T_e/T_n	$\chi_s = 70^\circ$ T_e/T_n	$\chi_s = 80^\circ$ T_e/T_n
50	1.10 (1.18)	1.10 (1.18)	1.10 (1.18)	1.10 (1.18)
60	1.51 (1.94)	1.51 (1.95)	1.51 (1.95)	1.52 (1.95)
70	2.33 (3.44)	2.92 (4.02)	3.09 (4.19)	3.16 (4.26)
80	1.09 (1.17)	1.41 (2.01)	3.60 (5.48)	5.19 (7.47)
90	1.02 (1.03)	1.09 (1.13)	1.38 (1.66)	2.18 (4.26)
100	1.00 (1.00)	1.00 (1.00)	1.09 (1.12)	1.33 (1.50)
110	1.00 (1.00)	1.00 (1.00)	1.00 (1.00)	1.02 (1.02)
ΔdB =	3.2 (5.0)	4.8 (7.4)	5.8 (8.5)	6.2 (9.0)
h_r =	95	99	102	109
h_{Te} =	69 (70)	72 (73)	76 (77)	81 (82)

NOTE: T_e/T_n = heated to unheated electron temperature ratio

ΔdB = change in 2.4 MHz absorption due to heating

h_r = approximate reflection height (level where plasma and angular wave frequencies are equal).

h_{Te} = height of peak in heated electron temperature distribution.

Values given in parentheses are for 120 MW of heating while the other values are for 60 MW.

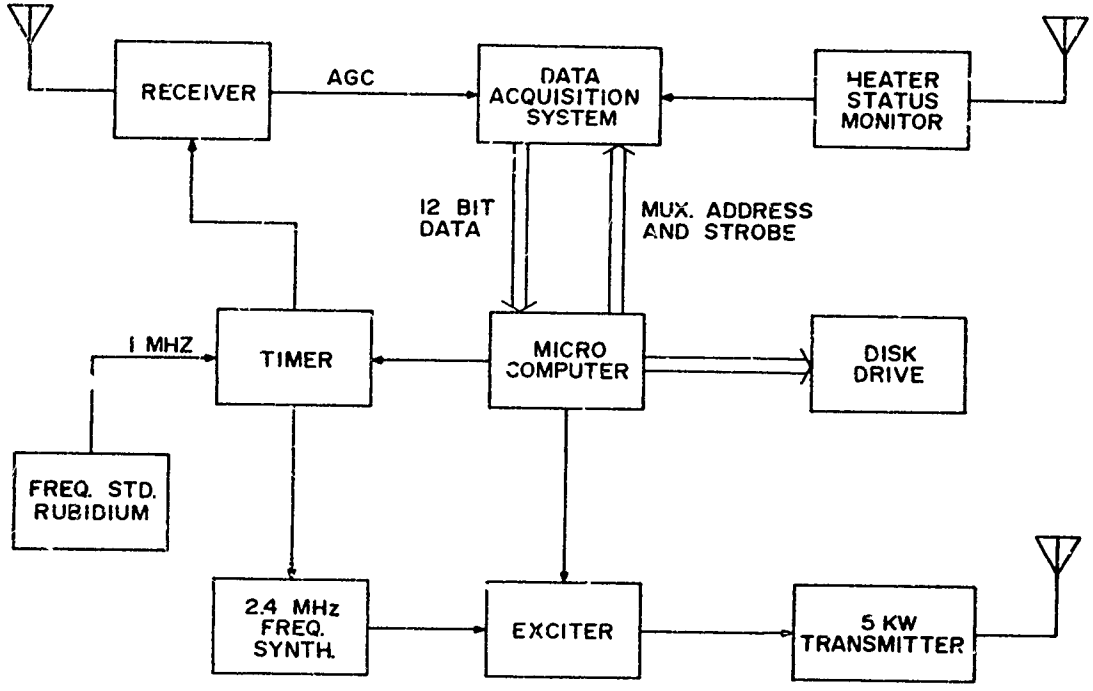


Figure 1: 2.4 MHZ sounding system.

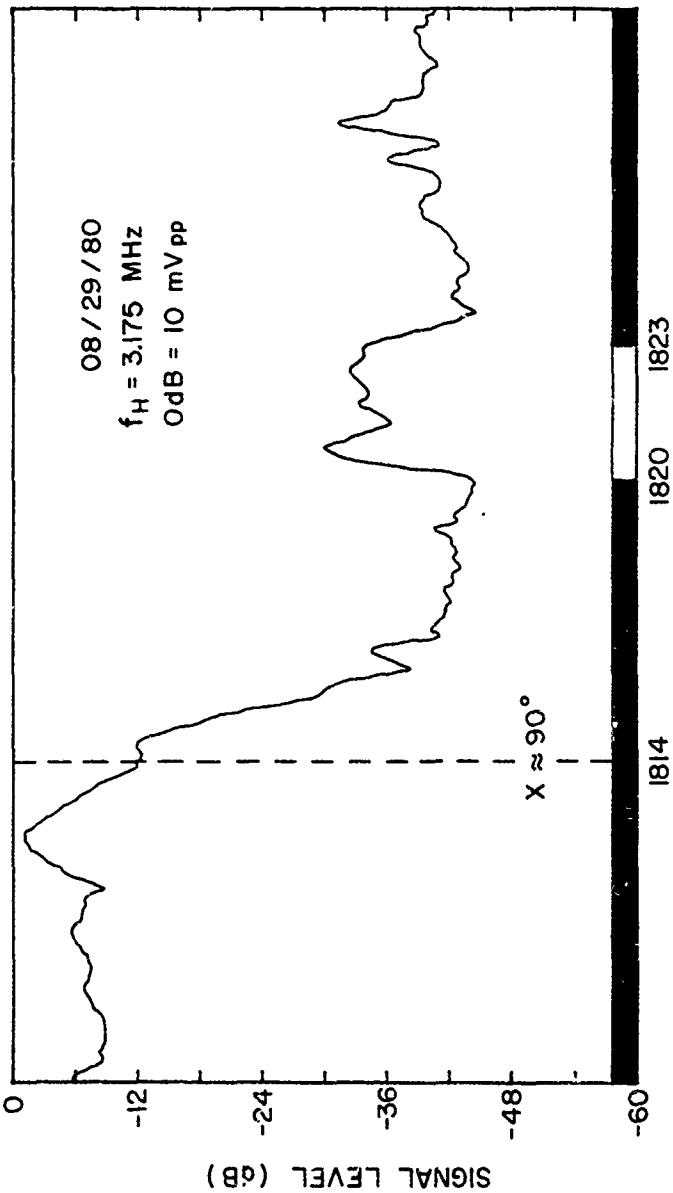


Figure 2: 2.4 MHz absorption during CW heating.

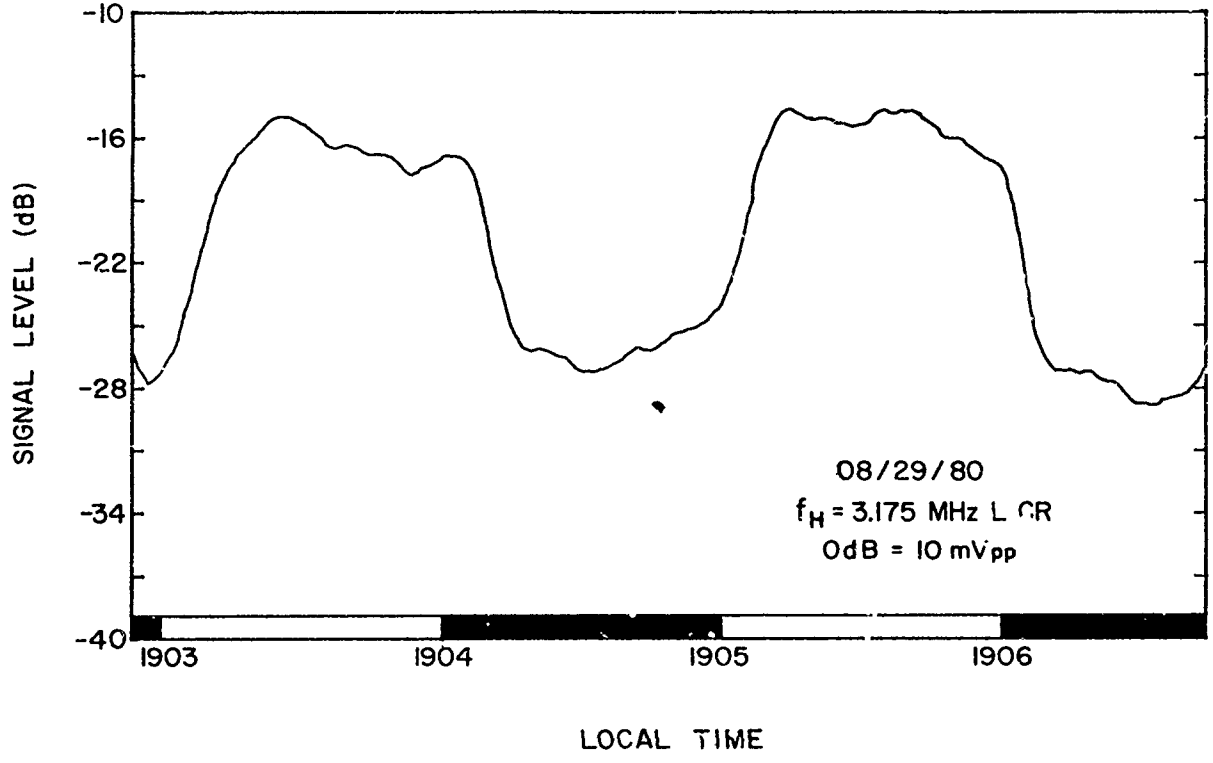


Figure 3: 2.4 MHz absorption during 1 minute heater on/off cycles.

Wireless Generation of ELF/VLF Radiation in the Ionosphere

C. L. Chang,* V. Tripathi,*,+ K. Papadopoulos,*,+
J. Fedder,++ P. J. Palmadesso,++ and S. L. Ossakow++

*Science Applications, Inc., McLean, VA 22102

+Permanent Address: University of Maryland,
College Park, MD 20742

++Geophysical & Plasma Dynamics Branch, Plasma Physics Division
Naval Research Laboratory, Washington, D.C. 20375

ABSTRACT

The ionospheric plasma may be used as a nonlinear active medium for the generation of radio waves in the ELF/VLF frequency range. A number of different concepts have been proposed which seek to avoid difficulties associated with large and inefficient ground or satellite based antennas by creating a "virtual" antenna in the ionospheric plasma. There are two major approaches to the problem of creating a radiating current pattern in the ionospheric plasma without the use of solid conductors. The first approach involves the use of ground based pulsed r.f. transmitters or "heaters" to generate a periodic variation in the ionospheric conductivity in regions where strong natural current systems exist. The result is an A.C. component of current which radiates at the desired frequency. Modulation of the parallel conductivity in the auroral electrojet, for example, leads to the formation of a closed circular Hall current pattern which acts like a vertical magnetic dipole antenna. Modulation of the Pedersen or Hall conductivities in this region produces a double vortex current pattern with a strong linear component in the center, which acts like a horizontal electric dipole radiator. Studies are now being conducted to determine the optimum altitude for modulation, power requirements, etc. Preliminary results suggest that systems of this type should be feasible. The second approach to the creation of a virtual antenna involves the generation and mode coupling of two higher frequency waves to produce the ELF/VLF signal. In simplest form, this method involves the transfer of energy from two waves $(\omega_o, \underline{k}_o)$ and $(\omega_1, \underline{k}_1)$ to a low frequency third wave $(\omega, \underline{k})$ which satisfies the condition $\omega = \omega_o - \omega_1 << \omega_o$, ω_1 and $\underline{k} = \underline{k}_o - \underline{k}_1$ in some region of the ionosphere. Thus one can use efficient high frequency transmitters radiating either from the earth or from satellites to produce low frequency radiation. Many variations of this scheme are possible, depending on the choice of high frequency modes, excitation geometry, etc. Again, results of preliminary studies of systems of this type are promising. While some of these ideas have been suggested previously, considerable additional research is needed in order to establish an adequate theoretical base for experiment design and for assessment of the feasibility of these concepts for communications.

INTRODUCTION

A key Navy communications channel, especially for submarine communications, utilizes VLF and ELF frequencies. The technical and political problems associated with direct generation of ELF waves by ground based large antenna arrays[1] (i.e., Sanguine and Seafarer projects) have been recently reviewed by Starkey.[2] However, recent progress in theoretical and computational plasma physics,[3] allows us to look confidently to the possibility of ELF generation and coupling in the earth ionosphere waveguide, by utilizing the ionospheric plasma as an active medium.[4] The basic idea involves the formation of a virtual ELF antenna in the ionospheric plasma, using modulated HF ionospheric heaters.[4] The resulting interactions are mainly nonlinear and go much beyond the usual Luxemberg effect type ideas. Conceptually, a variety of possible schemes could arise, depending on the use of particular nonlinear interactions and HF heaters based on the ground or space. In this paper we present preliminary results, with respect to two potential schemes. The first one relies on modulating pre-existing ionospheric currents, such as the auroral or equatorial electrojet, at ELF frequency, by modulating the conductivity of the ionospheric plasma through which they flow. The second relies on modified stimulated forward Brillouin scattering, of two HF modes off a low frequency compressible Alfven mode.[4,5] The excited Alfven ionospheric eigenmode, couples in the

earth ionosphere waveguide and the horizontal part of its electric field penetrates into the water. It should be stressed that a major advantage of exciting the waveguide from the ionosphere rather than the ground is the fact that the relevant wavelengths in the dielectric are 10^3 smaller than in free space due to the high refractive index. We proceed below to discuss the two concepts separately.

CURRENT MODULATION

The first approach involves using a powerful ground antenna to modulate the ambient electron temperature T_e at a low frequency. The T_e variation changes the electron-neutral collision rate ν and indirectly modulates the ambient plasma conductivities through their ν dependence, i.e.,

$$\begin{bmatrix} \sigma_{Pe} \\ \sigma_{He} \\ \sigma_{\|e} \end{bmatrix} = \frac{N_e e^2}{m_e} \begin{bmatrix} \nu/(\Omega^2 + \nu^2) \\ \Omega/(\Omega^2 + \nu^2) \\ 1/\nu \end{bmatrix} \tag{1}$$

where Ω is the electron gyrofrequency and σ_{Pe}, σ_{He}, $\sigma_{\|e}$ are the Pedersen, Hall and parallel conductivities of the electrons, respectively. Any ionospheric current, therefore, oscillates in the modified region at the specified frequency. The current perturbation serves as a "virtual" antenna to radiate electromagnetic waves. In estimating the ELF efficiency of the concept, the following computations should be performed: (a) Determination of the ionospheric region where T_e relaxes to its ambient value at a rate τ^{-1} much faster than the ELF radiation frequency ω (i.e., $\omega\tau << 1$); (b) the energy deposition altitude profiles of the HF power, and the resulting modifications in the T_e and the electron conductivity profiles; (c) the modulated current patterns self consistently created by the conductivity modulation; and (d) the resulting antenna strength as a function of the frequency and the HF radiated power. We have developed a series of coupled numerical schemes to achieve the above tasks. We present below some very preliminary results, for a typical situation of an HF antenna with ERP (ERP = power x antenna gain) 100 MW, modulated at 100 Hz, for auroral electrojet parameters. We again caution the reader that the results are very preliminary and should be used with caution.

The answer to the first question can be found by the following considerations. The large heat capacity of the neutrals does not allow any significant ion temperature change. This restricts the interesting altitudes to regions where the electron conductivity dominates. According to Fig. 1, which plots the ambient conductivity profiles versus the altitude, this region lies below 100 km in altitude. Another important factor in determining the optimal heating region arises from the relaxation process of electron energy. Specifically, in order to modulate the electron temperature at ELF, T_e must relax to its ambient value in a time scale shorter than the modulation period. Therefore, we are restricted to conduct the experiment in a region where $\delta\nu > f$, with δ the average fraction of energy lost in each collision ($\delta \sim 10^{-3}$) and f the modulation frequency. This additional requirement further limits the interesting region to under 90 km in altitude.

In computing the conductivity modifications the electron temperature T_e (τ) is given by

$$\frac{dT_e}{dt} = \left(\frac{2}{3N_e}\right) \frac{Q_o K(z)}{4\pi z^2} \exp\left[-2\int_o^z K(z')dz'\right] - \delta\nu_e(T_e - T_{eo}) \tag{2}$$

where Q_o is the effective radiation power (ERP). The absorption coefficient K(z) corresponds to the inverse of the local wave attenuation length and, for the case of parallel field propagation, can be expressed as

$$K(z) \simeq \frac{\omega_e^2}{2c} \frac{\nu}{[\nu^2 + (\omega_h \mp \Omega)^2]} \tag{3}$$

where the + and - signs correspond to the polarization of the ordinary (O) and extra-ordinary (X) mode, and ω_e is the plasma frequency. We note that at lower ionospheric heights the electron heat conduction is negligible and the electron density variations are due to changes in the recombination rate which are also negligible. Equations (2) and (3) constitute the basic heating equations and can be solved by numerical integration. Heating curves are presented in Fig. 2 for the typical case; we plot the electron temperature T_e as a function of altitude after 1 msec of X and O mode heating. Most efficient absorption is ensured by selecting the frequency at $\omega_h \sim \Omega$ which maximizes the absorption coefficient in (3). One important feature of this figure is that the wave energy of the X mode gets absorbed rapidly at 70 km with little heating observed above 78 km. This can be understood as follows. At $\omega_h \simeq \Omega$, the absorption coefficient has the simple relation $K(z) \propto N_e(z)/\nu(z)$. A sharp decrease in ν from 60 to 70 km leads to a marked increase in K(z) and suppresses dissipation by reducing the background neutral collisions. This creates profound peaks in the temperature profile. At a higher heating frequency $\omega_h \simeq 9\Omega$, as shown in the same figure, the absorption coefficient $K(z) \propto N_e(z)\,\nu(z)/\omega_h^2$ decreases gradually with height and the heated T_e profile exhibits no peak. This suggests that the heating characteristics are sensitive to both wave frequency and

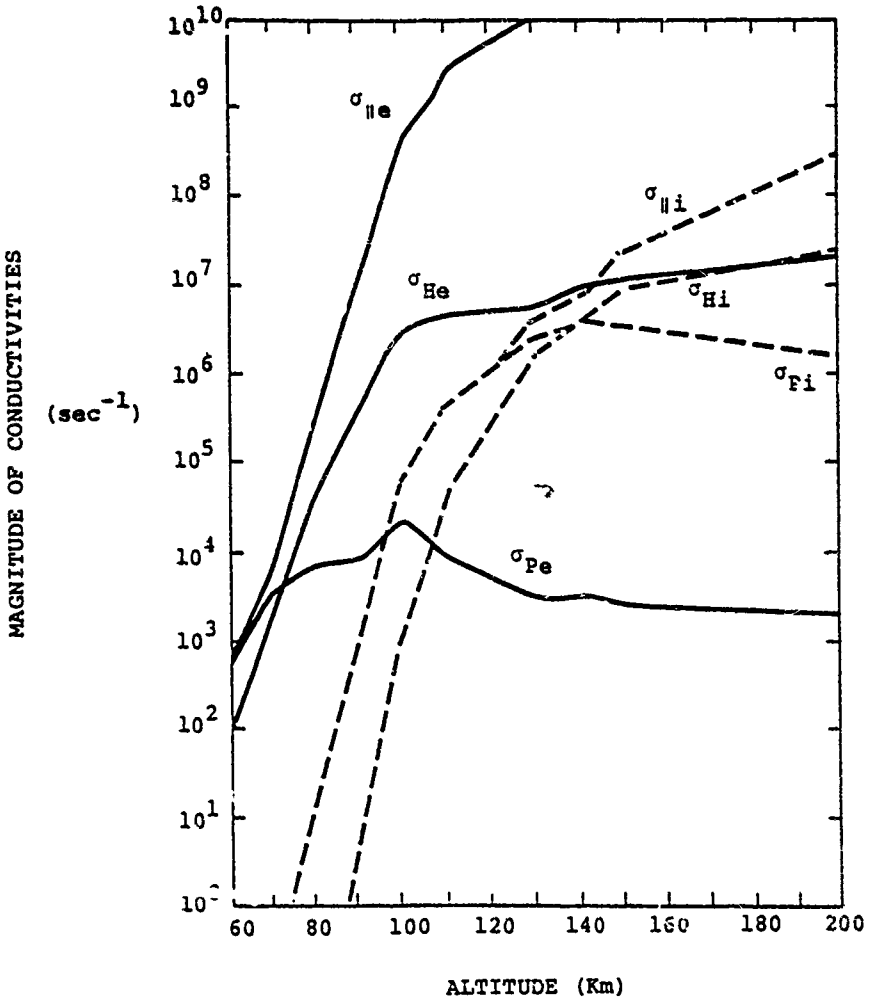


Figure 1. Electron and ion conductivity profiles as functions of altitude z.

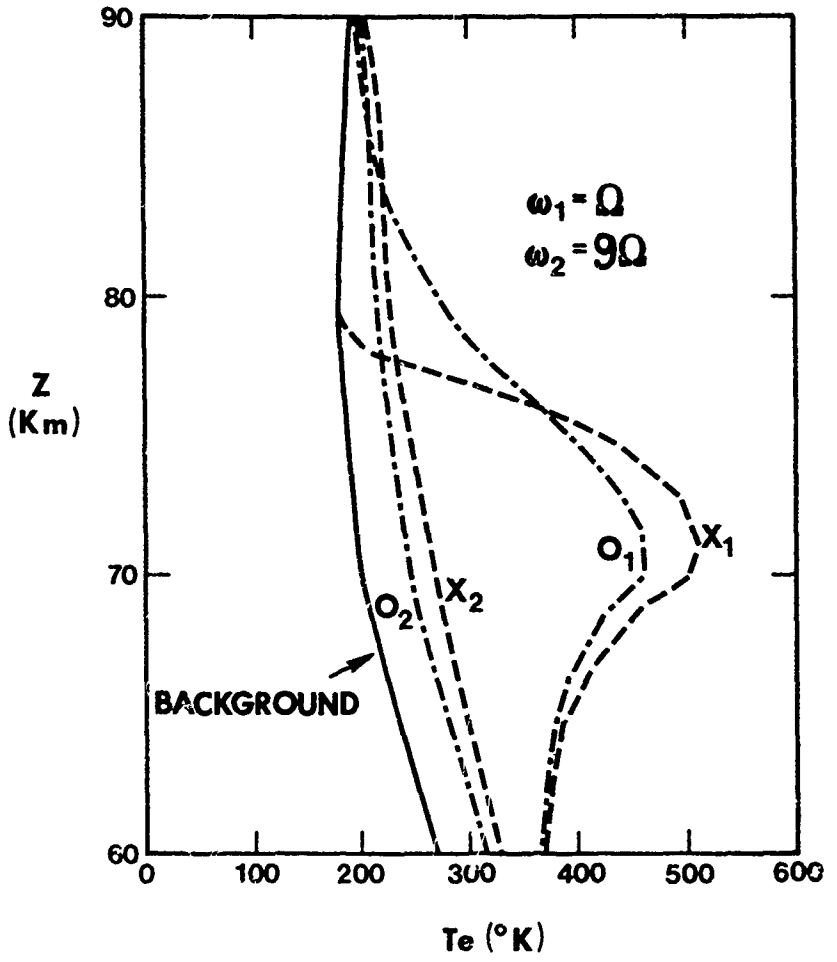


Figure 2. Modified electron temperature profiles $T_e(z)$ after 1 msec of heating at two different wave frequencies. X and O represent the extraordinary and ordinary modes of propagation.

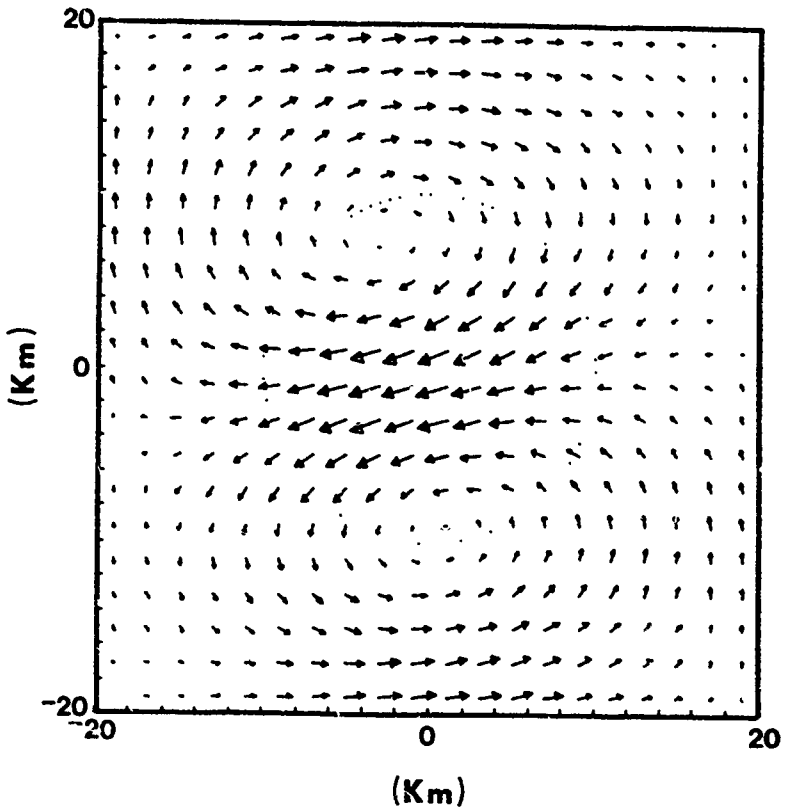


Figure 3. Illustration of the perturbed current pattern. The modified region is specified by the dotted circle with a radius of 10 km. The ambient electric field in this case is taken to be E_0 = 25 mV/m in the x direction and the magnetic field points into the paper.

background neutral and electron concentrations. In Fig. 2, the percentage changes of total conductivities at the temperature peak (for the X mode) are estimated to be -35% for σ_P and -70% for σ_H.

To simplify the description of the ionospheric current modulation, we reduce the problem to two dimensions by introducing the height-integrated conductivities Σ_P and Σ_H. The height-integrated current density is then

$$\underline{J} = \underline{\underline{\Sigma}} \cdot \underline{E}$$

where

$$\underline{\underline{\Sigma}}(x,y) = \begin{bmatrix} \Sigma_P & \Sigma_H \\ -\Sigma_H & \Sigma_P \end{bmatrix} \tag{4}$$

is the conductivity tensor. The total electric field $\underline{E}$ is the sum of an ambient field $\underline{E}_o$ and a perturbation $\underline{E}_1 = -\nabla\phi$, with $\underline{E}_o$ pointing in the x direction. From the current continuity equation $\nabla\cdot\underline{J} = 0$, we obtain

$$\frac{\partial}{\partial x}\left(\Sigma_P \frac{\partial\phi}{\partial x}\right) + \frac{\partial}{\partial y}\left(\Sigma_P \frac{\partial\phi}{\partial y}\right) + \left(\frac{\partial\Sigma_H}{\partial x}\right)\left(\frac{\partial\phi}{\partial y}\right) - \left(\frac{\partial\Sigma_H}{\partial y}\right)\left(\frac{\partial\phi}{\partial x}\right) = E_o \frac{\partial\Sigma_P}{\partial x} - E_o \frac{\partial\Sigma_H}{\partial y} \tag{5}$$

This equation is solved numerically for $\phi(x,y)$ with a given set of Σ_P and Σ_H profiles. Knowing ϕ, one can calculate the current perturbation $\underline{J}_1$ from the expression

$$\underline{J}_1 = -\underline{\underline{\Sigma}} \cdot \nabla\phi \tag{6}$$

An illustration of this current pattern is presented in Fig. 3.

From this figure we can see that the current perturbations form a double vortex flow pattern with a strong linear component in the center. This component is in the opposite direction of the ambient Pedersen current and is located inside the main lobe of the conductivity profile. As seen from Fig. 3 the flow pattern corresponds to an equivalent horizontal electric dipole radiation source. The strength of the dipole can be determined, by numerically integrating over the distribution $J_1(x,y)$. The resulting dipole moment for ionospheric electric fields of the order $E = 25$ mV/m, is on the order of 10^4 Am. Using Wait's formula,[6] we can estimate the ELF field on the ground as $E \approx 10^{-6}$V/m and $H \approx 10^{-7}$ A/m indicative of interesting signal to noise ratios.

PARAMETRIC EXCITATION

The parametric excitation of ELF, does not rely on the presence of ionospheric currents, providing us with greater freedom in site selection and relative independence from the ionospheric state. The particular nonlinear mechanism considered is the parametric decay of an HF pump radiowave (ω_o, k_o), where $\omega_o >> \omega_e$, into a low frequency (ω,k) compressional Alfven wave and a high frequency sideband $\omega_1 = \omega_o - \omega$, $\underline{k}_1 = \underline{k}_o - \underline{k}$, which is also directly driven from the ground. The double resonance excitation is advantageous over the single parametric excitation for ELF communications, because it tunes the desired ELF frequency and reduces the instability threshold. Since for ELF frequencies $\omega << \omega_o$, $|\underline{k}| << \underline{k}_o$, the process is analogous to forward stimulated scattering off Alfven waves. In order to determine the coupling efficiency and the instability threshold, we compute first the ponderomotive force due to the waves $(\omega_o,\underline{k}_o)$ and $(\omega_1,\underline{k}_1)$ and from this the growth rate γ_o for a homogeneous plasma. Following the derivation of Papadopoulos et al.,[5] we find

$$\gamma_o = \frac{k_x^2 V_o^2 \omega_i}{16\,\omega\omega_o} \tag{7}$$

where k_x is the wave-vector in the direction perpendicular to both $\underline{E}_o$ and the ambient magnetic field, ω_i is the ion plasma frequency and $V_o^2 = e^2E_o^2/m_e^2\omega_o^2$. As an order of magnitude result $\gamma_o \approx 10^2\text{sec}^{-1}$, for a 10 MW, 5 MHz transmitter with a $G = 10^3$ at ionospheric altitudes of 120 km.

Due to the existence of the vertical density profile and the finite region of the interaction, the wave resonance conditions are satisfied only locally. We should therefore examine the boundary value inhomogeneous problem. Using the Rosenbluth techniques,[7] and taking the x axis as the vertical (parallel to ∇n) and z as the horizontal, we find that in the WKB approximation the mode coupled equations read[5]

$$\frac{\partial\epsilon}{\partial t} + \underline{V}_g \cdot \frac{\partial\epsilon}{\partial\underline{x}} = \gamma_o(z)\,\epsilon_1 \exp\left(-\frac{ik' x^2}{2}\right) \tag{8}$$

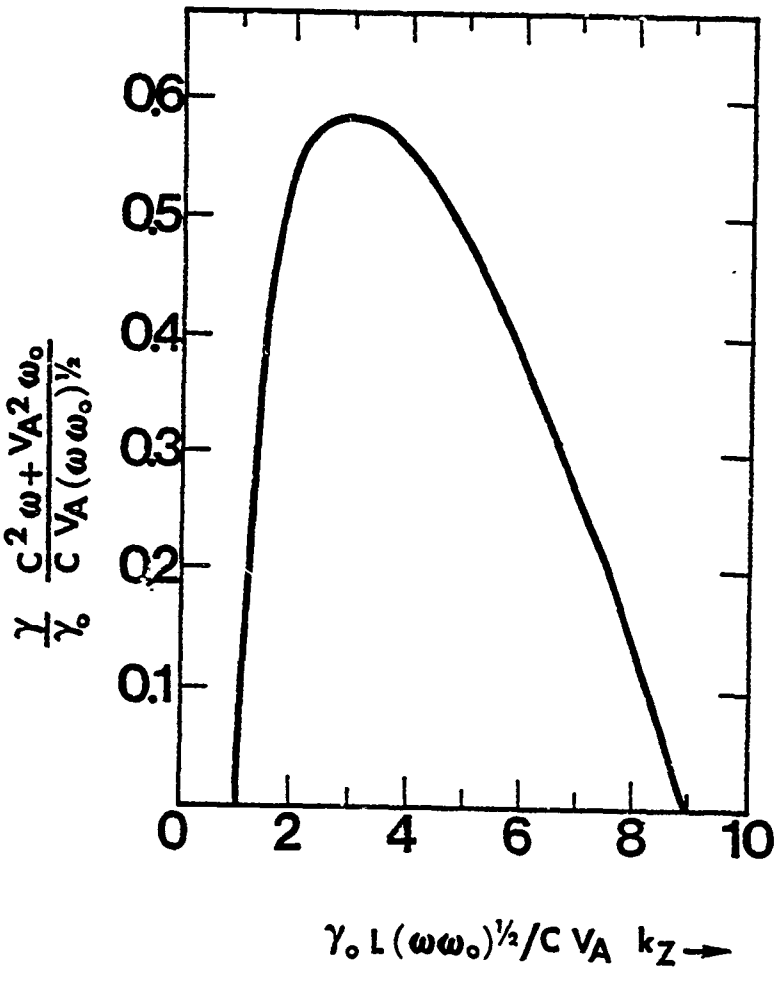


Figure 4. Growth rate of the fastest growing Alfven wave for a gaussian pump in an inhomogeneous plasma. The parameters are: rf power density ~ 0.5 W/m^2, $\omega_o \sim 5$ MHz, $\omega \sim 10^2$, $\Omega_i \sim 2\times10^2$rad. sec^{-1}, $c/V_A \sim 10^3$, $L \sim 3$ km.

$$\frac{\partial \epsilon_1}{\partial t} + \underline{V}_{g1} \cdot \frac{\partial \epsilon_1}{\partial \underline{x}} = \gamma_o(z)\epsilon \exp\left(\frac{ik' x^2}{2}\right) \tag{9}$$

where $\epsilon_1 = E_1 V_A/c$, $\epsilon = E_y(\omega_o/\omega)^{\frac{1}{2}}$, $\underline{V}_g = V_A^2\, \underline{k}/\omega$, $\underline{V}_{g1} = c^2\,(\underline{k}-\underline{k}_o/\omega)$, and V_A is the local Alfven speed. Details of the numerical solution of the above Eqs. can be found in Papadopoulos et al.[5] We restrict ourselves here to some of the key points. For gaussian pump profile e^{-x^2/L^2}, the growth rate as a function of the extent of the pump is shown in Fig. 4. This was computed for the particular case where the parameter

$$\lambda = \frac{\pi \gamma_o^2}{V_{gx} V_{gx_1}} \left(\frac{dk}{dx}\right)^{-1} \tag{10}$$

was $\lambda = 2$ (it corresponds to HF power density .5 W/m^2 at 5 MHz, for $\omega = 10^2$ at 120 km altitude). It is obvious from Fig. 4 that the instability is absolute for values of $\gamma_o L(\omega\omega_o)^{\frac{1}{2}}/k_z c V_A$ between 1 and 9. The growth rate maximizes at

$$\gamma_{max} \approx 0.2\gamma_o \tag{11}$$

This condition can be achieved by a 20 MW transmitter having a large gain $G \approx 10^4$ (i.e., $L \approx 3$ km). For more details on this scheme we refer the interested reader to Papadopoulos et al.[5] A determination of the amplitude of the excited waves requires a nonlinear theory and is presently under consideration. The theoretical upper limit will be given by the Manley-Rowe relations, which will give power levels in the source region of the order of (ω/ω_o) ERP $\approx 10^3$ watts. We are therefore aiming at only a few percent of the theoretical limit.

SUMMARY AND CONCLUSIONS

We have presented here some preliminary results indicating that with present day technology ELF waves at powers of interest for submarine communications, could be generated. In the current analysis only the simplest approach has been examined. Effects such as the three dimensionality of the currents, the role of the ponderomotive force on local density profile modifications, anomalous HF absorption and many others are presently under study and will be reported in time.

At this point we should comment on the experimental stage of the ideas. A number of results have appeared in the literature, following the initial suggestion by Papadopoulos[4] of their potential interest for Navy communications. The most extensive work has been done in the USSR; however, very few results appeared in the literature. The most interesting result was reported by Kapustin et al.,[8] which showed 2 kHz generation from 5 MHz modulated pumps. The opening of the Tromso, Norway, heating facility last year by the Max Planck Institute, offers the possibility of testing some of these ideas.[9] Actually, Kopka and Stubbe (private communication, 1980) have observed, on the ground near the antenna site, low frequency waves (500 Hz - 2 kHz) when modulating auroral electrojet currents at the appropriate frequency. They also detected micropulsations. Finally, we should mention an ongoing effort at the Pennsylvania State University, which is still at a very preliminary stage.[10] In summary, the experimental results are very rare, not well documented and not properly guided by theoretical planning to produce conclusive evidence of the particular interactions occurring and their potential optimization. The preliminary evidence is, however, positive.

In concluding we should mention that a well coordinated experimental and theoretical effort is urgently needed in this area. The design and optimization aspects require extensive theoretic nonlinear plasma physics input, and cannot be achieved by random experimentation. From the other side, idle theorizing without experimental guidance is apt to be an equally futile exercise, as far as practical systems are concerned.

ACKNOWLEDGEMENT

One of the authors (K.P.) greatly appreciates many discussions with Drs. T. Coffey, P. Stubbe, H. Kopka, A. Drobot, R. Sharma, and P. Sprangle. This work was supported at SAI by NRL Contract No. N00173-80-C-0194 and at NRL by ONR and NAVELEX.

REFERENCES

1. S. L. Bernstein, M. L. Burrows, J. E. Evans, A. S. Griffiths, D. A. McNeill, C. W. Niessen, I. Richer, D. P. White, D. K. Willim, "Long Range Communications at ELF," M.I.T. Lincoln Laboratory Report, 1973.

2. R. Starkey, Military Electronics/Countermeasures, p. 26, Jan. 1981.

3. C. S. Liu, "Parametric Instabilities in Inhomogeneous Unmagnetized Plasmas," Advances in Plasma Physics, Volume 6, pg. 83, 1976.

4. K. Papadopoulos, "Nonlinear Production of ELF Waves by Two High Frequency Waves," March 1975 (NRL Plasma Physics Division, internal report).

5. K. Papadopoulos, R. R. Sharma, and V. K. Tripathi, SAI Report 81-343-WA (1981). (Phys. Rev. Lett., communicated).

6. J. Wait, Electromagnetic Waves in Stratified Media, Pergamon Press, 1962.

7. M. Rosenbluth, Phys. Rev. Lett., 29, 565, 1972.

8. I. N. Kapustin, et al., JETP Lett., 24, 228, 1977.

9. P. Stubbe and H. Kopka, J. Geophys. Res., 82, 2319, 1977.

10. A. J. Ferraro and H. S. Lee, Penn. State University Report PSU-IRL-FR-79/1 (1979).

IONOSPHERIC HEATING EFFECTS ON LF PROPAGATION

FRANCIS J. KELLY
E.O. HULBURT CENTER FOR SPACE RESEARCH
NAVAL RESEARCH LABORATORY
WASHINGTON, D.C. 20375

D. N. MADHUSUDHANA RAO
ANDHRA UNIVERSITY
WALTAIR, INDIA

ABSTRACT

During an ionospheric modification experiment on September 13, 1980, the Arecibo Observatory Heater was operating in an on/off mode with a one minute period for each. A U.S. Navy low frequency transmitter at Isabella, Puerto Rico emitted a continuous radio wave at 123.7 KHz. Receivers at Vega Alta, Puerto Rico observed, at times, very strong (6dB or more) enhancements of the longitudinal magnetic component of the received LF radio wave, when the heater changed state from off to on or reverse. The greatest relative enhancements occurred when the time-varying LF ionospheric signal was weakest. A possible explanation of this observation may be based on the concept of interference because of multipath propagation within the ionosphere.

Introduction

Previous HF ionospheric heating studies have not shown a very strong heater effect on low frequency (LF) electromagnetic waves reflected from the lower ionosphere [Reference 1, 2, 3]. However, in a recent experiment the action of an ionospheric heater at the Arecibo Observatory* twice immediately changed the amplitude of a radio wave (123.7 KHz) by up to 10 dB. This result indicates that the strength of reflected LF waves can become extremely sensitive to the state of the ionosphere. Possibly, the observed [4,5] extreme variability of other long distance VLF propagation paths has a similar cause.

Experiment

On the night of September 13, 1980 the Arecibo heater was illuminating the overhead ionosphere at a frequency of 3.2 MHz with a repetitive one minute on, one minute off schedule. Continuous LF (123.7 KHz) transmissions from the U.S. Navy broadcast station at Isabella, Puerto Rico were received at Vega Alta, Puerto Rico on two orthogonal loop antennas (Figure 1). Wideband tape recordings and narrowband real-time amplitude measurements were made. Because the distance between the LF transmitter and the receiver was only 80 kilometers, the field measured by the transverse loop was dominated by the groundwave and was relatively constant. The longitudinal loop was insensitive to the groundwave and measured the ionospherically reflected fields, which varied over a range of about 20 dB over periods of several minutes. The basic cause of this field strength variation is undoubtedly slow changes in the ionosphere possibly caused by the motion of "patches" of electrons and ions in the D-region. However, the magnitude of the variation (15dB) was surprising. Figure 2 is a segment of the strip chart data gathered at this time. The on and off times of the heater are indicated by labels. The events of interest as shown at times when very rapid and abrupt increases of field strength coincided with heater transitions from on to off and from off to on. The validity and synchronism of this effect was confirmed by playing back the tape recorded data containing accurate time, a wideband replica of the LF signal, and a recording of the AGC of an HF receiver tuned to the heater signal. The effects were not accidental coincidences with bursts of wideband atmospheric noise; rather, the ionospherically reflected wave itself rapidly increased at both instants. When the LF signal was weak (20 dB below normal), the heater induced modification in the ionosphere increased the signal by 10 dB.

*The Arecibo Observatory is part of the National Astronomy and Ionosphere Center, which is operated by Cornell University under contract with the National Science Foundation.

Discussion

The behavior of the LF waves reminds one of a multipath interference situation in which several propagating waves cancel one another at a receiver location. When the cancellation is very complete, only a small change in either propagation path is needed to break the exact out-of-phase relationship between the waves and produce a large relative increase of field strength. To see whether such an explanation is tenable, ionospheric reflection and transmission coefficients were calculated as a function of frequency for the Puerto Rico experimental conditions using a nighttime ionospheric profile (Fig. 3) taken from Galejs [6]. The full-wave program develped by Pitteway and Smith [7] produced the reflection and transmission coefficients shown on Figures 4 and 5. The rapid oscillations of the reflection coefficients with frequency add weight to the idea that a multipath fading phenomenon was at work. However, until controlled experiments over a wide LF bandwidth can be performed, we cannot completely eliminate the unlikely possibility that some wideband absorptive phenomenon is at work in the Puerto Rican ionosphere and that the wideband phenomenon was modified by the action of the heater.

Acknowledgement

The authors wish to thank UNESCO for a study grant to Dr. Rao for work at the Naval Research Laboratory. We also wish to thank L. Quinn, L. Harnish, F. Djuth, W.E. Gordon, the staff of the Arecibo Observatory, and the personnel of Naval Receivers Facility, Sabana Seca, Puerto Rico for their assistance and cooperation in this work. We thank especially Mr. Martinez of Vega Alta, Puerto Rico for his generous hospitality during this experiment.

References

1. F. W. Perkins and R. G. Roble, "Ionospheric Heating by Radio Waves: Predictions for Arecibo and the Satellite Power Station", Journal of Geophys Res., Apr. 1, 1978, 83, 1161-1123 (No. A4).

2. T. B. Jones, K. Davies, B. Weider, "Observation of D-Region Modifications at Low and Very Low Frequencies", Nature 238, July 7, 1972.

3. C. M. Rush, E. J. Violette, R. H. Espeland, J. C. Carroll and K. C. Allen, "Impact of SPS Heating on VLF, LF and MF Telecommunications Systems Ascertained by Experimental Means", U.S. Dept. of Commerce, NTIA-R-80-37 (April 1980).

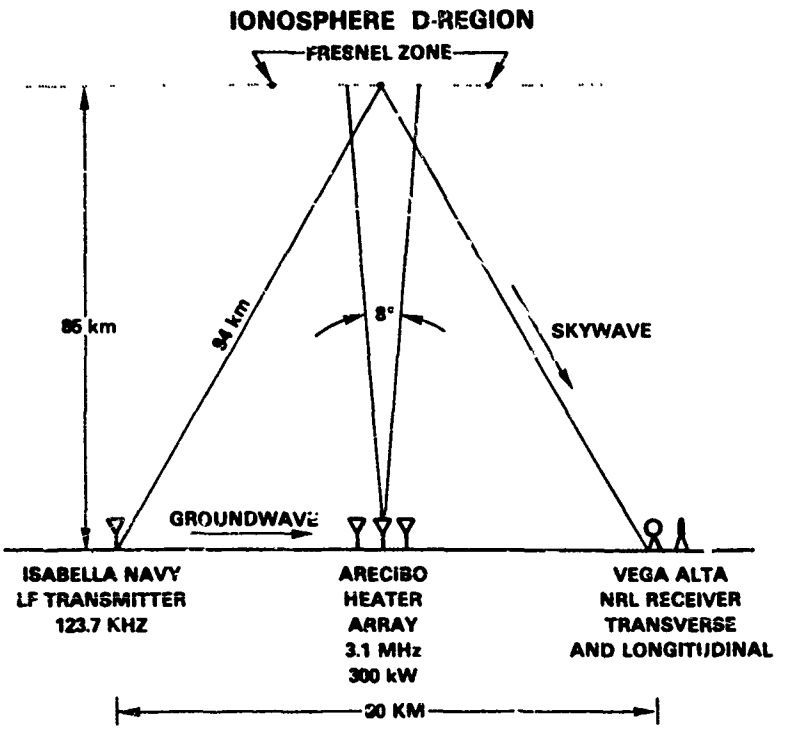
4. Tolstoy and F. Reder - (Private Communication)

5. J. E. Bickel, T. A. Ferguson and G. V. Stanley, "Experimental Observation of Magnetic Field Effects on VLF Propagation at Night", Radio Science, 5, pp. 19-25 (1970).

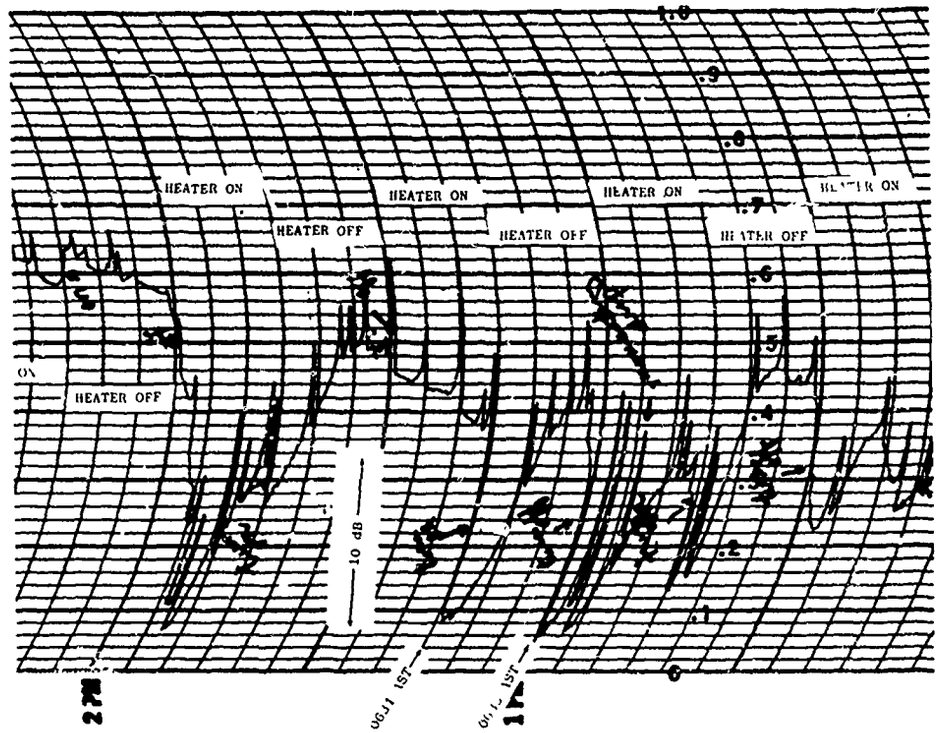
6. J. Galejs, "Terrestrial Propagation of Long Electromagnetic Waves," Pergamon Press (Oxford) 1972.

7. M.L.V. Pitteway, "The Numerical Calculation of Wave-Fields, Reflection Coefficients and Polarizations for Long Radio Waves in the Lower Ionosphere," Phil. Trans. R. Soc. A257, 219 (1965) and G.H. Smith and M.L.V. Pitteway, "Fortran Program for Obtaining Wavefields of Penetrating, Nonpenetrating, and Whistler Modes of Radio Waves in the Ionosphere," in ELF-VLF Radio Wave Propagation" edited by J. Holtet, D. Reidel Publishing Company (Dordrecht, Holland) 1974.

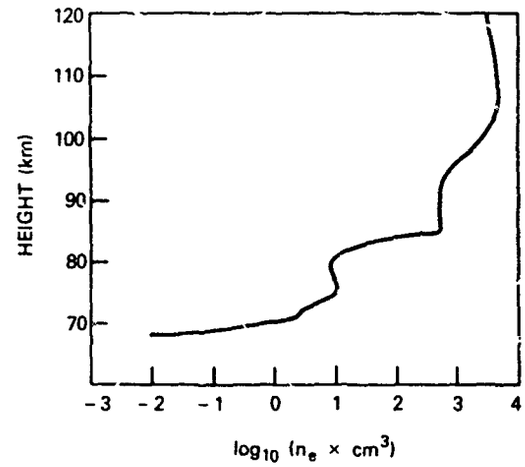
8. J.S. Belrose, "The Oblique Reflection of Low Frequency Radio Waves from the Ionosphere", pp. 149-168 in "Propagation of Radio Waves at Frequencies Below 300 Kc/s" edited by W.T. Blackband, The Macmillan Company (New York) 1964.



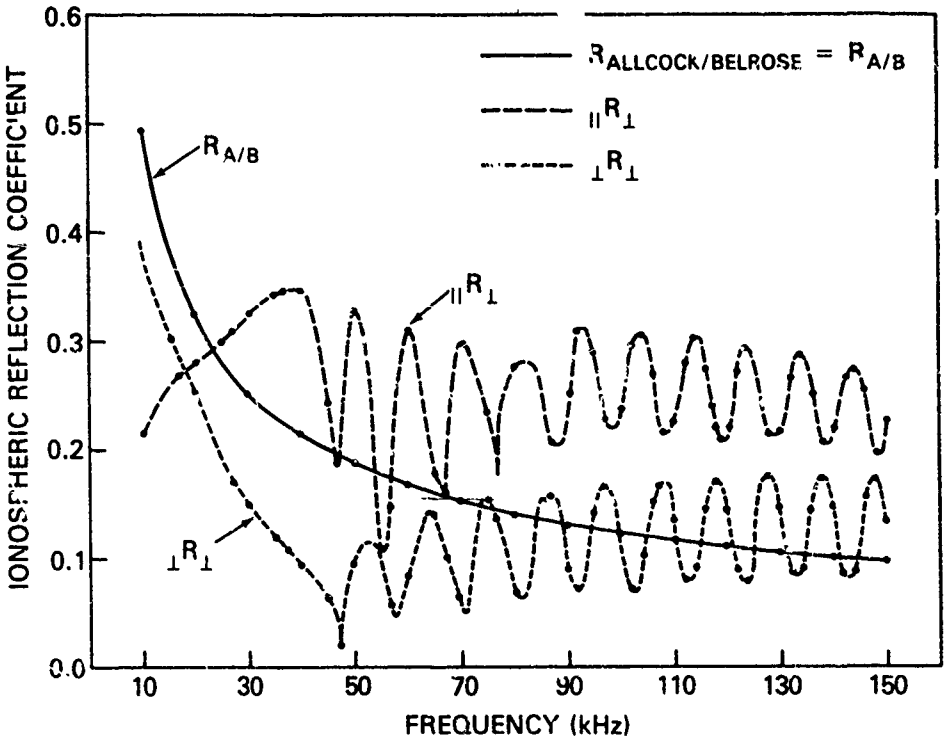
1. Schematic diagram showing the geometry of the experiment.



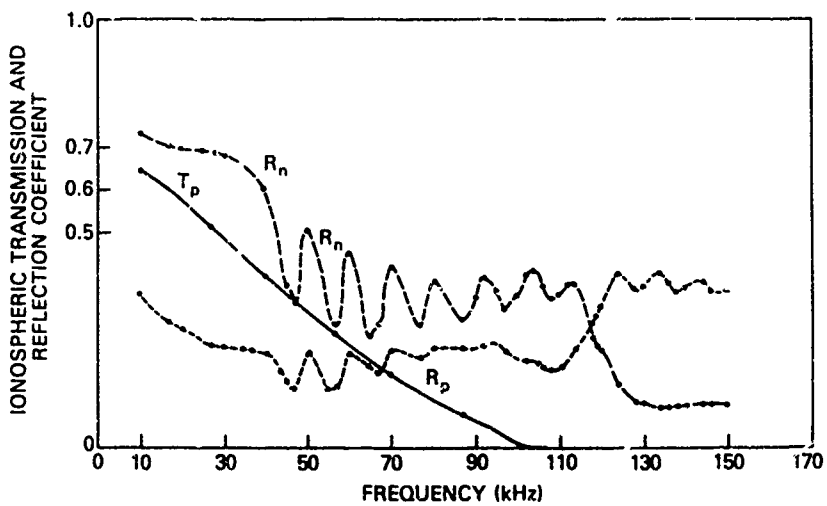
2. Strip chart data showing field strength of LF waves on longitudinal loop antenna. On and off periods of the heater are indicated by the shading.



3. Electron density profile for nighttime calculations.



4. Graph of reflection coefficients versus frequency. The $R_{A/B}$ curve is a plot of an empirical reflection coefficient approximation given by Belrose [6] using an equivalent frequency concept of Allcock. The $_{\parallel}R_{\perp}$ and $_{\perp}R_{\perp}$ reflection coefficients follow the convention of [7].



5. Graph of reflection and transmission coefficients versus frequency. Definitions of R_n, R_p and T_p are given in [7]. Data points are indicated because not enough points were obtained to completely define the curves.

Project HIPAS - High Power Active Stimulation of the Auroral Plasma

A. Y. Wong, J. Roederer*, R. Hunsucker*,
G. Dimonte, P. Viezbicke, J. Gregorio and J. Santoru
University of California, Los Angeles, CA 90024
*University of Alaska, Fairbanks, Alaska 99701

ABSTRACT

Laboratory experiments in which high power, pulsed electromagnetic waves interact with an inhomogeneous plasma indicate that the generated nonlinear plasma phenomena depend on peak incident power and not on pulse length. The electromagnetic waves can penetrate beyond the cutoff and produce large, enhanced electrostatic fields at the critical layer within 100 electron plasma periods. The enhanced electric field pressure can be comparable to the thermal pressure and can accelerate ions and electrons to velocities much greater than their thermal speed. Large density cavities (with $\delta n/n \gtrsim 10\%$) can be created in a time shorter than the usual ion response time because of the accelerated ion dynamics. These laboratory results have been extended to create a new and generalized concept to actively stimulate space plasmas with high power RF pulses of short duration. Recent observations at Arecibo Observatory support such short-pulse operations. A pulsed, high power HF radio wave transmitting system will be used for the stimulation of the auroral ionospheric plasma. The HIPAS field system, to be located near Fairbanks, Alaska, consists of a 2 MW transmitter and a crossed dipole antenna array. The transmitter consists of a low power oscillator and a series of amplifiers currently under construction at UCLA. This system will produce a 1 msec pulse at 2 MW output power with a 1 Hz repetition rate. The antenna system will consist of an array of crossed dipoles producing either left or right circularly polarized radiation with a power density of approximately 1 mW/m^2 at 100 km altitude during the first stage of operation. Modularity of the transmitter-antenna elements will permit easy expansion of the system to much higher powers.

INTRODUCTION

Recent results in laboratory microwave-plasma interaction experiments indicate that RF pulses of high peak power induce significant nonlinear effects at the critical layer where the incident frequency matches the plasma resonance. In this paper we describe an experiment that will extend this method to the high power pulsed excitation of the ionosphere in the auroral region [1]. Very short HF pulses, one to ten milliseconds in duration, will suffice since the relevant electron dynamics involves a submicrosecond time scale and the ion dynamics involves a tens of microseconds scale. Additionally, laboratory experiments have demonstrated that only the peak RF power, not the total energy, determines the amplitude of the excited modes. This permits constructing a pulsed HF system at a fraction of the cost of a CW system of comparable power. A distinct advantage of very high power excitation is that even in a collisional medium the nonlinear effects dominate because the fields are enhanced in a time much shorter than the collisional period and particles are accelerated over a distance much shorter than the mean free path. The active stimulation of the auroral plasma has many significant advantages: (1) The density gradient scale length in the auroral ionosphere can be as short as 1 km and the critical heights can be as low as 100 km: (2) Energetic auroral electrons provide free energy for the growth of certain plasma modes: (3) The site of the experiment allows the HF radio waves to intercept the earth's magnetic field over a wide range of angles, permitting the excitation of a greater number of plasma modes in the auroral ionosphere.

The HIPAS field system would form the core of a program involving field, laboratory and theoretical studies aimed at the understanding of phenomena produced during high power, pulsed ionospheric modification experiments. The new laboratory facility currently under construction at UCLA will utilize a large vacuum chamber (2.3 m long and 1.4 m diameter), an RF produced plasma and a high power microwave source (1 MW) to simulate pulsed ionospheric experiments. To provide an accurate simulation of the ionosphere, certain parameters including λ_0/L, the ratio of the free space electromagnetic wavelength to the density gradient scale length and ω_p/ω_{ce}, the ratio of the plasma frequency to the electron cyclotron frequency will be scaled to the ionospheric parameter

range.

The HIPAS experiments include two classes of phenomena, namely those produced by single frequency excitation and those by two or more frequency (double resonance) excitation. The phenomena of the first class include the excitation of plasma waves which enhance the plasma line, the generation of highly localized intense electric fields, particle energization, caviton formation, density profile modification, and variation of the auroral optical emission. The phenomena of the second class which uses the modulation of the high frequency resonance to couple to low frequency resonances include the excitation of low frequency modes such as ion acoustic, electrostatic ion-cyclotron, whistler, and lower hybrid waves.

EXPERIMENTAL BASIS FOR PULSED IONOSPHERIC EXCITATION EXPERIMENTS

The field excitation studies described in this paper are based on previous field experiments and on evidence that laboratory experiments can be scaled up to give data relevant to ionospheric plasma applications. Table 1 demonstrates that the ranges of relevant parameters observed in the ionospheric and laboratory plasma regimes are indeed similar.

TABLE 1 Critical Parameters for Evaluating Scaling Relationships

Parameter	Laboratory	Ionosphere
$\frac{L}{\lambda_0}$	10 to 10^3	10^1 to 10^3
$\frac{\nu}{\omega_0}$	10^{-5} to 10^{-3}	10^{-5} to 10^{-3}
$\frac{E^2}{4\pi n_0 kT_e}$	10^{-6} to 10 (achieved)	10^{-6} to 10 (expected)

In this table L is the density gradient scale-length, ω_0 and λ_0 are the frequency and wavelength respectively of the EM radiation, E is the enhanced electric field at the critical layer, ν is the electron-neutral collision frequency, T_e is the electron temperature and n_0 is the electron density at the critical layer, $n_0 = m_e\omega_0^2/4\pi e^2$.

The initial experiments were performed in large (2 m diameter x 2 m length) unmagnetized laboratory plasma systems at UCLA. They involved linear conversion of electromagnetic (EM) to electrostatic (ES) waves, resonant enhancement in the presence of electron-neutral collision, caviton formation and particle acceleration [2,3].

An electromagnetic wave (with frequency ω_0 and electric field E_0) incident on an inhomogeneous, unmagnetized plasma will penetrate the plasma until it reaches the reflection point. However, the component of E_0 parallel to the density gradient may be enhanced by plasma resonance at the critical layer where the EM frequency is equal to the plasma frequency ω_p. A simple physical picture explains this enhancement. The incoming EM energy flux travels into the critical layer at the speed of light. Whatever energy is transferred by resonance to the plasma convects out of the critical layer in the form of ES waves at approximately the group velocity, which is much smaller than the speed of light with which the EM energy convects in. The result is that a large enhanced electric field is produced close to the critical layer.

The linear conversion of EM waves to ES waves has been observed in the laboratory [2,3] where the propagation of electron plasma waves down the density gradient from the critical region was studied. Enhancements of 10^3 in the electric field were observed. The enhanced electric field was found to rise quickly with a rise time of $10(2\pi/\omega_p)$. It is on the basis of this rapid enhancement that a pulsed, high power field experiment is being set up. The region of enhanced electric field is approximately 20 Debye lengths (λ_D) in thickness parallel to the density gradient; it is centered close to the critical layer. A recent experiment at high electron-neutral collision frequencies with $\nu/\omega_0 \approx 2 \times 10^{-3}$ (auroral E layer conditions) revealed that the amplitude of the enhanced field increases as the electron-neutral collision frequency is increased (within certain conditions). This occurs because collisions reduce the number of fast electrons in the background plasma; these fast electrons can be accelerated by the localized RF field through transit-time damping, inhibiting the growth of the enhanced electric field.

At high power, the enhanced, localized electric field can induce substantial nonlinear perturbations in the density profile. The gradient of the enhanced electric field pressure, or the ponderomotive force, drives electrons from the critical layer and ambipolar forces cause ions to follow the electrons, producing a density cavity, or caviton. The length of EM pulse determines the shape of the density cavity. For short EM pulses, the caviton has the structure shown in Figure 1 [3]. Here the width of the caviton is about 20 λ_D and the density perturbation is approximately 30%. As the duration of the EM pulse increases ($\sim 2 \times 10^{-5}$ sec) more plasma is expelled from the critical layer, resulting in a step profile. The initial caviton structure decays within tens of ion plasma periods as the electrostatic wave is launched in the direction of lower electron density.

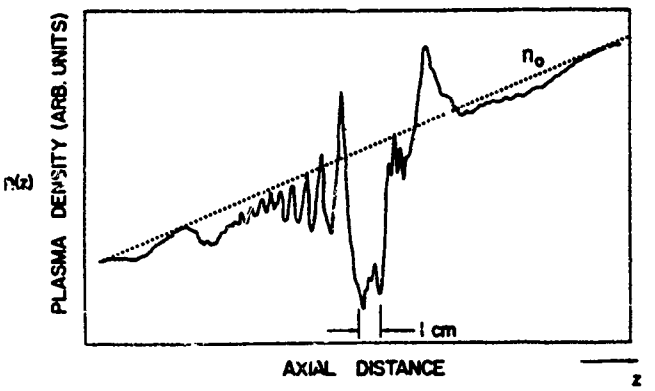


Figure 1: Caviton structure observed in the laboratory with short EM pulses (0.2 μsec, 10 kW, 1 GHz). The original linear density profile is also shown [3].

The double resonance method has been used in laboratory experiments where two EM waves with frequencies separated by a lower resonant frequency of the plasma are used to excite the low frequency modes. Ion acoustic waves in unmagnetized plasmas and electron cyclotron waves (up to the 7th harmonic, and drift waves [4] in magnetized plasmas have been generated in this manner.

There are also field experiments, using CW HF systems, which indicate that the experimental approach described in this paper is very promising. Conducted at Arecibo, Puerto Rico, they have produced high signal-to-noise ratio Thomson radar return signals. For example, Wong and Taylor [5] produced enhancements in the Thomson radar incoherent backscatter spectrum with 100 kW radio waves at 5.62 MHz. The enhanced plasma line had signal-to-noise ratio of 10^2 and was displayed directly on an oscilloscope. The risetime of plasma waves has been found in a recent experiment [6] to be as short as 0.6 msec which is less than or comparable to one ion acoustic period. This observation supports our view that linear conversion plays an important role in the generation of electrostatic waves.

In double resonance experiments at Arecibo, electrostatic waves excited by two EM frequencies, interact most strongly when the frequency separation between the EM waves is of the order of the ion acoustic frequency [7]. When the frequency separation is about twice the ion plasma frequency, the higher frequency plasma line decreases while the lower frequency plasma line increases, indicating energy flow from the higher frequency wave to the lower frequency wave and its sidebands. The enhanced ion wave activity is deduced from their complementary high frequency sidebands separated by the ion wave frequency. In a recent experiment at Arecibo [6] we have observed enhanced ion wave activities localized at the resonant height using the Thomson radar operated with the Barker code. The enhancement of ion waves is a sensitive function of the frequency separation between two pumps.

HIPAS FIELD SYSTEM

The two major components of the HIPAS HF transmitting system are the transmitter and crossed dipole antenna system.

Transmitter

The pulsed transmitter, currently under construction at UCLA, consists of a crystal oscillator and a series of amplifiers tuned to 4.9 MHz which lies in the midrange of ionospheric plasma frequencies. The transmitter is shown in Figure 2.

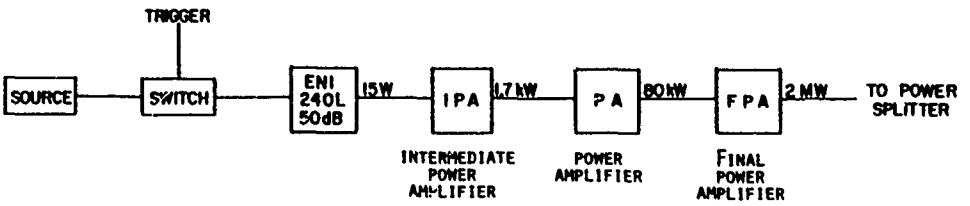


Figure 2: HIPAS Transmitter. The source will be a crystal oscillator tuned to 4.905 MHz. The switch will be used to pulse the drive signal to the amplifier chain. The Final Power Amplifier (FPA), utilizing a water cooler Machlett 7560 triode in a grid driven design has 18 dB gain.

The transmitter will generate up to 2 MW peak power with a 0.1% duty cycle. Typical operation will be a one millisecond pulse at 1 Hz. A system upgrade will be implemented this year to increase the duty cycle up to 1%, increasing the repetition rate to 10 Hz. Longer pulses at reduced power are possible so long as the maximum duty cycle is not exceeded.

Antenna System

The antenna system consists of a ring array of seven crossed dipoles arranged at equally spaced intervals along the circumference of a circle of radius R, $R/\lambda_o = 1.7$, and a single crossed dipole at the center. Calculations predict a peak gain of 18 dB with a half power beam width of 13° [8]. The power density with 2 MW transmitting power at 100 km altitude will be approximately 1 mW/m^2.

Each crossed dipole consists of four aluminum tubes (6 inch O.D.) of length ℓ, $\ell/\lambda_o = 0.23$, extending from a central feed point. Each of the four arms is fed 90° out of phase with respect to adjacent arms so that either left or right circularly polarized radiation can be transmitted. Additionally the beam can be steered along the earth's magnetic field through phasing of individual antennas. The present schedule requires installation of the antenna and transmitter by mid summer 1981. High power testing and initial experiments should begin by early fall.

The system can be easily expanded by using one 2 MW amplifier to drive each antenna. The amplitude and phase of the low level drive signal would be derived from a single source so that various pulse scenarios and directions of beam transmission could be obtained.

PLANNED EXPERIMENTS

The characteristics and expected energy density ratios of field systems with two different values of transmitted power are summarized in Table 2.

TABLE 2 Enhancement for Proposed Transmitted Power

	Total Power	Power Density	EM Field	Energy Density Ratio	Enhancement in Strong Collision Regime	
	P_o(MW)	ρ(mW/m^2)	E_o(V/m)	Λ_o	E(V/m)	Λ
E Layer	2	1	0.6	0.4×10^{-2}	15	2
100 km	16	8	2	3×10^{-2}	40	16
F Layer	2	0.2	0.3	0.5×10^{-4}	6	0.03
250 km	16	1	0.7	4×10^{-4}	16	0.2

In this table, $\Lambda_o = E_o^2/4\pi n_o kT_e$ and $\Lambda = E^2/4\pi n_o kT_e$. Although this analysis neglects the earth's magnetic field, initial calculations of the complete problem of ionospheric modification by high power HF radio waves indicates even stronger linear conversion may occur in certain situations. This is because in the magnetized plasma, EM waves are reflected at an altitude higher than the location where EM and ES modes couple to one another. That is the amplitude of the incident EM waves is not reduced by an intervening evanescent region as in the magnetic field free case [9].

We expect that the ES waves generated by the high power EM waves will be detected by the Chatanika incoherent backscatter radar. Our computations show that at the position where the ES wavelength is half that of the radar (located about 160 m from the generation region) the ES amplitude is $\delta n/n \approx 0.01$. The radar wave power returned at the plasma line would be $\approx 10^{-7}$ erg/sec and the corresponding receiver temperature $\approx 10^{4}$°K. An interesting expansion of this analysis is to consider the interaction of auroral electrons and the excited ES waves. Free energy from the electrons could be transferred to the waves when the phase velocity of the wave is equal to that of the electrons. We estimate possible wave amplitude increases of a factor of ~ 5 under certain conditions. For the high peak power regime of the experiment, we predict that cavitons will be detectable with the Thomson radar, in spite of their thinness. The estimated return signal (at the central line, because cavitons are static) again will be approximately 10^{4}°K.

Finally, we predict that the large electric fields in cavitons during the initial phase (before the ambipolar fields are set up) can accelerate a flux of secondary auroral electrons to levels that will yield optically measurable effects. Electron acceleration has indeed been confirmed in the laboratory through the detection of ArII 4610 Å emissions from the vicinity of caviton region [10]. Calculations for the auroral ionosphere (assuming several cavitons in a "linear acceleration" configuration) predict measurable optical effects caused by the acceleration of 1 eV secondary auroral electrons to 10 eV and to 16 eV, two values which correspond to the threshold levels for the N_2 first positive band, and to the second positive band and the OI 7774 Å emission, respectively. These optical emissions are prompt and the predicted emission rates (25 kR, 3 kR, 0.7 kR) would be distinguishable from the auroral background.

ACKNOWLEDGEMENT

This work is supported by the Office of Naval Research contract N00014-78-C-0754.

REFERENCES

1. A. Y. Wong, J. Santoru and G. G. Sivjee, UCLA PPG-430, April 1980.
2. R. L. Stenzel, A. Y. Wong, and H. C. Kim, Phys. Rev. Lett. 32, 654 (1974).
3. A. Y. Wong, in: Laser Interaction and Related Plasma Phenomena, Vol 4B, Edited by H. J. Schwarz and H. Hora (1977).
4. A. Y. Wong, D. R. Baker and N. Booth, Phys. Rev. Lett. 24, 804 (1970).
5. A. Y. Wong and R. J. Taylor, Phys. Rev. Lett. 27, 644 (1971).
6. A. Y. Wong, D. Eggleston, J. Santoru, G. Morales and R. Behnke (Private communication) 1981.
7. A. Y. Wong, W. F. DiVergilio and K. Iizuka, Phys. Lett. 75A, 144 (1979).
8. M. Ma (Private communication) 1981.
9. R. Sharma, M. Shoucri, J. E. Maggs and G. J. Morales, Bulletin of the American Phys. Soc. 25, 916 (1980).
10. W. F. DiVergilio, C. Kim, J. Tang and A. Y. Wong, in: Technical Digest Inertial Confinement Fusion Conference, San Diego, CA, February 1978.

Ionospheric Predictions -- A Review of the State of the Art

Kenneth Davies
National Oceanic and Atmospheric Administration
Environmental Research Laboratories
Space Environment Laboratory
Boulder, Colorado 80303 USA

ABSTRACT

This paper reviews the state of ionospheric predictions following the Solar-Terrestrial Predictions Workshop held in April 1979 at Boulder. The review includes: the various uses of the ionosphere, the physical bases, long-term and short-term variations, the techniques used for their prediction and the current state-of-the-art of ionospheric predictions. The different requirements of various users (system designers and planners) are discussed together with the needs for timely dissemination systems.

Because ionospheric predictions and forecasts have been underway for many years the advances have been of second order. The main advances include the incorporation of features; such as the auroral oval, the magnetospheric cleft, the mid-latitude trough, etc; the positions of which vary with local time and solar-terrestrial activity. Other advances include four dimensional models of the ionosphere, which yield electron density profiles useful in raytracing and some new methods for predicting E-layer and E_s-layer characteristics, transionospheric propagation, which requires prediction of total electron content and scintillation characteristics, has come to the fore. In recent years, advances have been made in the use of empirical and physical models of the ionosphere.

Short-term ionospheric forecasting has also progressed in recent years primarily through technological advances in data acquisition (e.g. satellite sensors), data processing (computers) and dissemination of information. Nevertheless, solar-flare forecasting -- a vital ingredient of magnetic/ionospheric storm forecasting -- is still largely an art in which the forecasters' experience and skill, in recognizing and interpreting solar features, plays the dominant role.

INTRODUCTION

The aim of this paper is to review the state of ionospheric predictions in the light of presentations made at the Solar-Terrestrial Predictions (STP) Workshop held April 23-27, 1979 at Boulder, Colorado (Donnelly, 1979). In an earlier discussion of ionospheric predictions (Davies and Smith, 1978), the author discussed the various uses of the ionosphere, societal concerns, the physical bases, long-term and short-term variations and the techniques used for their prediction. Since that study was published, the STP Workshop presented a unique opportunity for predictions users and scientists to exchange information concerning prediction techniques and recent scientific advances in the solar-terrestrial field. While the Workshop covered a broad field of solar-terrestrial concerns, we shall confine our attention to ionospheric predictions. An important question raised at the STP Workshop was "Can we significantly improve predictions?" Judged on this criterion ionospheric predictions fared poorly, perhaps, because it is a relatively old field which has been well farmed.

In indexing the papers appearing in the proceedings the following system has been used for brevity: e.g. Klobuchar (II, 217) represents the paper whose first author is Klobuchar, appearing in Volume II starting on page 217. Some useful contributions to the ionospheric literature are provided by the papers reviewing activities at various institutions, e.g. Reddy (I, 118; II, 203) and Larsen (II, 617), and by the Working Group Reviews and recommendations by Mitra (II, 203) on Communications Predictions, Vondrak (II, 476) on Magnetosphere-Ionosphere Interactions, Hunsucker (II, 513) on High-Latitude E- and F-Region Ionospheric Prediction, Rush (II, 562) on Mid- and Low-Latitude E and F Region Working Group and Thrane (II, 573) on D-Region Predictions.

In this paper we shall review: (1) some of the modern demands that relate to solar-terrestrial predictions, (2) the physical picture of the earth's environment, (3) the existing prediction tools

for the D, E and F layers for different latitude zones, (4) the increasing importance of transionospheric (satellite-to-ground) communication systems, (5) the roles of numerical and physical models and (6) short-term ionospheric forecasting.

MODERN REQUIREMENTS FOR IONOSPHERIC PREDICTIONS

In earlier days the requirements for predictions were confined to maximum usable frequencies, lowest usable frequencies, signal-to-noise ratios and average angles of elevation etc. for long-term planning of radio circuits and to alert the radio operator to change frequency in the event of circuit failure. While these predictions have been of inestimable value to radio system designers, the advance of communications technology has placed greater demands on the ionospheric predictor. Users of advanced systems need predictions tailored to their particular system and interpreted directly in their own system parameters. Some examples of system needs are: (1) communication systems which need predictions of multipath, dispersion, fading characteristics etc., (2) satellite systems that require parameters including the total time delay, phase coherence over the receiver bandwidth and depth and occurrence of amplitude scintillation, (3) navigation systems that require propagation time delays, (4) radars which need range and angle corrections. Furthermore, while climatological (median) predictions are useful, the designer of a radio system needs to have some idea of "worst-case" conditions. This is particularly true in high-latitudes where the ionosphere is frequently disturbed.

The main categories of users of predictions are:

(a) Planners and designers who require long-term predictions for the determination of system coverage and frequency allocation for communications, surveillance and navigation systems.

(b) System operators who are interested mainly in short term forecasts or even knowledge that a geophysical disturbance is in progress.

(c) Users of navigation and timing systems who depend on the stability of the reflecting region for reliable time and position determination.

(d) Communicators and broadcasters who use radio signals reflected from the ionospheric E and F layer and which are absorbed in the D-region of the ionosphere.

(e) Those concerned with the lifetimes of satellites which are decreased by increased drag resulting from atmospheric heating.

(f) Researchers who study the ionosphere and upper atmosphere or whose work is adversely affected by the ionosphere, e.g. radio astronomers, satellite trackers.

A primary need of users of ionospheric predictions is a timely dissemination system, for example, there is little use in receiving a prediction after a disturbance has started. Telemetry services limit the timeliness. Many forecasters do not realize that priority TWX messages can take over 24 hours to reach their destinations - a time span over which forecasts are not always reliable. At, or shortly after the commencement of an event (e.g. a polar cap absorption event), forecasts of its subsequent behavior are often quite accurate.

Individual users need specific information for example: navigators want the magnitude of distance errors, communicators want maximum and minimum frequencies. Furthermore, an aircrew operating in the arctic and changing position rapidly has different requirements from those of a stationary or a shipborne operator. In the field of HF data links, predictions are required of effects on the error rate of fading, multipath and signal/noise level. Thus the prediction requirements depend upon the complexity and sophistication of the system in use. For many users, simple schemes, for a desk or pocket calculator, may suffice as for example, the method of Picquenard and de Paula (IV, D2-41) and of Meisel (IV,D2-31).

SOME PHYSICAL CONSIDERATIONS

During the past decade or so a great deal has been learned of the physical processes that create, maintain and remove ionization in high latitudes. A pictorial representation of the earth's neighborhood is shown in Figure 1 which include such features as the solar wind, the magnetosphere, the plasmasphere and the radiation belts all of which affect the ionosphere. More detailed features that influence the ionosphere in high latitudes include (1) auroral ovals, (2) polar caps (3) mid-latitude troughs and (4) magnetospheric clefts.

The high latitude ionosphere is profoundly affected by particle precipitation, electric currents, electric coupling, plasma transport and thermospheric heating. Depending on the level of disturbance, the energy input above 90 km by particle precipitation and Joule heating is between 10^{10} and

10^{12} watts, which is comparable with the 6 x 10^{11} watts of solar electromagnetic energy absorbed above 90 km (Vondrak, II, 476). Energy dissipated by ionospheric currents is transferred directly to the neutral atmosphere and produces convection (winds) flowing to lower latitudes. This introduces chemical changes in the F region that results in higher electron loss and causes negative F-region storms in middle latitudes. In terms of relative magnitudes, about 30% of the average energy input is by charged particles and 70% by Joule heating. The equatorward edge of the auroral oval marks the boundary between the corotating magnetic field (the plasmasphere) and the field lines that are dragged into the magnetotail by the action of the solar wind. It is believed that low energy (< 400 eV) particle precipitation also defines the equatorward boundary of diffuse aurora. Poleward of the diffuse aurora is the region of discrete auroral arcs which are associated with layers of dense electron density in the E region. During disturbed periods the oval expands equatorward and increases in intensity. There are three auroral absorption regions, one just equatorward of the diffuse aurora near midnight, a pre-noon region and one in the early morning during disturbed periods (see Figure 2).

Another area of particle precipitation into the high-latitude ionosphere is the dayside cleft. The cleft is typically 2° to 4° wide in longitude and between 78° and 80° invariant latitude near local noon. This is a region where the solar wind has, more-or-less, direct access to the upper atmosphere and where enhancements of electron density have been observed. The cleft region moves with geomagnetic activity and with the interplanetary magnetic field (IMF).

A third feature of the E and F regions is the electron density trough in the night ionosphere in which the density is low. There is a sharp increase in density at the poleward boundary caused by particle ionization. The main trough is a winter nighttime phenomena. Because the trough moves with time and geomagnetic disturbance and because of the sparseness of the ionospheric sounding network little is known of its variations with solar cycle, disturbance, etc. It is known that during large disturbances the trough moves to lower latitudes as seen in Figure 3. Furthermore, trough-like structures can be produced by large-scale traveling ionospheric disturbances (gravity waves) traveling equatorward from the polar regions. Electron density gradients in the trough are important to long distance communications on HF.

The fourth major feature of the high-latitude ionosphere is the polar cap over which intense D-region absorption events, called polar-cap absorptions (PCA), cause disruption of radio circuits.

The major disturbances in the magnetosphere, ionosphere and geomagnetic field are called storms. Storms often follow the appearance of solar flares but which are also of the recurrent type due to the rotation of the sun. Storms are believed to result from the injection of hot plasma from the geomagnetic tail into the auroral oval, and by Joule heating by the convection electric field. The size and location of the auroral oval is fairly well monitored by ground and satellite detectors. Advances in the understanding of storms have enabled construction of semiphysical predictive indices based on solar wind characteristics, so that qualitative prediction is possible. Presently, warnings of disturbances in high latitudes can be issued but little in the way of quantitative prediction can be made.

In middle latitudes, the plasmasphere acts both as a reservoir and as a transmitter of ionospheric plasma. Thermal plasma flows from one hemisphere to the other along magnetic flux tubes thus coupling conjugate ionospheres that are in different (summer-winter) seasons. During the day thermal plasma flows from the ionosphere, where it is created, into the plasmasphere and during the night the ionospheric F-layer is partially maintained by ionization from the plasmasphere. During a plasmaspheric storm (Kersley et al., 1978) there is a depression in the plasmaspheric content which reaches its lowest value on the third day and is followed by a slow recovery which may last up to 14 days.

The electron and ion distribution in the F2 layer in middle latitudes is strongly influenced by electric fields and neutral winds that cause vertical motion of the plasma. Thus information on the temporal and spatial structures of the electric field and neutral wind are essential to the calculation of F2-layer behavior.

The equatorial ionosphere is dominated by vertical electrodynamic drift, which is particularly important after sunset and which causes the equatorial anomaly in which the maximum electron densities occur on either side of the magnetic-dip equator. Bubbles of depleted plasma are thought to be caused by the intrusion of neutral gas into the F2 layer resulting in an unstable situation in which the bubbles rise and produce small scale irregularities (e.g. Costa and Kelley, 1976) that cause spread F on ionograms and gives rise to radio scintillations.

The low latitude F-region, unlike that in middle and high latitudes, is largely protected from particle effects. It is dominated by the equatorial anomaly in which the maximum electron densities occur around ± 15° of the magnetic dip equator as seen in Figure 4. This feature is caused by electromagnetic uplift of ionization followed by diffusion down the magnetic field and away from

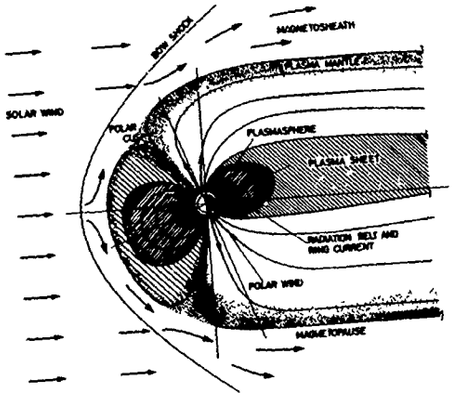


Fig. 1

The earth's environment

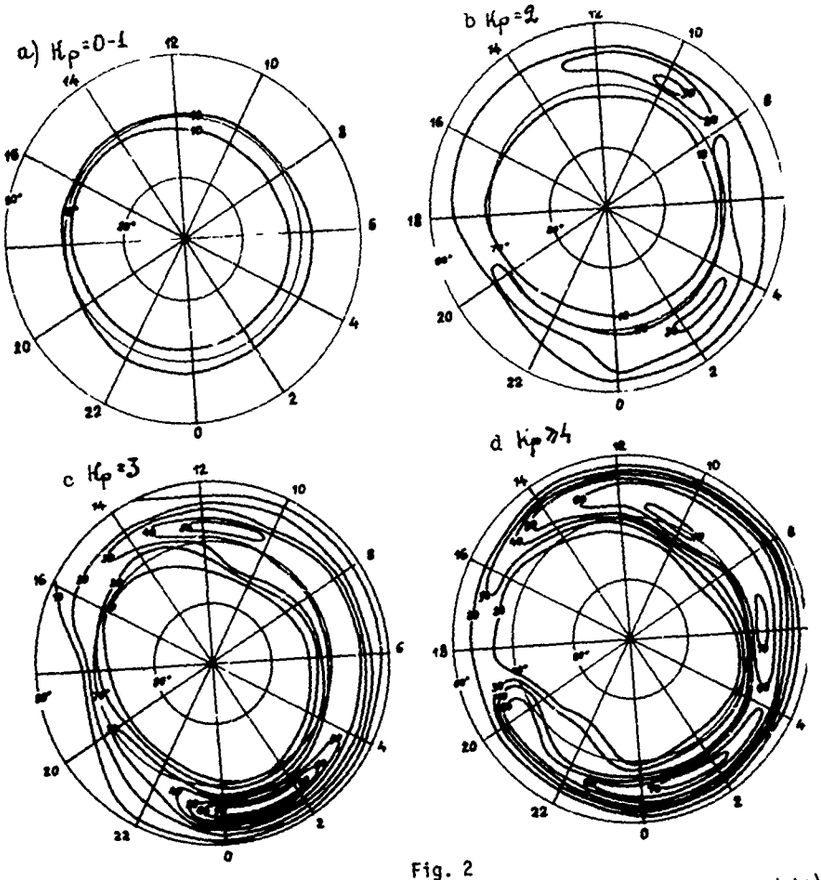


Fig. 2

Polar plots of the probability of auroral absorption (from 30 MHz riometer data) occurrence for values ≥ 0.5 dB in winter 1964, in terms of corrected geomagnetic latitude and time (from Besprozvannaya II, 528).

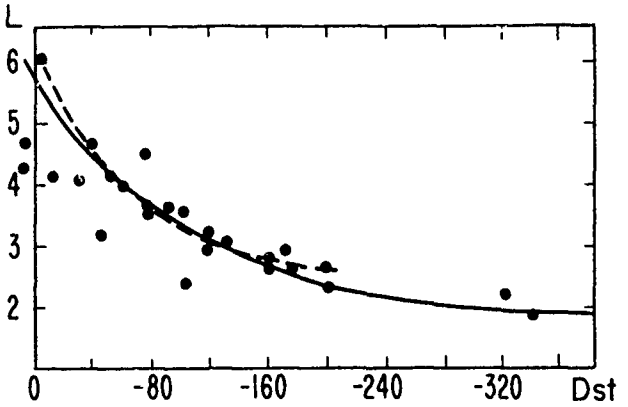


Fig. 3

Average position of the ionospheric trough minimum in the night sector for 25 global magnetic storms versus Dst (see Besprozvannayo, II, 528).

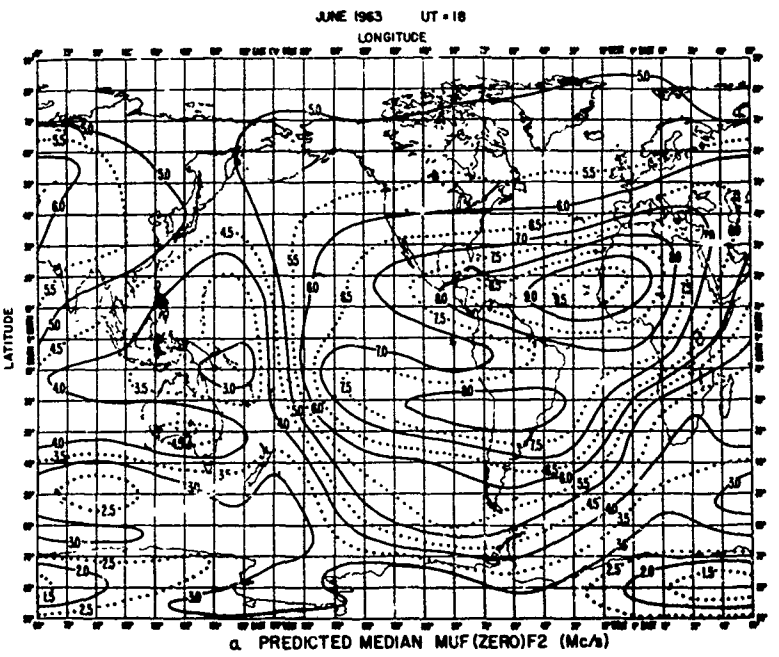


Fig. 4

Numerical map of F2

the equator giving rise to a "fountain-like" plasma motion that results in a trough of ions near the equator and humps on either side. Because of day-to-day changes in the eastward electric field, ion production etc., there are high fluctuations in F-region parameters in low latitudes. The positions of the density "humps" vary in latitude from day-to-day, seasonally, with sunspot numbers and during storms. Thus a given location on the surface of the earth may be outside or inside the equatorial trough or at a peak depending on prevailing conditions.

The E-region in low latitudes has a distinctive daytime type of equatorial sporadic E which is patchy and transparent to radio signals reflected from the upper layers. This sporadic E is closely associated with the equatorial electrojet current that flows along the magnetic equator and which gives rise to the polarization electric fields that produce the F-region trough. The sporadic E has irregularities aligned with the magnetic field.

SUMMARY OF OLDER PREDICTION METHODS

F2 Layer Maps

Ionospheric predictions are valuable in the design of and the operation of point-to-point radio communications, broadcast and satellite-to-ground systems. Information needed include, for example: the useable range of frequencies (for frequency management), the height of reflection of radio waves (for antenna design), attenuation (for estimating required transmitter power and/or service area), fading or scintillation characteristics, total electron content (for time delay) and radio noise levels. Data on maximum (critical) frequencies, layer heights, and radio-wave absorption collected at a large number of sounding stations have been used to construct computerized contour maps of the ionosphere of the type illustrated in Figure 3. The most comprehensive of these mapping methods is that developed at the Boulder Laboratories of the U. S. Department of Commerce (Jones and Gallet, 1962; CCIR 1978). The method has been extended to three dimensions by Jones and Stewart (1970) to represent electron density profiles. In addition, the technique has been applied to represent departures of foF2 from the monthly median during an individual ionospheric storm (Jones et al., 1973). A numerical representation of the critical frequency of the F1 region has been given by Rosich and Jones (1973) and by CCIR (1978).

The afternoon equatorial trough and the "humps" in maximum or penetration frequency, at approximately 15°N and 23°S, are clear in Figure 4. However, the high latitude features associated with the troughs, auroral zones, magnetospheric cusps are poorly represented. One of the reasons for this is that the locations of these features varies with local time, magnetic disturbance (see Figure 3), etc., and tend to be smoothed out in monthly median maps.

The E Region

There are two methods in use for predicting the maximum penetration frequency (foE) of the regular daytime E layer: those of Muggleton (1975) and Leftin (1976). The nightime foE predictions are based on the work of Wakai (1971). The day-to-day fluctuations of foE about the monthly median values are of the order of 6 per cent.

Sporadic E(E_S) is highly variable as compared with the normal E-layer and numerical maps of sporadic E characteristics have been constructed by Leftin and Ostrow (1969). Near the magnetic dip equator sporadic E is regular in occurrence and CCIR Report 259-3 (1978a) provides a method for calculating signal strength of E_S reflections as a function of foE_S, circuit distance and antenna characteristics. A new prediction method for South America (Giraldez, IV,C-87) gives the dependence of the blanketing frequency on solar zenith angle, sunspot number and latitude. Presently, however, there is no practical computer program for the determination of HF skywave characteristics for E_S propagation on a global scale.

The D Region

For medium frequency and high frequency absorption the most widely used prediction scheme is that described by Barghausen et al., (1969). In this program the absorption is expressed by semi-empirical formulas in terms of frequency, distance, sunspot number, solar zenith angle and geographical position. The absorption predictions are useful and reasonably accurate for low latitudes and for middle latitudes in summer.

Of special interest is the prediction scheme used for the VLF OMEGA navigation system (Swanson 1977) which may be the most used of all ionospheric predictions, being in daily use by thousands of navigators. The method assumes that the phase of a signal can be expressed as a sum of incremental phase shifts over path increments and includes both ground and ionosphere variations. Present system accuracy is significantly degraded by prediction inaccuracies.

ADVANCES IN SYNOPTIC MAPPING

General Remarks

Because predictions of the ionosphere, and especially its critical frequencies, have been made for nearly fifty years, it was unlikely that major advances would be forthcoming at the Workshop, and such was the case. Nevertheless, some important efforts were discussed concerning the application of recent data to update the CCIR maps, for short-term prediction and for large-scale features, such as the mid-latitude trough and auroral oval.

The F2 Layer in Middle Latitudes

The most widely used method of F2 predictions is that employing the numerical coefficients described in the CCIR Atlas of Ionospheric Characteristics (CCIR, 1978). One useful extension of this basic system to four dimensions is that of the U.S. Air Force Global Weather Central to represent electron density versus height profiles at any longitude, latitude and time (Tacione I, 367). From radio soundings, the ionospheric layers are represented by empirical orthonormal functions between heights of 95 km and 2000 km. The calculated electron density profile is calibrated by minimizing the difference between the calculated total electron content (TEC) and the observed TEC by adjusting the scale height of the F2 layer. This four-dimensional model is useful for tracing radio rays through realistic ionospheres.

At Stanford Research Institute (SRI) International, Hatfield (IV, D2-1) has incorporated the latest information on ionospheric features that affect HF radio propagation into a code (AMBCOM) that is easy and economical to use. This code represents the electron density profiles by three parabolas and includes ionospheric tilts by specifying the parabolic parameters along the circuit. These parameters are obtained either from the numerical maps or they may be obtained directly from ionospheric soundings. The code allows reflection from both topside and bottomside of sporadic E layers and includes a model of the auroral ionosphere, with auroral absorption, that depends on geomagnetic activity. In addition it has a homing procedure for point-to-point communications, a surveillance capability for over-the-horizon radar and the output includes contour maps of signal/noise plus raypath and wavefront plots.

In the Soviet Union (Anufrieva IV, C-57), a prediction system was proposed based on global maps that include, in addition to $N_{max}F2$, the maximum height gradient in N to specify the electron density profile. This permits the prediction of the maximum frequencies in an ionospheric duct (Shlionsky, IV, D3-60) which exceed considerably the MUF of the F2 layer by factors of from 3 to 8.

In the long-term HF predictions system of the Deutsche Bundespost (Damboldt, I, 25) predictions are made of critical frequencies, signal strength, etc. for a given sunspot number. The technique incorporates signals on frequencies above the "classical MUF" that propagate by paths involving ionospheric and/or ground scatter, E_S, ionosphere-ionosphere reflections, ducting, etc. An empirical factor is applied to the "standard MUF", calculated from the maps, to obtain the "operational MUF".

A procedure used in Japan (Maeda, I, 212), based on the CCIR maps, provides monthly median predictions three months in advance and includes frequency-versus-time data and path loss for fixed circuits and for ship-to-shore between foreign harbors or fishing banks and Tokyo. These forecasts are available to both domestic and foreign users.

A method of predicting foF2 by the monthly ratio method has been proposed by All India Radio (Murthy, IV, D2-54) in order to overcome the "saturation" effect of the dependence of monthly median foF2 on the twelve month average sunspot number. The monthly ratio (MR) is given by:

$$MR = \frac{foF2\ (\text{month } M+1, h_i)}{foF2\ (\text{month } M, h_i)} ,$$

where i is the month number, and it has been calculated for every month of a previous sunspot cycle. To predict a future median foF2 one takes the median for the latest month and multiplies by the appropriate MR.

The influence of horizontal structure (gradients) on radio propagation received considerable attention. The techniques of ray tracing used at IZMIRAN, USSR for calculating hop distance, MUF, angles of arrival, etc. were discussed by Kerblay (IV, D2-65). At the National Physical Laboratory in India, special attention is paid to the effect of the equatorial (or Appleton) anomaly on radio propagation and to keeping prediction users aware of the conditions that prevail in low latitudes.

The F2 Layer in High Latitudes

Predictions of foF2 are relatively poor in the auroral zone partly because the critical frequency itself is ill defined and partly because of the high variability. However, additional data have been incorporated into the CCIR (1978) model by the Air Force Geophysics Laboratories (AFGL) and the Rome Air Development Center (RADC) to furnish an improved model which more accurately represents the spatial structure. A new model developed by Stanford Research Institute (Hatfield, IV, D2-1) retains the foF2 and h_mF2 in the RADC-1976 Polar Model (Vondrak et al., 1977). It incorporates an auroral-E layer and a valley, based on incoherent scatter measurements at Chatanika, Alaska, as well as an absorption map that depends on geomagnetic activity. The new SRI code makes an important advance in prediction technique by representing auroral oval phenomena as functions of magnetic activity, geomagnetic latitude, time and solar zenith angle. The midlatitude trough is a feature of the nightside. The auroral morphology is implemented with a K_p-dependent foF2 (Elkins and Rush, 1973), fE_s, and auroral absorption.

The three-dimensional models developed by RADC (Miller and Gibbs, 1974) and SRI are suitable for ray tracing which enables computation of off-great-circle propagation by refraction and by scattering from field-aligned irregularities. A major problem with such models is that averaging, in order to obtain representative critical frequencies and heights, generally results in reducing the horizontal gradient. Because of the complicated ionospheric structure in the trough, auroral and polar-cap regions, specific ray paths are sensitive to the location of the structures. Thus the question arises as to whether a "strongly averaged" model gives better predictions than a model with misplaced individual features. This is particularly important to those system designers who are concerned with worst-case estimates which are more adequately represented even by misplaced structures than by average morphologies. In spite of these limitations it should be realized that the modern 3-D models do provide a distinct improvement over the older models which simply assumed horizontal reflecting layers. It is quite likely that in high latitudes, average models will never be of much value to the radio operator. Perhaps the best hope for useful predictions is in near real-time input from satellite sensors.

Several attempts to improve short-term predictions of the high latitude ionosphere have been made by Soviet workers. The day-to-day variability of such parameters as foF2, h_pF2, and h'F may be predicted using the vertical and radial components of the interplanetary magnetic field (IMF) (Zevakina IV, C-27). Graphs giving the ranges of variabilities of ionospheric parameters for disturbed high latitude ionosphere were given by Kovalevskaya (IV, D2-16) while Besprozvannaya et al., (II, 528) treat the large scale features of the polar ionosphere that accentuate the variability.

The E Layer in Middle Latitudes

A new approach to E-layer prediction by Ivanov-Kolodny and Nusinov (IV, C-82) incorporates the dependence of the layer production jointly on solar X-rays and extreme ultraviolet. Their formula for the E-layer critical frequency is:

$$(foE)^4 = I \cos^p \chi \left[\frac{R_o}{1+R_o} \cdot \frac{X}{X_o} + \frac{1}{1+R_o} \frac{U}{U_o} \right]$$

where X and U are values of ionizing fluxes in the X-ray (8 Å to 20 Å) and EUV (304 Å) parts of the solar spectrum and X_o and U_o are the corresponding reference values. The exponent p is approximately 1.07 in summer and approximately 1.23 in winter and

$$R_o = 0.23 + 0.07\,\phi(D)$$
$$p = 1.15 + 0.08\,\Phi(D)$$
$$I = I_o\,[1 - 0.06\,\psi(D)]$$

where ϕ, Φ, and ψ are functions of day number D. I_o is best obtained from observed foE, e.g. at Moscow I_o = 190 MHz^4 (i.e. foE = 3.71 MHz).

For the prediction of the horizontal structure of sporadic E, Kerblay and Nosova (IV, C-77) use the formula

$$(fE_s) = (fE_s)_o \left[1 - K \sin\left(\frac{2\pi x}{\ell_1} + \phi\right) \cos\left(\frac{2\pi y}{\ell_2} + \psi\right)\right]$$

where K is a horizontal structure parameter, ℓ_1 and ℓ_2 are the scales of E_s in the x and y directions and ϕ and ψ are appropriate phase shifts.

The E Layer in High Latitudes

Auroral electrons with kilovolt energies are the primary source of the high-latitude E region, the E region structure closely resembles the pattern of auroral particle precipitation. In future, particle precipitation data may be the bases of high-latitude prediction systems. In addition to the kilovolt particles, higher energy particles deposit energy in high latitudes during relativistic electron precipitation and polar cap absorption events, which affect primarily the D region.

Besprozvannaya et al., (II, 528) have prepared maps depicting E_S ionization distribution for quiet and disturbed geomagnetic conditions at high-and low-sunspot numbers. Under disturbed conditions the region of enhanced E_S coincides with the auroral oval and the zone of 1.3 keV electron precipitation. These workers find also that auroral E occurrence has a horse-shoelike pattern open on the dayside as shown in Figure 5.

The D Region in Middle Latitudes

The D region is important to radio usage from two aspects: (1) the absorption of radio signals and (2) the phase stability of signals. The former affects the signal strengths and, therefore, the required transmitter powers while the latter influences ELF, VLF and LF timing and navigation systems. Among the recommendations made at the STP Workshop were: (a) those to improve the utility of existing knowledge and (b) improvements in fundamental knowledge required for the future. Among those in category (a) are (i) timely dissemination, (ii) automatic devices indicating user response, (iii) recalculation of existing absorption maps using the Dyson and Bennett (1979) formula, (the new formula will reduce the electron densities needed to account for the observed absorption in low latitudes and thus reduce the size of the "equatorial anomaly" in the D-region (George 1971), and (iv) more simple measurements of the D-region (e.g. absorption, sudden disturbances, etc). In category (b) the most promising line is the use of satellite observations to identify the positions and intensities of disturbed areas.

It is convenient to distinguish areas of the globe where specific processes dominate: (a) low-geographic latitudes (± 30°), (b) middle latitudes (> 35° geographic), (c) high latitudes (> 60° magnetic or L > 4) Area (a) is relatively free of particle precipitation effects except near the South Atlantic anomaly in the geomagnetic field. George (1971) showed that the D region is geomagnetically controlled because there are HF absorption maxima near 20° north and south of the magnetic equator, possibly the result of downward transport of nitric oxide from the F region. In middle latitudes high variability is introduced by energetic particle precipitation, aeronomical and dynamical factors. For example, there is a strong "storm after-effect" in which HF absorption and VLF propagation effects linger for several days after a major magnetic storm. Also the neutral D region is influenced by dynamical coupling with the lower atmosphere (Gregory and Manson, 1969). The winter anomaly of high absorption in Europe can be explained by meteorological conditions in the mesosphere (Offerman, 1977; Thrane II, 573). Thus our ability to predict the middle latitude D region in winter months depends on our ability to predict the meteorology of the mesosphere and lower thermosphere on the appropriate spatial scales.

Workers at All India Radio (Bhatnagar IV, D3-55) have found radio noise measurements on low frequencies and medium frequencies (transmitted via the D region in India) are lower than those predicted by CCIR (1964), which therefore needs revision. Work on absorption of radio waves in low latitudes was reported also by Oyinloye (IV, D3-1) and by Kotadia (IV, D3-20). A formula used at the Space Research Centre, Poland, for estimating D region, absorption of radio waves resulting from solar flares is given by Klos and Stasiewicz (I, 67).

Ionization of the D and E regions, by precipitation of electrons with energies of the order of several hundreds of keV, takes place during the recovery stage of a magnetospheric storm. Spjeldvik and Lyons (IV, B-59) propose a prediction scheme, based on a physical model, which predicts order-of-magnitude effects on electron density enhancements that may be used to forecast VLF to MF ionospheric radio propagation disturbances.

The D Region in High Latitudes

In high laitutdes the D region is influenced by disturbances: auroral absorption, polar cap absorption and relativistic electron precipitation associated with magnetic storms. Standard prediction methods take account of high-latitudes only in a very rudimentary way. For example, the Barghausen et al., (1969) program introduces "excess system loss at high latitudes", and OMEGA predictions include average auroral and polar effects on VLF phase velocity. Auroral absorption is well mapped, in a statistical sense, by riometers (Hargreaves, 1969) but the maps have not yet been developed into a useful prediction scheme. A computer program has been developed at the Appleton Laboratory, Great Britain (Bradley, 1975) that includes a statistical distribution of auroral absorption in time and space. Soviet workers (Besprozvannaya II, 528) have constructed polar plots of the probability of auroral absorption for various values of K_p - see Figure 2. With increase of K_p, auroral

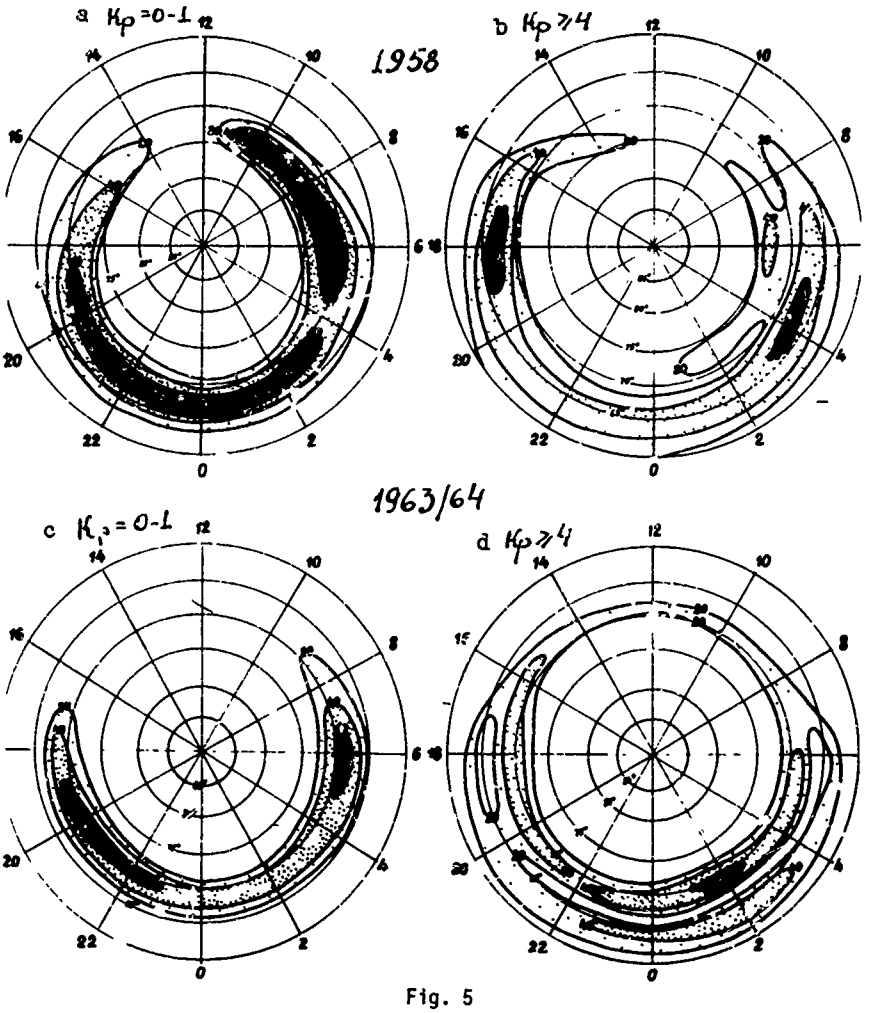


Fig. 5

Polar plots of occurrence probability of the auroral E in winter. The coordinate grid is corrected geomagnetic latitude and time (from Besprozvannaya et al., II, 528).

sorption moves to lower latitudes and into two maxima ($\sim$ 0200 and 0600 GLT). Of course, these maps are for averaged conditions and, therefore may not represent an individual disturbance.

Because the D-region in high latitudes is susceptible to short-term disturbances it is dealt with later under "Ionospheric forecasting".

TRANSIONOSPHERIC PARAMETERS

Quiet Behavior of Total Electron Content

In the past decade an increased emphasis has occurred in communication involving satellite transmissions and, therefore, transionospheric propagation, compared with those using reflected radio signals. There is a fundamental difference in approach to the ionosphere of operators using the two types of systems. In the case of reflected waves the ionosphere is essential and the radio user needs to know the variations of the ionosphere to make maximum use of his circuits. On the other hand, to the satellite-systems operator the ionosphere is simply a nuisance and what he/she needs to know is whether an ionospheric disturbance will be an inconvenience and what correction can be made to the system to allow for the ionosphere; this is true, of radio astronomers, satellite trackers, satellite geodesists and so on.

The integrated columnar electron content (TEC) of the ionosphere is of importance not only for satellite ranging corrections, satellite navigation systems, for corrections to radio astronomical refraction, and satellite time disemmination but as a general index of the overall state of the ionosphere, e.g. storm response. Two approaches are used to predict TEC: (1) extrapolate from observations, and (2) use numerical maps to obtain N_{max}, e.g. from the CCIR (1978) maps, together with assumed electron density profile shape. Approach (1) is very limited because of the scarcity of the TEC observations, especially in the southern hemisphere. For periods of a few months in advance, TEC predictions are accurate to $\pm$ 20 per cent (Klobuchar, II, 217). A single daytime prediction has a 25 per cent RMS error from the monthly mean. Future progress in this area requires that computer programs be small, quick running and capable of forecaster interaction. Input parameters must be readily available and the output should be directly usable. It appears that the main improvements will arise from a larger data base, rather than from more theory, and updating with real time TEC measurements.

With approach (2) using existing numerical maps, the RMS error in foF2 at solar minimum is approximately 0.5 MHz and at solar maximum around 0.8 MHz (see Klobuchar, II, 217). The RMS deviation is lowest in summer (May-July) and is highest ($\sim$ 12%) around the equinoxes (September-March). It should be realized, however, that the foF2 provides a good predictor of total electron content.

The total electron content N_T between ground and a synchronous satellite can be divided into two parts: (1) the Faraday content N_F up to a height of about 2000 km below which the electrons are effective in rotating the polarization of radio waves and (2) the plasmaspheric content N_p between 2000 km and the satellite (Davies, 1980). In middle latitudes, to a first approximation, N_p is independent of time of day but, because of the pronounced diurnal variation of N_F, the ratio N_p/N_F varies between about 15% during the early afternoon to around 50% just before dawn. The vertical plasmaspheric content over the U.S.A. is about 3×10^{16} el m^{-2} at low sunspot numbers ($\sim$ 20).

Storm Behavior of Electron Content

The solar disturbance (local time) response of the Faraday content to an ionospheric/geomagnetic storm over North America (Mendillo, IV, C-1), is as follows: if the storm starts early in the (solar) day there is an enhancement in N_F on the afternoon of that day followed by a sharp drop around sunset and below normal content for the next 2 or 3 days. Mendillo and Klobuchar (IV, C-15) show that the diurnal variations of TEC departures from the monthly mean on quiet and disturbed days are virtually "mirror images". Thus a general knowledge of geomagnetic conditions may enable a real-time update to monthly mean predictions.

The storm behavior of the plasmaspheric content is as follows: there is a decrease in N_p (by a factor of up to 5) reaching a minimum on day 3 followed by a slow recovery lasting up to about two weeks under quiet conditions. At moderate to high sunspot numbers the plasmasphere may never be allowed to recover to its quiet state before being depleted by a subsequent storm.

Gross averaging of storm variations in TEC is of only limited value for prediction of storm behavior because the response of the ionosphere to a given Kp or A_p differs from one storm to another.

Predictive Capability

Lunar tidal variations of up to 15 per cent in N_F can be predicted with reasonable accuracy provided that a sufficient data base exists (Bernhardt, 1979)

One way of predicting TEC is to update monthly median curves using real-time data. For a 50% reduction in error the prediction lead times are as follows:

	Day	Night
Solar Maximum	3 hours	1 hour
Solar Minimum	1 hour	1/2 hour

The STP Workshop suggests the following ways of improving transionospheric progagation predictions: (1) use of contour maps of TEC such as those of Davies, et. al. (1977) shown in Figure 6, (2) use of standard data formats, calibrations and interpretation of TEC measurements, (3) increased participation of World Data Centers for Disemmination of TEC data. One way of improving predictions is to use scaling procedures such as that of Donatelli and Allen (IV, D1-65) to reduce the residual error in ionospheric refraction. Donatelli and Allen (IV, D1-65) give curves showing the residual error in range correction using an updating technique. Their curves, shown in Figure 7, are interpreted as follows: a given curve starting at a given hour, 08 say, has zero error because the range is calibrated against observation at that time. At 10 LT the error in the predicted value is about 5 per cent while at 12 LT is is around 11 per cent. It appears that after only a few hours a prediction is no better than a prediction of the mean.

Irregularities

Small scale irregularities are of considerable importance to transionospheric radio because they cause scintillations in amplitude and phase which distort the radio signals and limit the channel capacity. The primary scintillation models available for prediction are those of Fremouw and Rino (1978), of Singleton (IV, D1-16) and of Aarons et al., (1978). Amplitude scintillations obey a Nakagami distribution defined by the S_4 index and t_c the fade coherence time plus a coarse coherence bandwidth. Provided that t_c is very much longer than the time for one bit or baud, bit error can be expressed in terms of S_4. As t_c decreases to the baud interval, waveform distortion produces catastrophic failure.

Scintillation models for predictive purposes can be constructed based on the following parameters: (1) the perturbation strength (2) spectral index (3) height and thickness of irregularity layer (4) axial ratio, i.e. elongation and (5) the inner and outer scales. The irregularity strength is obtained directly, from in-situ satellite measurements (Rino, 1979a, b) and, in the equatorial region, varies by six orders of magnitude. Alternatively Singleton (IV, D1-1) uses spread F from ionograms to infer the irregularity strengths, however, one problem with this approach is that the irregularities may be above the F2 peak, such irregularities produce scintillation but might not be seen on conventional ionograms.

The past ten years has seen considerable improvement in our knowledge of scintillation morphology particularly that in the equatorial regions. Equatorial scintillation peaks at latitudes $(\lambda_m) \pm 10°$ of the magnetic equator and decays with a Gaussian shape viz $\exp[-\{(\lambda_m - 10)/10\}^2]$, the scintillation index S_4 decreases with radio frequency f as $f^{-1.5}$ while the RMS phase fluctuations ϕRMS vary as f^{-1}. Both S_4 and ϕ_{RMS} depend on sunspot number (R) as (1 + 0.04 R) (see Singleton, IV, D1-1; Goodman, IV, D1-50). The power (W) in the scintillation falls off with fading frequency ν according to a power law

$$W \propto \nu^{-p}$$

where

$$2 < p < 3$$

This frequency exponent for S_4 is valid for weak scatter only which holds when S_4 is less than about 0.5. Goodman (IV, D1-50) reports that scintillation bandwidth is so broad that frequency diversity is a practical impossibility as is polarization diversity but that time and space diversity are promising in overcoming the effects of scintillation on radio systems. However, space diversity, requiring separations of the order of 1000 m, cannot readily be accommodated on a ship. The only other solution is to go to higher carrier frequencies (e.g. $\gtrsim$ 10 G Hz). Time diversity especially for aircraft usage, seems to be the best solution.

MODELLING

There are two basic categories of models namely: (1) numerical or empirical models and (2) physical models. Numerical maps of ionospheric parameters are examples of empirical models. Physical models are based on ion production and loss which in turn requires models of ionizing radiation and neutral atmosphere.

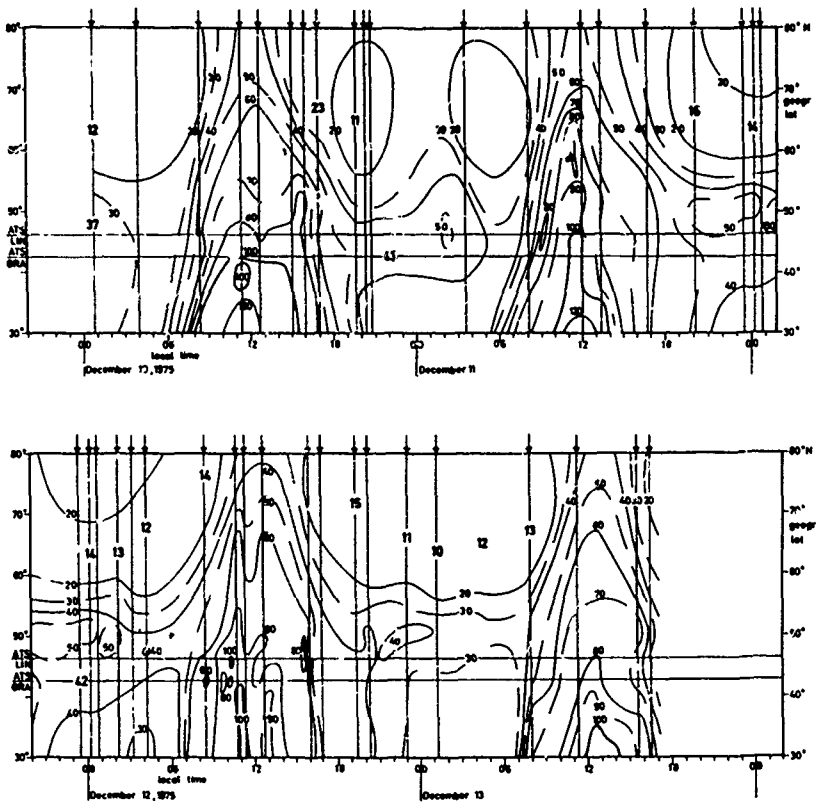


Fig. 6

Contours of electron content in a local time versus latitude grid. The electron content is in units of 10^{15} m^{-2}. Separation of solid lines $\approx 20 \times 10^{15}$ m^{-2}. Obtained from a combination of ATS6 beacon data and NNSS data.

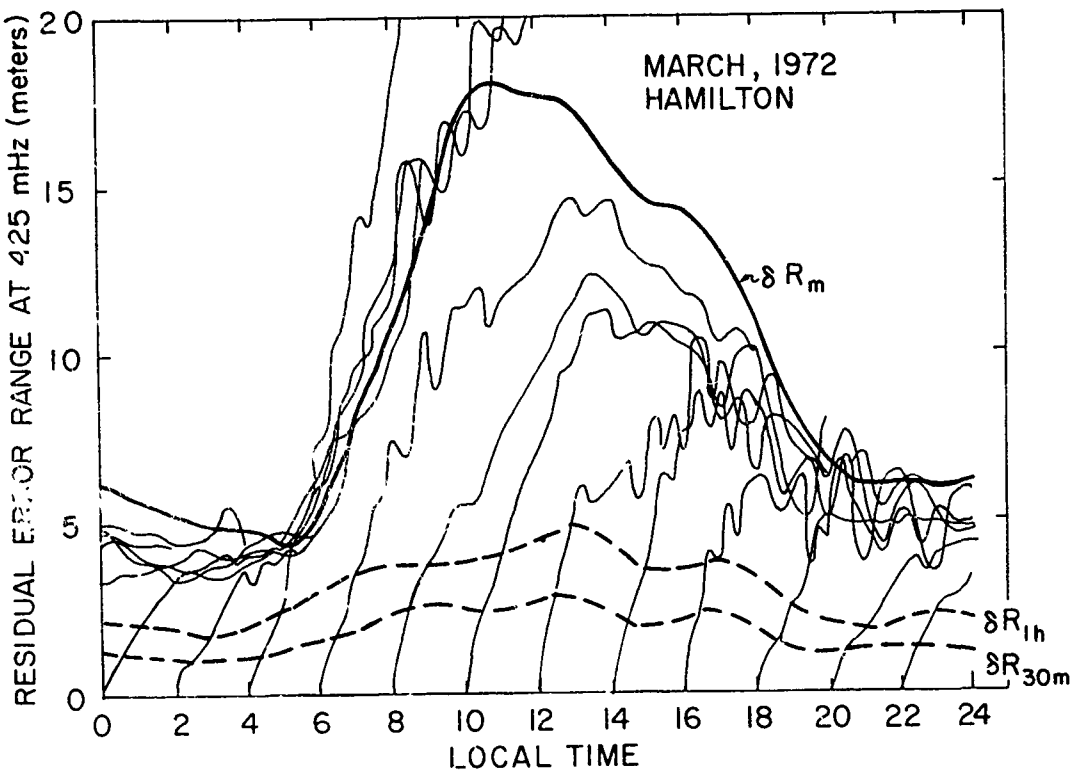


Fig. 7

The residual error in range correction using an updating technique ranging from zero error at the time of calibration to the residual error using only the median correction, δR_m, at a time depending primarily on the interval between calibration and subsequent terminator.

Empirical Models

These, such as the COSPAR International Reference Atmosphere (CIRA 1972) and the International Reference Ionosphere (Rawer et. al., 1978), which provide average height variations of the main atmospheric parameters such as pressure, density, temperature, electron density profiles, etc. The models are based on experimental data using interpolation and extrapolation techniques. The Reference Ionosphere is a new and useful tool but, unfortunately the number of profiles is small and, therefore, the uncertainty large. This is particularly true in regions of high variability such as in high latitudes. Davis and Berry (1977) have produced a D-region model by a best fit to all measured profiles in the literature (McNamara, 1978). This model is usable for VLF and LF phase and amplitude predictions. A list of some electron density profile models is given in Table 1. A model of phase and amplitude scintillation based on in-situ density fluctuation measurements is also available (Basu and Basu, D1-32).

Physical Models

The second approach involves basic physics. Given the flux of ionizing radiation and neutral atmosphere properties the electron profile can be computed. The ionization processes in the ionosphere are fairly well known but selective ionization of minor constituents may dominate the ion production in the D region (see Swider, II, 599). For this reason the transport of minor constituents into and out of the D region may significantly influence the ionization balance. The loss process is very complicated. Above about 85 km ions of molecular oxygen and nitric oxide determine the loss whereas just below this level hydrated positive ions dominate.

In the F region the ion production and loss processes are relatively simple but plasma transport dominates e. g. neutral winds, electromagnetic drift, plasma diffusion.

The use of theoretical and/or semiempirical models for ionospheric prediction was discussed by Vlasov (IV, C-41). In this way one can predict: (1) average plasma properties for radio propagation and (2) neutral atmosphere properties that affect satellite orbits. For such models we need a reliable theory for ionosphere-magnetosphere coupling (e.g. Murphy et al., 1980) and for ionosphere-lower atmosphere coupling.

While empirical models have the advantage of being based on measurements, physical models are much more versatile in dealing with man made modifications (e.g. chemical, radio heating) of the ionosphere. The main value of physical models is in estimating the reaction of the ionosphere to a change in the forcing function, e. g. chemical composition, electric field, etc. They are not likely to replace morphological models as regards accuracy.

IONOSPHERIC FORECASTING

We define ionospheric forecasting as predictions of ionospheric conditions for periods less than one solar rotation (27 days) and especially those effects associated with solar flares and solar-terrestrial disturbances (see Davies and Smith, 1978). These forecasts depend primarily on solar observations (both in the visible and x-ray parts of the spectrum) and are supplemented by ground-based observations of the geomagnetic field and of the ionosphere together with satellite observations of energetic particles.

Heckman (I,322) discussed the prediction activities of the Space Environment Services Center in Boulder, lists the current users of the forecasts, data sources and distribution systems (e.g. telephone, teletype, computer links, WWV broadcasts, mail). A summary of users of the SESC and types of activity affecting their systems is given in Table 2. Prediction products include: geomagnetic indices, proton events (polar cap absorption), radiation level, solar flares (onsets, intensities, durations), sudden storm commencements, etc.

Similar radio warnings are issued in Japan (Maeda, I, 223; Marubashi, I,182) by the Tokyo Regional Warning Center in which forecasts are continually updated and issued 12 hours in advance based mainly on ground-based data. Daily and weekly forecasts are also made and distributed via the International Ursigram and World Data Service (IUWDS) and broadcast on station JJD. The forecasts are used extensively for ship-to-shore communications, satellite tracking, control and telemetry. The Japanese predict also a MAGNIL indicating the end of a disturbance. Increases in signal strength prior to a sudden commencement (10 to 20 dB for $\tilde{>}$ 10 hours) accompanied by increases in f_0F2 are used to predict communication disturbances on polar paths (Ondoh, IV, D2-21).

There is little doubt that the major advances in ionospheric forecasting in the past 20 years has been brought about by technological improvements in data aquisition (e.g. satellite sensors), data processing (computers) and dissemination. This is illustrated by the geophysical forecasting at the U.S. Air Force Global Weather Center (Thompson, I,350) where the major operational advances include:

Table 1

Some Electron Density Profile Models for E and F layers *

Model (year)	E Region	F1 Region	F2 Region	Remarks
CCIR (1970)	Parabola		Parabola	
Damon and Hartranft (1970)	Chapman Layer	Chapman Layer	Chapman Layer	
Nisbet: Penn State MKI (1970)	Solution of Equation of Continuity		Solution of Diffusion Equation	Daytime Model
Bent (1972)	Parabola Squared	Parabola Squared	Parabola	Topside: Exponential Decay Functions
Cookingham (1972)	From Nisbet	From Nisbet	Chapman, Parabola, or Sine Squared Function	
Bradley and Dudeney (1973)	Parabola below E layer max	Linear up to level of 1.7 foE	Parabola	
Ching & Chiu (1973)	Phenomenological	Phenomenological	Phenomenological	
Miller and Gibbs (1973)	Parabola	Parabola	$sech^2$ in daytime parabola at night	Topside: Exponential Decay functions
Rush and Miller (1973)	Parabola	Parabola	$sech^2$ in daytime parabola at night	Topside: sech function
Jasperse (1976)	Solution to electron distribution function			Daytime Bottomside Model

*From Tascione, I, 367, consult this paper for references.

Table 2

Summary of users of the SESC and types of activity affecting their systems *

Customer	Type of Activity Producing Effect
Civilian satellite communication	Magnetic storms
Commercial Aviation--mid-latitude communication (VHF)	Solar radio emissions
Commercial aviation--polar cap communication (HF)	PCA, magnetic storms
Commercial aviation navigation (VLF)	PCA, magnetic storms
Electric power companies	Magnetic storms
Long lines telephone communication	Magnetic storms
High altitude polar flights radiation hazards	Solar proton events
Civilian HF communication Coast Guard, GSA, commercial companies, VOA	X-ray emission, U.V. emission, magnetic storms
Geophysical exploration	Magnetic storms
Satellite orbital variation military and civilian	U.V. emission, magnetic storms
DoD SATCOM communication	Magnetic storms
DoD HF communication	X-ray emission, U.V. emission, PCA, magnetic storms
DoD reconnaissance	PCA, magnetic storms
DoD navigation	X-ray emission, U.V. emission
ERDA communication prospective customers	X-ray emission, U.V. emission magnetic storms
International community	All
Scientific satellite studies IMS, Solar Maximum mission, Shuttle, solar constant measurements, stratospheric ozone variation, interplanetary missions	Optical solar flares, magnetic storms, X-ray emission, U.V. emission, solar proton events, solar features
Scientific rocket studies IMS, magnetosphere, ionosphere upper atmosphere, sun	Optical solar flares, solar features, magnetic storms, solar proton emission, X-ray emission
Scientific ground studies IMS, sun, interplanetary, magnetosphere, ionosphere, upper atmosphere, stratosphere, troposphere, seismological/geomagnetic	Optical solar flare, magnetic storms, solar proton emission, X-ray emission, U.V. emission, solar features

*From Heckman, I, 322

(1) the solar observing optical network and (2) the radio solar telescope network and (3) in the use of a method for data processing and handling developed by meteorologists. The technological advance is illustrated also by the PROPHET system (Argo I, 312) which is tailored to the needs of specific users. It is found that the use of PROPHET is more beneficial than signal coding and information processing. In Australia, a computer presentation called GRAFEX (Turner, IV, D2-85) converts the predictions to a form suitable for a fast printer. The KAZIA system, used in Poland, employs a computer code for processing URSIgram data for forecasting ionospheric activity and radio circuit performance (Stasiewicz I,61).

In the USSR (Avdyushin I, 104) three principal predictions are issued: (a) short term predictions of ionospheric and magnetic disturbances (b) long term maps of MUF etc. and (c) long term predictions of radio propagation along fixed paths. Predictions are issued: monthly; 5 days; 2 days; and twice daily. The monthly forecast relies mostly on recurrence, whereas in the shorter time-scales the emphasis is on real events. Soviet forecasts include: (1) changes in F2 critical frequency (see Kuleshova, IV, C-37), (2) auroral absorption (see Zevakina, IV, D3-14) and (3) polar cap absorption (Akinyan III, D-14).

Attempts to improve day-to-day predictions based on a combination of magnetic K index and 10.7 cm solar radio flux (Wilkinson, I, 259) show little improvement in the "noise level". Predictions of foF2 one day ahead are disappointing while extension of prediction to geographically remote regions show no substantial improvements.

CONCLUDING REMARKS

The Solar-Terrestrial Predictions Workshop was a success from the viewpoint of bringing together predictors, forecasters, users and researchers in the field for the first time.

Because the field of ionospheric radio involving reflected waves is relatively old, the advances reported at the workshop were of second order. Some of these are valuable as, for example, the development of four-dimensional ionospheric models, and the incorporation (at long last) of time dependent features (e.g. troughs, auroral zones, clefts, etc.) into global ionospheric maps. This extension of numerical mapping is of considerable value to those concerned with radio ray tracing. On the other hand considerable advances have been made in our knowledge of the temporal, spatial and storm behavior of ionospheric characteristics involved in transionospheric radio, e.g. Faraday content, plasmaspheric content and scintillation. Short-term ionospheric predictions have been improved by confining attention to localized geographical areas e.g. Europe (Lassudrie-Duchesne, I,12) and India (Reddy, I,118; Aggarwal I,134).

Improvements in recent years in ionospheric forecasting are due mainly to efficient data acquisition and assessment e.g. "PROPHET" (Argo I,312). Thus warning services, as distinct from services prediction, have undergone considerable improvement in recent years by the installation of systems such as SELDADS (Williams, 1976) which employ satellite communications systems. Real time monitoring of the ionosphere-magnetosphere system is done using both satellite and ground-based sensors: particle detectors, DMSP optical auroral photography, plasma probes, auroral radars, Faraday polarimeters, riometers, ionosondes. These sensors enable timely warnings to be disseminated.

Solar flare forecasting -- a vital ingredient of magnetic/ionospheric storm-forecasting -- is still largely an art in which the forecasters' experience and skill in recognizing and interpreting solar features plays the dominant role. This expertise has been developed to a fairly high level and a forecaster can often forecast that a flare will occur within one to three days but without being precise about just when the flare will occur, its magnitude nor its terrestrial consequences.

Likewise, forecasting the onset, duration and magnitude of an ionospheric storm is still, to a large extent, dependent on personal intuition. One aspect of this, that was repeatedly emphasized at the STP Workshop, was the role of the magnetosphere in ionospheric disturbance, e.g. energy input at the magnetospheric cleft, input into the auroral zone of particles energized by electric fields parallel to the geomagnetic field or by Alfvén waves, plasma convections over the polar cap which maintain the winter polar ionosphere. In the immediate future real-time forecasts of polar propagation will be best made by the use of satellite sensors (see Cauffman, III, B-48) to identify positions and intensities of large scale features e.g. areas of polar-cap absorption and auroral absorption, sporadic E, small scale irregularities.

Another aspect that received considerable emphasis was the role of the neutral atmosphere. Joule heating of the neutral atmosphere by auroral currents produces circulation that results in stronger winds and chemical changes (richer in molecules). The former leads to a lifting of the ionosphere in middle latitudes and consequently a lower decay rate of the plasma while the latter increases the loss rate and produces negative storms that markedly disrupts HF radio communications. The role of neutral air dynamics in influencing the state of the D region was brought out in the meeting. In particular, it is known that the transportation of minor constituents in the middle atmosphere exercizes a

crucial role in the ionization balance in the D region both in middle and in high latitudes.

The four volume proceedings of the STP Workshop provides an invaluable document for future predictors, forecasters, users and scientists working both in ionosphere physics and radio propagation. We owe a debt of gratitude to Dr. R. F. Donnelly both for organizing the meeting and for editing the proceedings.

ACKNOWLEDGMENTS

I wish to express my appreciation to R. F. Donnelly for suggesting that I write this review and for his helpful criticism of the typescript. I am indebted to the following for their constructive comments on the typescript: G. K. Hartmann, R. Hunsucker, J. Klobuchar, C. Rush and E. V. Thrane.

JOURNAL REFERENCES

Aarons, J., E. Mackenzie and K. Bhavnani, 1978, Equatorial and high latitude empirical models of scintillation levels, AGARD Conference Proceedings No. 238, Operational Modelling of the Aerospace Propagation Environment.

Barghausen, A. F., J. W. Finney, L. L. Proctor and L. D. Schultz, 1969, Predicting long-term operational parameters of high-frequency sky-wave telecommunication systems, ESSA Tech. Rept. ERL-110-ITS 78.

Bernhardt, P. A., 1979, Digital processing of ionospheric electron content data, IEEE Trans. ASSP-27, 705.

Bradley, P. A., 1975, Long-term HF propagation predictions for radio circuit planning, The Radio and Electronic Engineer, 45, 31.

CIRA, 1972, COSPAR International Reference Atmosphere, Akademie-Verlag, Berlin.

CCIR, 1964, World distribution and characteristics of atmospheric noise, Report 322, International Telecommunications Union, Geneva.

CCIR, 1978, CCIR Atlas of Ionospheric Characteristics, Report 340-3, Documents of 14th Plenary Assembly, Kyoto, International Telecommunications Union, Geneva.

CCIR, 1978a, VHF propagation by regular layers, sporadic-E and other anomalous ionization, Report 259-3, International Telecommunications Union, Geneva.

Costa, E. and M. Kelley, 1976, Calculation of equatorial scintillation at VHF and gigahertz frequencies based on a new model of the disturbed equatorial ionosphere, Geophys. Res. Letters, 3, 677.

Davies, K., 1980, Recent advances in satellite radio beacon studies with particular emphasis on the ATS-6 radio beacon experiment, Space Sci. Rev., 25, 357.

Davies, K. and E. K. Smith, 1978, Ionospheric Predictions, in Geophysical Predictions, 82-95, National Academy of Sciences, Washington, D. C.

Davies, K., G. K. Hartmann and R. Leitinger, 1977, A comparison of several methods of estimating the columnar electron content of the plasmasphere, J. Atmos. Terr. Phys. 39, 571.

Davis, R. M. and L. A. Berry, 1977, A revised model of the electron density in the lower ionosphere, Technical Report TR111-77, Defense Communication Agency.

Donnelly, R. F., 1979, Solar-Terrestrial Predictions Proceedings, U. S. Gov't Office, Washington, D. C. 20402.

Dyson, P. L. and J. A. Bennett, 1979, General formulae for absorption of radiowaves in the ionosphere, J. Atmo. Terr. Phys. 41, 367.

Elkins, T. and C. Rush, 1973, A statistical predictive model of the polar ionosphere, in An Emperical Model of the Polar Ionosphere, editor T. Elkins, AFCRL-TR-73-0331, Air Force Cambridge Research Laboratories, Hanscom AFB, MA.

Fremouw, E. J. and C. L. Rino, 1978, A signal-statistical and morphological model of ionospheric scintillation, in AGARD Conference Proceedings No. 238, Operational Modelling of the Aerospace Propagation Environment.

George, P. L., 1971, The global morphology of the quantity Nvdh in the D- and E-regions of the ionosphere, J. Atmos. Terr. Phys., 33, 1893.

Gregory, J. B., and A. H. Manson, 1969, Seasonal variations of electron densities below 100 km at mid-latitudes- II. Electron densities and atmospheric circulation, J. Atmos. Terr. Phys. 31, 703.

Hargreaves, J. K., 1969, Conjugate and closely spaced observations of auroral absorption - I, Planet. Space Sci., 17, 1459.

Jones W. B. and R. M. Gallet, 1962, Representation of diurnal and geographic variations of ionospheric data by numerical methods, J. Res. NBS 66D, 419.

Jones, W. B. and F. G. Stewart, 1970, A numerical method for global mapping of plasma frequency, Radio Sci., 5, 773.

Jones, W. B., R. M. Gallet, M. Leftin and F. G. Stewart, 1973, Analysis and representation of the daily departures of the foF2 from the monthly median OT Report 73-12, U.S. Department of Commerce.

Kersley, L., H. Hajeb-Hosseinieh and K. J. Edwards, 1978, Post geomagnetic storm protonospheric replenishment, Nature, 271, 429.

Leftin, M., 1976, Numerical representation of monthly median critical frequencies of the regular E region (foE), OT Report 76-88, U.S. Department of Commerce/Office of Telecommumications.

Leftin, M. and S. M. Ostrow, 1969, Numerical maps of f_bE_s for Solar Cycle Minimum, ESSA Technical Report ERL 124-ITS87.

McNamara, L., F., 1978, Ionospheric D-region profile data base, Report UAG-67, World Data Center A, Boulder, CO 80303.

Miller, D. C. and J. Gibbs, 1974, Ionospheric analysis and ionospheric modelling, AFCRL-TR-0364, Air Force Cambridge Research Laboratories, Hanscom AFB, MA 01731.

Muggleton, M. L., 1975, A method of predicting foE at any time and place, Telecommun. J., 42, 413.

Murphy, J. A., G. J. Bailey and R. J. Moffett, 1980, A theoretical study of the effect of quiet-time electromagnetic drifts on the behavior of thermal plasma at mid-latitudes, J. Geophys. Res., 85, 1979.

Offerman, D., 1977, Some results from the European Winter Anomaly Campaign 1975/6, in Dynamical and Chemical Coupling of Neutral and Ionized Atmosphere, ed. B. Grandel and J. A. Holtet, D. Reidel Publ. Co., Dordrecht, Holland, page 235.

Rawer, K., S. Ramakrishnan and D. Bilitza, 1978, International Reference Ionosphere 1978, URSI, Brussels, Belgium.

Rino, C. L., 1979a, A power-law phase screen model for ionospheric scintillation (1) Weak scatter, Radio Sci., 14, 1135, 1979b (2) Strong scatter, Radio Sci., 14, 1147.

Rosich, R. K. and W. B. Jones, 1973, The numerical representation of the critical frequency of the F1 region of the ionosphere, OT Report 73-22, U. S. Department of Commerce/Office of Telecommunications.

Swanson, E. R., 1977, Propagation effects on OMEGA AGARD Conf. Proc. No. 209.

Vondrak, R. R., G. Smith, V. E. Hatfield, R-T. Tsunoda, V. R. Frank and P. D. Perreault, 1977, Chatanika model of the high-latitude ionosphere for application to HF propagation prediction, Rome Air Development Center Technical Report RADC-TR-78-7, U.S.A.

Wakai, N., 1971, Study on the night-time E region and its effect on the radio wave propagation, Journal of Radio Research Labs. (Japan), 18, 245.

Williams, D. J., 1976, SELDADS, NOAA Tech. Rep. ERL 357, SEL 37. U. S. Government Printing Office, Washington, D. C.

List of References from Solar Terrestrial Predictions Proceedings, Volume I, II, III and IV, editor, Richard F. Donnelly, Space Environment Laboratory, Environmental Research Laboratories, National Oceanic and Atmospheric Administration of the U. S. Department of Commerce and for sale by the Superintendent of Documents, U. S. Government Printing Office, Washington, D. C., 20402 (Stock No. 003-023-00041-9). An entry (II, 217) signifies Volume II, page 217.

Aggarwal, S., D. R. Laksmi and B. M. Reddy, A simplified indexing of F-region geophysical noise at low latitudes, (I,134).

Akinyan, S. T., I. M. Chertok and V. V. Fomichev, Quantitative forecasts of solar protons based on solar flare radio data, (III, D-14).

Anufrieva, T. A., T. L. Gulyaeva, G. K. Kadukhin, Prediction of the parameters of the maximum of the vertical electron density gradient, (IV, C-57)

Argo, P. E. and I. J. Rothmuller, PROPHET: an application of propagation forecasting principles, (I, 312).

Avdyushin, S. I., A. D. Danilov, A. B. Malishev, G. N. Novikova and P. N. Svidsky, Forecasting ionospheric and geomagnetic conditions at the IAG Forecasting Center (I, 104).

Basu, S. and S. Basu, Model of phase and amplitude scintillations from in-situ measurements, (IV, D1-32).

Besprozvannaya, A. S., A. V. Shirochkov, and T. I. Shchuka, On the approach to the forecast of the polar ionosphere conditions, (II, 528).

Bhatnagar, A. K. and M. Sain, Atmospheric radio noise measurements in LF, MF bands, (IV, D3-55).

Cauffman, D. P., Space plasma monitoring: two new capabilities for the coming decade (III, B-48).

Damholt Th., Propagation predictions for the HF range by the Research Institute of the Deutsch Burdepost, (I, 25).

Donatelli, D. E. and R. S. Allen, Ionospheric refractive correction using an adaptive procedure (IV, D1-65).

Giraldez, A. E., Daytime sporadic-E blanketing frequency predictions, (IV, C-87).

Goodman, J. M., A resume of anticipated Fleetsatcom and Gapfiller scintillation effects during the peak of solar-cycle 21 (1980-1982), (IV, D1-50).

Hatfield, V. E., HF communications predictions 1978 (An economical up-to-date computer code, ABCOM) (IV, D2-1).

Heckman, G., Predictions of the Space Environment Services Center (I, 322).

Hunsucker, R. D., High-latitude E- and F- region ionospheric predictions (II, 513).

Ivanov-Kolodny, G. S. and A. A. Nusinov, Forecast of critical frequency and height of maximum forecast of mid-latitude E-layer (IV, C-82).

Kerblay, T. S., E. M. Kovalevskaya, E. M. Zhulina and L. M. Ishkova, Prediction of the characteristics of radio signal reflected from horizontally-inhomogeneous ionosphere and the relevant requirements for prediction of ionospheric parameters (IV, D2-65).

Kerblay, T. S. and G. S. Nosova, Statistical prediction of E_s-layer parameters and echo-signal characteristics (IV, C-77).

Klobuchar, J., Transionospheric propagation predictions (II, 217).

Klos, Z. and K. Stasiewicz, Estimation of solar flare influence on radio circuit transmission loss (I, 67).

Kotadia, K. M., A. Gupta and R. M. Kotak, Determination of solar cycle variation of HF radio wave absorption at low latitude (IV, D3-20).

Kovalevskaya, E. M. and E. M. Zhulina, On the statistical properties of the disturbed high-latitude ionosphere in radio wave propagation computations (IV, D2-16).

Kuleshova, V. P., E. V. Lavrova and L. N. Lyakhova, Forecasting of ΔfoF2-variations for ionospheric disturbances (IV, C-37).

Lussudrie-Duchesne, P. A. M. Bourdilla and H. Sizun, The French short term radio propagation predictions in the decameter band (I, 12).

Larsen, T. R., Effects of D-region ionization on radio wave propagation (II, 617).

Maeda, R., Radio propagation prediction services in Japan (I, 212).

Maeda, R. and H. Inuki, Radio disturbance warning issuance system (I, 223).

Marubashi, K., Y. Miyamoto, T. Kidokoro and T. Ishii, Forecasts of geomagnetic activities and HF radio propagation conditions made at Hiraiso, Japan (I, 182).

Meisel, D. D., B. Duke and W. D. Savedoff, Minicomputer simulation of ionospheric radiowave propagation at decametric wavelengths (IV, D2-31).

Mendillo, M. and J. A. Klobuchar, A morphology-based prediction scheme for the coupled latitudinal and local-time development of F-region storms (IV, C-15).

Mendillo, M., F. X. Lynch and J. A. Klobuchar, Geomagnetic activity control of ionospheric variability (IV, C-1).

Mitra, A. P. Communications predictions (II, 203).

Murthy, P.S.N., C. S. R. Rao and M. Sain, Prediction of foF2 by the monthly ratio (MR) method (IV, D2-54).

Ondoh, T., and K. Obu, Prediction of HF communications disturbances by pre-sc field increases on polar paths crossing the auroral zone (IV, D2-21).

Oyinloye, J. O., Prediction of radio wave absorption in the ionosphere (IV, D3-1).

Picquenard, A. A. E. and E. R. de Paula, A simplified computer method for long-term calculation of HF sky-wave circuits (IV, D2-41).

Reddy, B. M., S. Aggarwal, D. R. Lakshmi, S. Shastri and A. P. Mitra, Long term solar activity and ionospheric prediction services rendered by the National Physical Laboratory, New Delhi (I, 118).

Reddy, B. M., Ionosphere-reflected propagation (II, 203).

Rush, C., Report of the mid- and low-latitude E and F region working Group (II, 562).

Shlionsky, A. G., Prediction of waveguide propagation of radio waves using the extremal-parametric method on the basis of prediction of the key ionospheric parameters (IV, D3-60).

Singleton, D. G., An improved ionospheric irregularity model (IV, D1-1).

Singleton, D. G., Predicting transionospheric propagation conditions (IV, D1-16).

Spjeldvik, W. N. and L. R. Lyons, On the predictability of radiation belt electron precipitation into the earth's atmosphere following magnetic storms (IV, B-59).

Stasiewicz, K., M. Maksymienko and M. Jakimiec, KAZIA -- A computer system for solar-terrestrial data processing and objective prediction (I, 61).

Swider, W., Ion production in the D-region (II, 599).

Tascione, T. F., T. W. Flattery, V. G. Patterson, J. A. Secan and J. W. Taylor, Ionospheric modeling at Air Force Global Weather Central, (I, 367).

Thompson, R. L. and J. A. Secan, Geophysical forecasting at AFGWC (I, 350).

Thrane, E. V., et al., D-region predictions (II, 573).

Turner, J., Grafex predictions (IV, D2-85).

Vlasov, M. N, Fundamentals of the physical forecast of ionospheric plasma (IV, C-41).

Vondrak, R. R., Y. T. Chiu, D. S. Evans, V. G. Patterson, G. J. Romick and K. Stasiewicz, Magnetosphere-ionosphere Interactions (II, 476).

Wilkinson, P.J., Prediction limits for foF2 (I, 259).

Zevakina, R. A., and M. V. Kiseleva, On the short-term prediction of the space-time distribution of auroral absorption (IV, D3-14).

Zevakina, R. A. and E. V. Lavrova, On the possibility to predict variations in the F2-region parameters as a function of the IMF direction (IV, C-27).

TABLES

1. Electron density profile models for E and F layers.
2. Summary of users of the SESC and types of activity affecting their systems.

LIST OF FIGURE CAPTIONS

Fig. 1 The earth's environment

Fig. 2 Polar plots of the probability of auroral absorption (from 30 MHz riometer data) occurrence for values $\geq$ 0.5 dB in winter 1964, in terms of corrected geomagnetic latitude and time (from Besprozvannaya II, 528).

Fig. 3 Average position of the ionospheric trough minimum in the night sector for 25 global magnetic storms versus Dst (see Besprozvannayo, II, 528).

Fig. 4 Numerical map of F2

Fig. 5 Polar plots of occurrence probability of the auroral E in winter. The coordinate grid is corrected geomagnetic latitude and time (from Besprozvannaya et al., II, 528).

Fig. 6 Contours of electron content in a local time versus latitude grid. The electron content is in units of 10^{15} m^{-2}. Separation of solid lines $\approx$ 20 x 10^{15} m^{-2}. Obtained from a combination of ATS6 beacon data and NNSS data.

Fig. 7 The residual error in range correction using an updating technique ranging from zero error at the time of calibration to the residual error using only the median correction, δR_m, at a time depending primarily on the interval between calibration and subsequent terminator.

NEW FORECASTI'.G METHODS OF THE INTENSITY AND TIME DEVELOPMENT OF' GEOMAGNETIC AND IONOSPHERIC STORMS

S.-I. Akasofu
Geophysical Institute
University of Alaska
Fairbanks, Alaska 99701

ABSTRACT

It is shown that geomagnetic storms, magnetospheric substorms and the associated auroral activity and ionospheric disturbances result directly from an enhanced power of the solar wind-magnetosphere dynamo, which is controlled by the solar wind parameter ε. This finding has raised the possibility that by monitoring $\varepsilon(t)$ upstream of the solar wind, one can now accurately forecast the occurrence and intensity of geomagnetic and ionospheric disturbances. In fact, this operation has already been implemented by using the ISEE/C satellite, providing us with ~1 hr lead time forecasting. However, in order to extend the above study in providing a practical and reliable forecasting of geomagnetic and ionospheric disturbances with a sufficiently long lead time (>6 hrs), an entirely different method must be used. Like weather forecasting, it is necessary to simulate numerically the propagation of solar wind disturbances caused by a solar flare and forecast numerically time variations of ε. The resulting ionospheric disturbances can be forecasted by knowing the simple relationship between ε and the AE index, namely $AE(\gamma) \sim \varepsilon(\text{erg/sec})/10^{16}$.

VARIETY OF DEVELOPMENT OF GEOMAGNETIC STORMS

Intense auroral and ionospheric disturbances tend to occur during the main phase of a geomagnetic storm. However, it has long been known that the main phase develops differently from one storm to another. Some storms begin with a very large storm sudden commencement (ssc), but they fail to develop a significant main phase. Such storms are associated only with minor auroral activity and thus with minor ionospheric disturbances. In some other storms, a small ssc is followed by a major main phase. Such storms are associated with intense auroral and ionospheric disturbances. Yet in some other storms, a major "main phase" breaks out without the storm sudden "commencement." Thus, the concept of "average storm" is of little use in forecasting the time development of individual geomagnetic storms. Since the magnitude of an ssc is a measure of an enhanced solar wind pressure ($\Delta p = \Delta(2mnV^2)$) and thus of an enhanced kinetic energy flux (ρV^3), the above findings indicate that geomagnetic storms are not simply caused by the impact of an enhanced solar wind flow. Therefore, it is practically impossible to forecast how a geomagnetic storm develops after a particular solar flare, unless one can find the solar wind parameter which directly controls the development of the main phase of geomagnetic storms. For the same reason, the forecasting of ionospheric disturbances and of the resulting disturbances on radiowave systems based on the "average storm" is bound to fail for individual storms.

SOLAR WIND-MAGNETOSPHERE ENERGY COUPLING FUNCTION

The variety of development of geomagnetic storms indicates that geomagnetic storms are not caused simply by an impact of an intense solar wind flow. After a long search, we have finally found the solar wind quantity which controls directly the development of the main phase of geomagnetic storms (Perreault and Akasofu, 1978). This quantity is given by

$$\varepsilon = VB^2 \sin^4\left(\frac{\theta}{2}\right) \ell_o^2$$

where V and B are the solar wind speed and magnetic field intensity, respectively, θ denotes the polar angle of the solar wind magnetic field projected onto the dawn-dusk plane, and ℓ_o is a constant (~7 earth radii). Furthermore, Kan et al. (1980) showed that ε can be identified as the power generated by the solar wind-magnetosphere dynamo. With these findings, it is possible to redefine a geomagnetic storm and auroral activity as follows:

Geomagnetic disturbances are caused by the magnetic fields of electric currents which are generated by the solar wind-magnetosphere dynamo. When the dynamo power exceeds 10^{18} erg/sec (10^{11} watts), the disturbance magnetic fields are intense enough to be recognized as the substorm fields. The excitation and ionization of upper atmospheric particles by the current-carrying electrons flowing down along the geomagnetic field lines also becomes appreciable. A variety of phenomena associated with these processes are called the auroral and ionospheric substorms (Akasofu, 1968; 1977). When the dynamo power exceeds 10^{19} erg/sec (10^{12} watts), the disturbance magnetic fields are intense enough to be recognized as the storm fields. Auroral and ionospheric substorms are considerably intensified during the storm period.

Here, we examine how the solar wind-magnetosphere coupling function ε controls the development of geomagnetic storms. The development can be monitored quantitatively by two geomagnetic indices, the Dst and AE. The Dst index is obtained by averaging longitudinally the disturbance field in low latitudes and is a good measure of the magnetic field of the ring current belt (the storm-time radiation belt) at any given time. The intensity of geomagnetic storms is defined in terms of Dst; $|Dst| < 50\gamma$, a weak disturbance; $|Dst| \sim 100\gamma$, a moderate storm; $|Dst| > 200\gamma$, a major storm. The AE index is constructed by using magnetic records from a number of auroral zone stations and is a measure of the intensity of the concentrated electric currents which flow in the auroral ionosphere. By using the Dst and AE indices, one can estimate the rate U_T of the total energy dissipated in the inner magnetosphere and the ionosphere (Akasofu, 1981). Most of the disturbance phenomena are manifestations of the energy dissipation processes, U_T becomes large for an intense geomagnetic storm.

Figure 1 shows, from the top, the kinetic energy flux $K = \rho V^3 \ell_o^2$ of the solar wind, the solar wind-magnetosphere energy coupling function ε, the rate U_T of the total energy dissipated in the inner magnetosphere and the ionosphere and the two geomagnetic indices, AE and Dst for the geomagnetic storm of March 31 - April 3, 1973. One can see that U_T is closely correlated with ε, but not with K, indicating that the intensity of the main phase is controlled by the power generated by the solar wind-magnetosphere dynamo, not simply by an intense impact of the solar wind flow. Note that an intense main phase began to develop when ε exceeded ~10^{19} erg/sec.

Figure 2 shows another example of the set K, ε, U_T, AE, and Dst for the storm of January 18-21, 1973. In this example, it is interesting to see that ε and U_T were very small for more than 12 hours after the storm sudden commencement. One can easily infer that during the first 12 hours after the ssc, there was no significant ionospheric disturbance, since the AE index was extremely low. The main phase began to develop only after 14 UT when ε increased significantly. There were three major increases of the AE index, corresponding to three increases of ε. It is not difficult to expect that the ionosphere was greatly disturbed three times during the storm, corresponding to the increases of ε.

Figure 3 illustrates what we might call a "mini-storm" which was associated with an intense impact of the solar wind flow, but a major main phase failed to develop (see the Dst index). The reason for this failure is quite obvious. This particular solar wind flow contained only a small amount of ε; compare ε in Figures 1, 2, and 3. In many events, such storms are associated with a large value of VB^2, since a blast wave from a solar flare can produce a large value of both V and B. However, if the vector $\underline{B}$ has a very large northward component, θ becomes very small, making ε rather small, and thus resulting in only a weak storm.

With this preparation, one can understand why there is so much scatter in the relation between the maximum Dst value for individual storms and the central meridian distance of the responsible solar flares, as Figure 4 indicates. One can infer that the envelope of the points in Figure 4 indicates approximately the dependence of the maximum value of the quantity VB^2 on the central meridian distance. The scatter of the points for a given central meridian distance comes

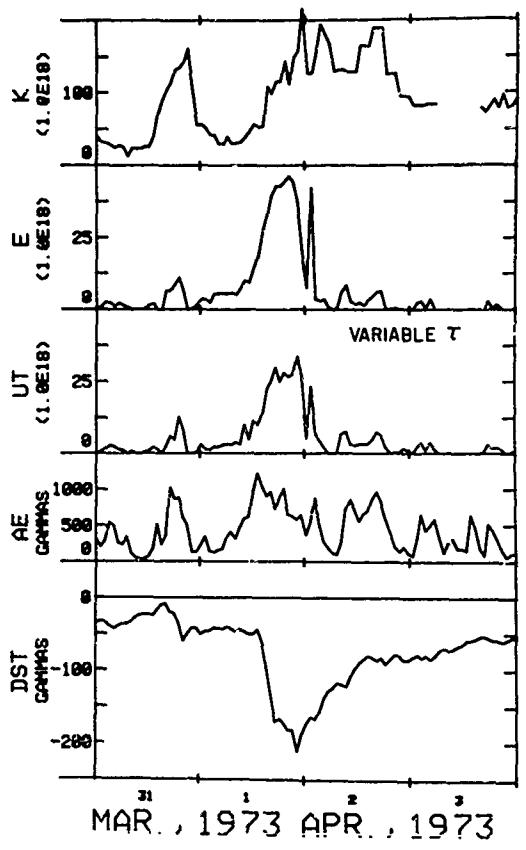


Fig. 1. From the top, the kinetic energy flux $K = \rho V^3 \ell_o^2$ of the solar wind, the solar wind-magnetosphere energy coupling function ε, the rate U_T of the total energy dissipated in the inner magnetosphere and the ionosphere, and the two geomagnetic indices AE and Dst for the storm of March 31 - April 3, 1973.

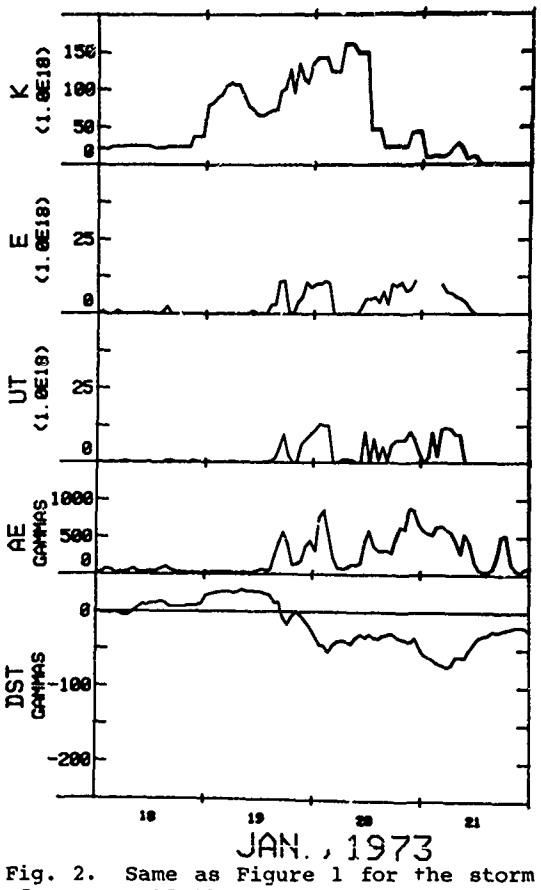


Fig. 2. Same as Figure 1 for the storm of January 18-21, 1973.

partially from the fact that θ varies considerably from one event to another. Therefore, from the forecasting point of view, we have no way of predicting the intensity of storms for an intense central meridian flare, unless one can infer how θ and ε vary in time in the enhanced solar wind flow.

From the above three examples, one can easily see that it is essential to monitor ε in forecasting the occurrence and intensity of geomagnetic storms. The kinetic energy flux K has little relation to the storm intensity, although it may be closely correlated with the intensity of solar flares. Once $\varepsilon(t)$ can be monitored, time variations of the AE index and the time development of auroral and ionospheric substorms can easily be inferred by a simple formula $AE(\gamma) \sim \varepsilon/10^{16}$; for example, for $\varepsilon \sim 10^{19}$ erg/sec, AE ~ 1000γ. The maximum Dst index during a geomagnetic storm is approximately related to the maximum ε, denoted by ε_{max}, by $|Dst| = 60(\log \varepsilon_{max} - 18)^2 + 25$. As far as ionospheric disturbances are concerned, the Dst index is an important parameter which determines the size of the auroral oval. The minimum geomagnetic latitude L_m of the auroral oval in the midnight sector is given approximately by $L_m \sim 60^o - (15^o/400)\ |Dst(\gamma)|$ for $|Dst| > 50\gamma$ (Akasofu and Chapman, 1963).

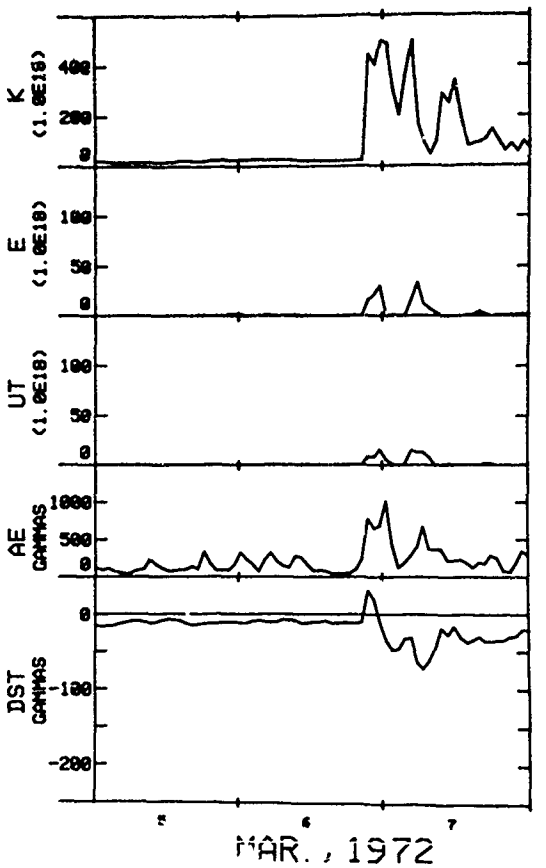


Fig. 3. Same as Figure 1 for the storm of March 5-7, 1972.

FORECASTING THE OCCURRENCE AND INTENSITY OF GEOMAGNETIC STORMS AND IONOSPHERIC DISTURBANCES

From the above study, it is obvious that one can reliably forecast the time development of a geomagnetic storm and the associated ionospheric disturbances, if the solar wind-magnetosphere energy coupling function ε can be predetermined as a function of time. There are obviously two ways of determining $\varepsilon(t)$. The first one is to monitor ε directly by a satellite or a space probe at the front of the magnetosphere. Figure 5 shows an example of such attempts (Akasofu and Chao, 1979). A medium intensity geomagnetic storm occurred at the time when the Mariner 5 space probe was located at a distance of about 460 earth radii from the earth on its way to Venus. At that time, the solar wind disturbance was also monitored by an earthbound satellite, Explorer 34. In Figure 5 we plot hourly average values of ε at Mariner 5 and Explorer 34 by shifting the Mariner data by 3.4 hours to allow for the transit time of ε; for the locations of the two spacecraft with respect to the magnetosphere, see the insert in Figure 5. Similarity of ε at the two locations suggests that ε was well retained during its traverse from Mariner 5 to Explorer 34, so that the forecasting of geomagnetic disturbances is possible. In this regard, we note that the ISEE/C satellite is located at the libration point, a geocentric distance of 1.5×10^6 km (~235 earth radii) toward the sun. By relaying the monitored data directly to the earth, it is possible to make a reliable forecast of the time development of a geomagnetic storm with a lead time of ~1 hr, assuming that ε "propagates" along the sun-earth line. In fact, such a forecasting system is now operative (Tsurutani and Baker, 1977).

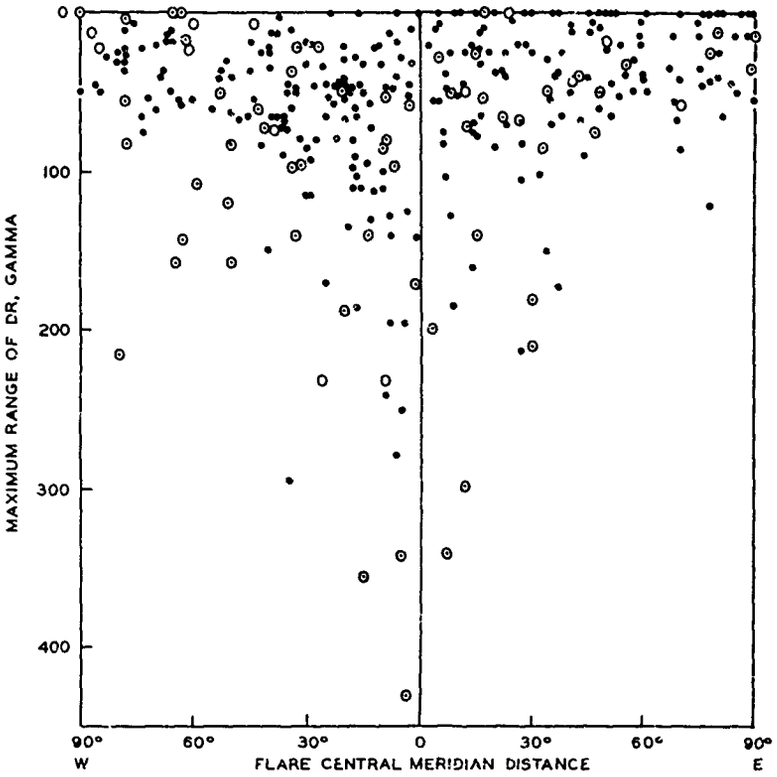


Fig. 4. Relationship between the maximum Dst (=DR) and the central meridian distance of the responsible solar flares. The circled dots are associated with solar proton events (Akasofu and Yoshida, 1967).

From a practical point of view, however, a lead time on the order of ~1 hr is too short to provide an adequate warning to many radiowave systems which will be affected by ionospheric disturbances. Unfortunately, the libration "point" is the only location where a space probe can be placed semi-permanently. Therefore, one must find an entirely new method in predetermining ε as a function of time with a lead time of longer than 6 hrs. One of the ways to achieve such a goal is to forecast numerically $\varepsilon(t)$, like numerical weather forecasting. Indeed, the task of numerically forecasting the time development of a geomagnetic storm is no longer an impossible one, since ε is the only parameter to be determined.

It is not certain if the global weather can be monitored by a single parameter like ε; so far meteorologists have not found one. On the other hand, the difficulty of forecasting ε is due to the fact that we must be able to simulate solar wind disturbances in space between the sun and the earth, without having observed values of V, B and θ. Fortunately, however, there has already been a considerable amount of effort made in this endeavor (Dryer, 1974; Wu et al., 1976). Most of these studies have so far been limited to simulate changes of V and B in the equatorial plane. In order to forecast ε numerically, we must extend such studies to a three-dimensional simulation or at least to a meridian plane simulation. The initial conditions can be set by the intensity of a solar flare, the location of a solar flare on the solar disk, and the location of the earth with respect to the solar current disk. The last point may require some explanation. It has been found that the sun has an extensive electric current disk (Schulz, 1973; Saito, 1975; Svalgaard and Wilcox, 1976; Smith et al., 1978). Furthermore, the current disk is greatly warped. As a result, as the sun rotates with the period of 27 days, the earth

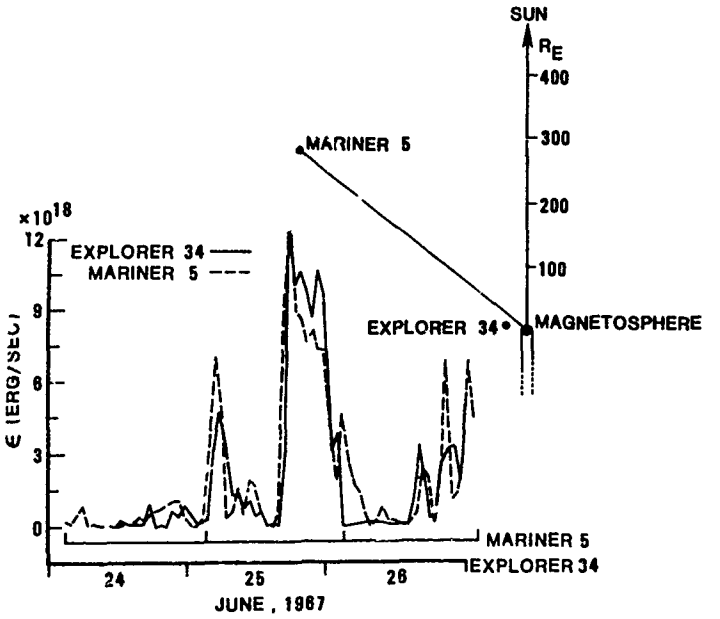


Fig. 5. Comparison of hourly values of ε monitored by the space probe Mariner 5 and the earthbound satellite, Explorer 34. The Mariner 5 data are shifted by 3.4 hours to allow for the transit time of ε.

changes its location with respect to the current disk. Figure 6 shows schematically this situation. In a simple situation, the earth is located above the current disk for 27/2 days and below it for 27/2 days; the current disk passes the earth's location twice during the 27 day period.

During the declining epoch of the sunspot cycle, a medium intensity storm tends to take place during a short period (±2 days) around the passage of the current disk. Weak geomagnetic disturbances follow the storm for about one week after the passage of the current disk (Wilcox and Ness, 1965). This is because both B and V tend to be large near the current disk. However, the actual occurrence of geomagnetic storms depends also on changes of θ. A geomagnetic storm will occur when B, V and θ happen to become simultaneously large. It appears that the warped current disk has various wavy features, causing θ to change. This situation will be simulated in the near future.

During the maximum epoch of the sunspot cycle, a major geomagnetic storm tends to occur also during the passage of the current disk, as well as during the rest of the period. However, even during the latter case, the current disk undergoes a temporal excursion at the location of the earth. It is not difficult to infer that a solar flare and the resulting solar wind disturbances cause a violent up-down motion of the current disk. Figure 7 shows a series of snapshots (the meridian cross-section) of the deforming solar current disk as the blast wave propagates outward. The resulting deformation of the current disk may be one of the causes for large changes of θ. If the resulting change increases θ and if this period coincides with the period of a large value of B and V, a major storm can break out. On the other hand, a storm will be weak if θ becomes small during the passage of the current disk.

Acknowledgments. The work reported here was supported in part by the U.S.A.F. Contract F19628-79-C-0067.

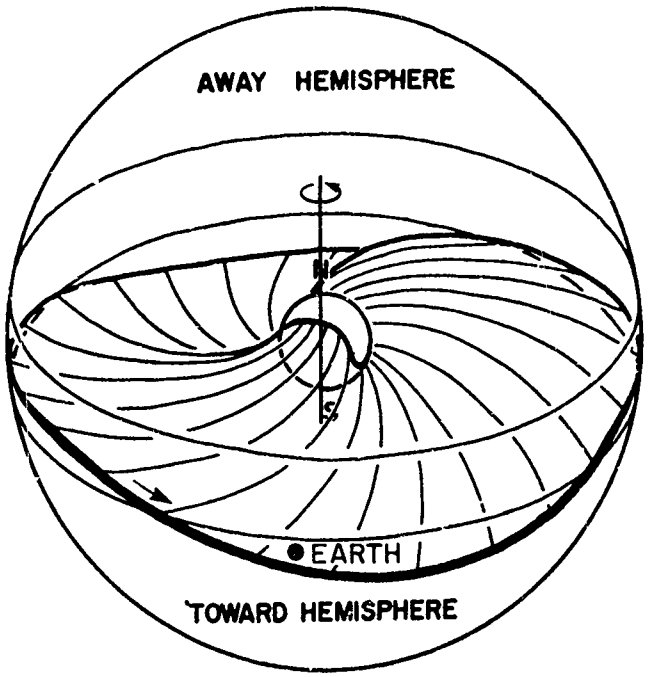


Fig. 6. The warped solar current disk and the earth. As the sun rotates with the period of 27 days, the earth's location with respect to the current disk varies.

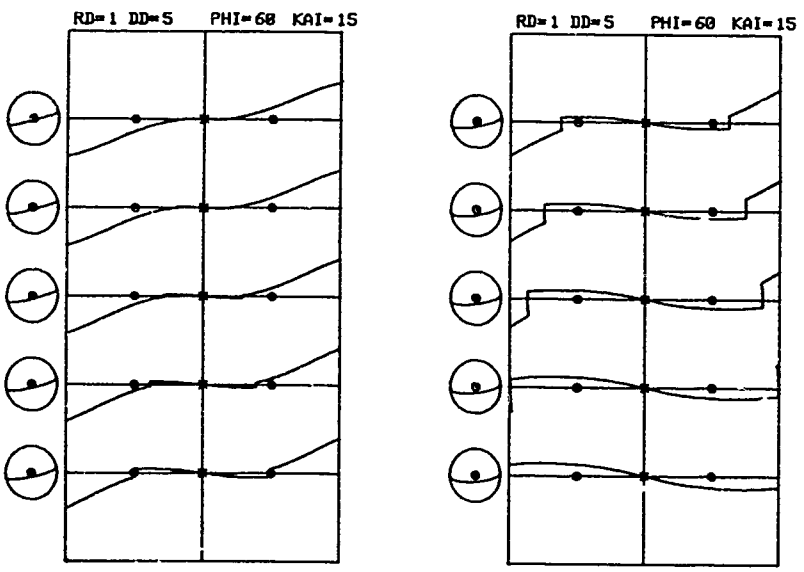


Fig. 7. A series of snapshots (the meridian cross-section) of the deforming solar current disk as the blast wave from a solar flare propagates outward.

References

Akasofu, S.-I., Polar and Magnetospheric Substorms, D. Reidel Pub. Co., Dordrecht, Holland, 1968.

Akasofu, S.-I., Physics of Magnetospheric Substorms, D. Reidel Pub. Co., Dordrecht, Holland, 1977.

Akasofu, S.-I., Energy coupling between the solar wind and the magnetosphere, Space Sci. Rev., (in press), 1981.

Akasofu, S.-I., and S. Chapman, The lower limit of latitude (U.S. sector) of northern quiet auroral arcs, and its relation to Dst (H), J. atmos. Terr. Phys., 25, 9, 1963.

Akasofu, S.-I., and S. Yoshida, The structure of the solar plasma flow generated by solar flares, Planet. Space Sci., 15, 39, 1967.

Akasofu, S.-I., and J. K. Chao, Prediction of the occurrence and intensity of magnetospheric substorms, Geophys. Res. Lett., 6, 897, 1979.

Dryer, M., Interplanetary shock waves generated by solar flares, Space Sci. Rev., 15, 403, 1974.

Kan, J. R., L. C. Lee, and S.-I. Akasofu, The energy coupling function and the power generated by the solar wind-magnetosphere dynamo, Planet. Space Sci., 28, 823, 1980.

Perreault, P., and S.-I. Akasofu, A study of geomagnetic storms, Geophys. J. Roy. astr. Soc., 54, 547, 1978.

Saito, T., Two-hemisphere model on the three-dimensional magnetic structure of the interplanetary space, Sci. Rep. Tohoku Univ., Ser. 5, 23, 37, 1975.

Schulz, M., Interplanetary sector structure and the heliomagnetic equator, Astrophys. Space Sci., 24, 371, 1973.

Smith, E. J., B. T. Tsurutani, and R. L. Rosenberg, Observations of the interplanetary sector structure up to heliographic latitudes of 16°: Pioneer 11, J. Geophys. Res., 83, 717, 1978.

Svalgaard, L., and J. M. Wilcox, Structure of the extended solar magnetic field and the sunspot cycle variation in cosmic ray intensity, Nature, 262, 766, 1976.

Tsurutani, B., and D. Baker, Substorm warnings: An ISEE-3 real time data system, EOS, 60, 701, 1979.

Wilcox, J. M., and N. F. Ness, Quasi-stationary corotating structure in the interplanetary medium, J. Geophys. Res., 70, 5793, 1965.

Wu, S. T., M. Dryer, and S. M. Han, Interplanetary disturbances in the solar wind produced by density, temperature, or velocity pulses at 0.08 a.u., Solar Phys., 49, 197, 1976.

RECENT HIGH-LATITUDE IMPROVEMENTS IN A COMPUTER-BASED SCINTILLATION MODEL

E. J. Fremouw and J. M. Lansinger
Physical Dynamics, Inc.
Bellevue, WA

ABSTRACT

An empirical model of the mean morphological behavior of ionospherically produced scintillation has been updated and improved. The model code employs the phase-screen propagation theory to calculate the power-law spectral index and strength (at a fluctuation frequency of 1 Hz) of phase scintillation and the standard deviation of phase for a user-specified system. For most applications the code employs an effectively infinite ionospheric outer scale, but means have been provided for including the effect of a finite outer scale for systems requiring extended phase stability. The intensity scintillation index, S_4, is calculated from weak-scintillation theory and corrected for multiple-scatter effects. The theoretical basis for the correction stems from Rician signal statistics, but it is believed to provide a practically useful transition between conditions of weak and saturated ($S_4 = 1$) scintillation. Morphologically, the model has been updated only in the auroral zone. Phase-scintillation data obtained from the DNA Wideband satellite at Poker Flat AK over a period of almost three years following solar minimum have been used to calibrate the high-latitude model. The update includes the strong geometrical dependence of phase scintillation stemming from three-dimensionally anisotropic irregularities and a very considerable increase in mean scintillation level with increasing solar activity. Limitations of the model stemming from uncertainties and unknowns about high-latitude scintillation are stated in the paper.

I. INTRODUCTION

A variety of modern military systems used for communication, navigation, and surveillance depends upon transmission of radio signals through the ionosphere. Thus, description and understanding of ionospheric structure in scale-size regimes that can distort transionospheric radio waves are research topics of interest to several elements of the Department of Defense. Moreover, high-altitude nuclear plasmas become structured under the influence of various instabilities, and the resulting irregularities scatter radio waves to produce phase and intensity scintillations.

As a result of the common physics underlying the scintillation phenomenon arising in both the naturally disturbed and nuclear-perturbed ionosphere, the Defense Nuclear Agency (DNA) fielded an experiment in which carefully designed radio signals were transmitted from the DNA Wideband satellite through the naturally structured ionosphere (Fremouw *et al*, 1974; Rino *et al*, 1977; Fremouw *et al*, 1978). The portion of the Wideband data base collected at Poker Flat AK (64.8° invariant; 65.1° N, 147.5° W) has now been summarized in an applications-oriented computer model, to be used for engineering evaluation of effects to be expected from auroral-zone scintillation. This paper briefly describes the model code, WBMOD, which is based on earlier work by Fremouw and Bates (1971), Fremouw and Rino (1973), and Fremouw *et al* (1977a).

II. OVERVIEW OF WBMOD

A. Structure of the Code

Figure 1 is a flow diagram of Program WBMOD. Upon initiation of the program, the user is asked for information regarding his computational scenario. The requested information includes parameters of the user's system, such as operating frequency and the longest time over which the system's mission requires phase stability. It also includes other aspects of the intended operation, such as transmitter and receiver location and time of day, plus characterization of the general state of solar-terrestrial disturbance by means of sunspot number and planetary magnetic activity index, K_p. Finally, the user specifies one of his input quantities as the independent variable (e.g., transmitter location or time of day). Indicators of scintillation strength (i.e., scintillation indices), are calculated as functions of the selected independent variable.

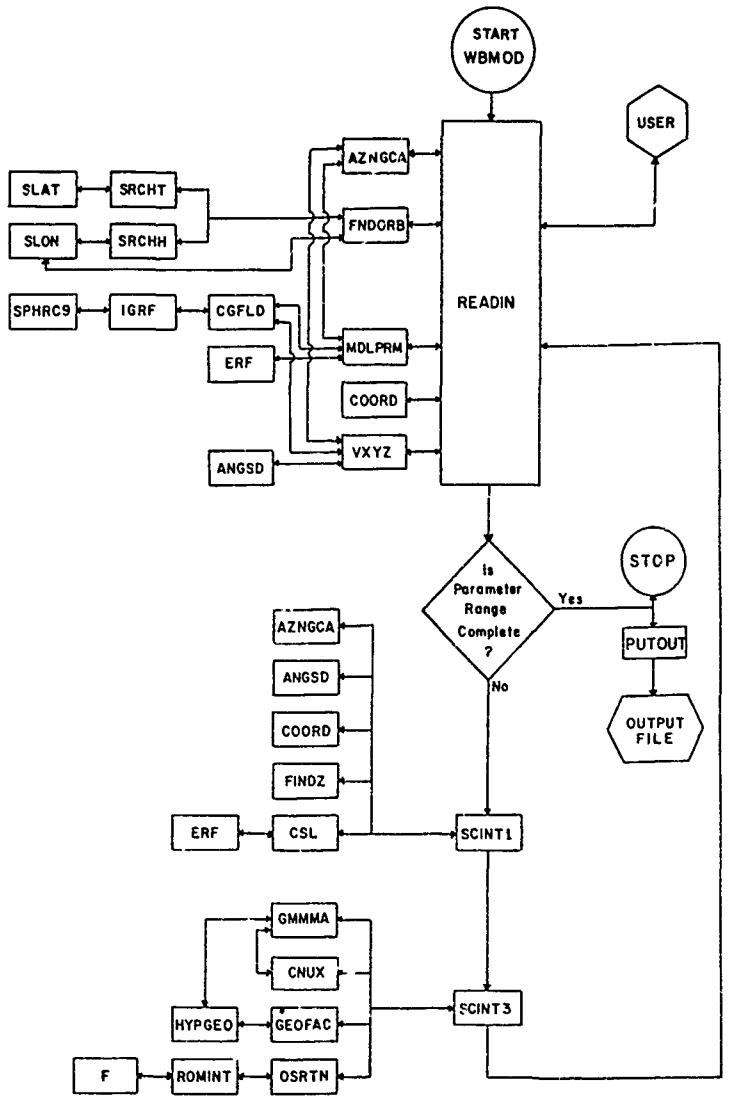


Figure 1. Flow diagram for Program WBMOD.

Control of WBMOD computations is quickly relinquished by the driver program to three major subroutines, READIN, SCINT1, and SCINT3. In addition to calling for program inputs, READIN controls many of the computations that are peripheral to calculation of the scintillation indices. Since scintillation severity is highly geometry-dependent, much of the code is involved in geometrical computations, and some of these are controlled by READIN.

Among the parameters that the user may choose to vary are the receiver or transmitter latitude/longitude coordinates (RCRD and TCRD respectively). This may be done either in an incremental but static manner or in an orbital mode (ORBT), in which the scanning motion of the line of sight is taken into account. (In all modes but ORBT, scintillation is taken to arise solely from drift of ionospheric irregularities across a stationary line of sight.) Whether the variable parameter is a terminal location or some other independent variable (e.g., K_p or sunspot number), its incrementing is controlled by READIN. The first implementation of the irregularity model is made by means of calls to MDLPRM, which calculates all parameters describing the irregularities except their

strength. Simple models for irregularity drift velocity and height of the scattering layer, upon which the line-of-sight scan velocity depends, are invoked at this time to complete certain geometry calculations.

If the user has not specified an irregularity drift velocity (i.e., if he has defaulted by entering "model" when asked), the code will employ the rudimentary drift model contained in MDLPRM. If the user has selected the ORBT mode, the drift velocity is added to the scan velocity calculated in Subroutine VXYZ and its subordinates. The scan velocity dominates over the drift velocity in typical low-orbiting scenarios, in which case the drift is not very important. While the geometry is calculated in geodetic coordinates, all of its aspects that control radiowave scatter and the development of scintillation must account for orientation of the geomagnetic field. Thus, MDLPRM and VXYZ rely upon Subroutine CGFLD, which sets up calls to the International Geomagnetic Reference Field (Trombka and Cain, 1974) contained in Subroutine IGRF.

As READIN increments the calculations, it also checks for completion of the number of increments specified by the user, whereupon it calls on Subroutine PUTOUT to list the computation results in an output file. Until completion is accomplished, READIN passes control back to the driver during each increment cycle, and the driver calls Subroutines SCINT1 and SCINT3 in sequence to carry out the main scintillation computations.

B. The Scintillation Calculations

Subroutine SCINT1 plays a preparatory roll for calculating the scintillation parameters. Its two main functions are (1) to establish the scattering geometry at the ionospheric penetration point of the line of sight and (2) to provide the strength of scintillation-producing irregularities at that point, calculated from the model as Function CSL. For its first task, SCINT1 relies upon several subroutines used earlier. For its second task, SCINT1 makes the code's only request for information on irregularity strength by invoking Function CSL. The other key information about the irregularities, which describes their three-dimensional configuration and height, has already been calculated by Subroutine MDLPRM and passed directly to SCINT3 in Common Block OMP.

The direct scintillation calculations are made in Subroutine SCINT3, which makes use of the phase-screen scattering theory of Rino (1979). The central quantity calculated is T, which in Rino's original infinite outer-scale formulation numerically equals the power spectral density of phase at a fluctuation frequency of 1 Hz. It is given by

$$T = \lambda^2 r_e^2 \frac{\sqrt{\pi}\,\Gamma(\nu)}{(2\pi)^{2\nu+1}\Gamma(\nu+\frac{1}{2})} C_s L(\sec\theta) G\, V_e^{2\nu-1} \quad , \tag{1}$$

where λ = radio wavelength,
r_e = classical electron radius,
θ = incidence angle of the propagation vector on the (horizontal) scattering layer,
and L = layer thickness.

The gamma functions arise from normalizing the three-dimensional ionospheric spectrum to the electron-density variance, $\langle(\Delta N)^2\rangle$, such that the structure function is

$$C_s = 8\pi^{3/2} \frac{\Gamma(\nu+\frac{1}{2})}{\Gamma(\nu-1)} \frac{\langle(\Delta N)^2\rangle}{\alpha^{(2\nu-2)}} \quad . \tag{2}$$

The spectrum itself is taken to be of the form $C_s \kappa^{-\eta}$ in the spectral regime responsible for scintillation, but to be held finite by an outer scale, α. The gamma-function arguments depend only upon the sharpness of electron-density gradients, expressed as a spectral parameter given by

$$\nu = \frac{\eta-1}{2} \quad . \tag{3}$$

The corresponding one-dimensional (*in-situ*) spectral index is $2\nu-1$ (Cronyn, 1970), and the two-dimensional (phase) spectral index is

$$p = 2\nu \quad . \tag{4}$$

The height-integrated spectral strength of the irregularities, $C_s L$, is the quantity obtained from Function CSL, and our model for it will be described in detail in Section III C.[5] The two remaining quantities in Eq. (1), G and V_e, describe respectively the static and dynamic aspects of geometrical control over phase scintillation. They are calculated in Subroutine GEOFAC, which is called by SCINT3, as follows:

$$G = \frac{ab \sec\theta}{\sqrt{AC-B^2/4}} \tag{5}$$

$$V_e = \frac{\left(CV_{sx}^2 - BV_{sx}V_{sy} + AV_{sy}\right)^{\frac{1}{2}}}{\sqrt{AC-B^2/4}} \tag{6}$$

where a = field-aligned axial ratio (ratio of irregularity size along the geomagnetic field to that normal to the field in a reference direction),

b = second axial ratio for describing sheetlike irregularities (ratio of irregularity size in direction normal to both the geomagnetic field and the reference direction to that in the reference direction),

and V_s = a foreshortened horizontal projection of the line-of-sight scan velocity calculated in Subroutine VXYZ and defined in Eq. (14) of Rino (1979).

The geometrical factors A, B, and C, which are defined in Eq. (41) of Rino and Fremouw (1977), depend upon the incidence angle, θ, and magnetic heading, ϕ, of the propagation vector in addition to a and b. They depend also on the geomagnetic dip angle, ψ, and on a final irregularity parameter, δ, which is the angle that sets the reference direction for defining b.

The physical fact accounted for by the static geometrical enhancement factor, G, is that the phase perturbation imposed on a radio wave propagating along an extended dimension of irregularities builds up quasi-coherently, as compared with that for propagation along a short dimension. That described by the effective velocity, V_e, is that a low-pass spatial spectrum results in stronger spectral density at a given temporal frequency (1 Hz) not only for greater scan velocities, but also for scans across short irregularity dimensions as compared with long ones.

The fundamental outputs from WBMOD are T and p, which respectively are measures of the strength and spectral character of phase scintillation. The power-law spectral index, p, of phase is obtained from Eq. (4), which ignores the effect of diffraction on the shape of the phase spectrum. Diffractive alteration of p is believed to occur, but to be quite subtle (Livingston *et al*, 1981). The code is structured so that future research results about spectral index could be incorporated in Subroutine MDLPRM, but WBMOD currently employs a constant value of 1.25 for ν and outputs the corresponding value of p (2.50).

Unlike p, the strength, T, of phase scintillation is highly variable. The large majority of WBMOD is given over to calculating T and two commonly used indices of scintillation activity based on it, one for phase and one for intensity. The scintillation index for phase is simply its standard deviation, σ_ϕ, which may be calculated by integrating the phase-scintillation temporal spectrum, $\Phi_\phi(f)$, as follows:

$$\sigma_\phi^2 = \int_{f_c}^{\infty} \Phi_\phi(f)\,df = \int_{f_c}^{\infty} \frac{T\,df}{(f_o^2+f^2)^{p/2}} \quad , \tag{7}$$

where

$$f_o = V_e/2\pi\alpha \quad . \tag{8}$$

[The outer scale, α, is measured in rad/m in the field-normal reference direction used in defining a and b, at the $2^{(\nu-\frac{1}{2})}$-point on the in-situ power spectrum.]

In Eq. (7), f_c is the lowest phase-fluctuation frequency to which the system is sensitive. For instance, in the Wideband satellite experiment with normal processing, f_c was 0.1 Hz (Fremouw *et al*, 1978) as set by phase detrending. In a coherently integrating radar, it would be the reciprocal of the time over which phase coherence is required. For systems not sensitive to phase instability in the propagation medium, f_c is effectively infinite, and the effective σ_ϕ is zero.

Equation (7) may be evaluated analytically for three ranges of the ratio f_c/f_o (Fremouw and Lansinger, 1981), but an analytical evaluation has not been found for the range f_c only slightly greater than f_o. The ionospheric outer scale, α, is sufficiently large that $f_c >> f_o$ over the range of effective velocity, V_e, encountered in the Wideband experiment, and we have employed the corresponding analytical evaluation of Eq. (7) for our modeling. Moreover, the magnitude and variational behavior of α are not known. While the outer scale appears to be quite large compared with the spatial windows of a number of ionospheric experiments, there is no assurance that the infinite outer-scale limit would apply for all systems in all operating scenarios. Accordingly, SCINT3 contains an efficient means (Subroutine OSRTN and Function F) for numerically evaluating Eq. (7), so that the code is not restricted inherently to application in the infinite outer-scale limit. At present, α is set at a very large constant value (10^6 m), so the limiting analytical evaluation of Eq. (7) is employed in any likely application. An option is provided for the user to override this default value should he want to investigate the effect of varying the outer scale. The main reason for coding Eq. (7), however, is to prepare for ready inclusion of any new results on the ionospheric outer scale that may be yielded by research programs.

The scintillation index for intensity is the ratio, S_4, of the standard deviation of received signal power to the mean received power (Briggs and Parkin, 1963). Unlike σ_ϕ, its relation to T is set not by a system or

ionospheric parameter, but by the diffraction process that gives rise to intensity scintillation. For weak to moderate levels of intensity scintillation, S_4^2 is very well approximated (Rino, 1979 and Fremouw, 1980) by

$$S_{4w}^2 = C(\nu)\frac{T}{V_e^{(2\nu-1)}}\ \frac{F}{G}\ z^{\nu-\frac{1}{2}} \quad , \tag{9}$$

where $C(\nu)$ is a normalization factor related to that in Eq. (1). The Fresnel filter factor, $F(a, b, \delta, \nu)$, describes the geometrical enhancement of intensity scintillation. It also accounts for diffraction, together with the Fresnel-zone size,

$$z = \frac{\lambda z \sec\theta}{4\pi} \tag{10}$$

in which z is the effective "reduced height" (including correction for wave-front curvature and curved-earth geometry) of the irregularities.

While Eq. (9) is a weak-scintillation formula, it may be generalized for practical purposes, to include the well-known saturation of S_4 at unity by writing

$$S_4^2 = 1-\exp(-S_{4w}^2) \quad , \tag{11}$$

which is exact for scintillating signals that obey Rice statistics (Fremouw and Rino, 1976). Use of Eq. (11) ignores some effects of geometrical-optics focusing, which can drive S_4 modestly above unity and which subtly alter the signal statistics accompanying scintillation (Fremouw, Livingston, and Miller, 1980). Comparison of the behavior of S_4 and σ_ϕ, as measured in the Wideband experiment, however, shows that Eq. (11) is quite adequate to represent the behavior of S_4 for presently identified applications of WBMOD, and it has been coded into SCINT3.

III. THE MODEL

A. Overview

In order to calculate T, p, σ_ϕ, and S_4, one must have values for eight parameters describing ionospheric irregularities. They are the height, h, and vector drift velocity, V_d, of the irregularities; an outer scale, α; four "shape" parameters describing the irregularities' three-dimensional configuration and spatial "sharpness", a, b, δ, and ν; and the height-integrated spectral strength, C_sL. Program WBMOD contains models for the foregoing eight parameters, but the degree of detail is very much less for some than for others.

As alluded to in Section II B, the "model" for outer scale is a single, effectively infinite (10^6 m), constant value. That for drift velocity is almost equally rudimentary, consisting of a geomagnetically eastward motion at a speed, in m/sec, of

$$V_{dy} = 50-15\left(1+\text{erf}\ \frac{\lambda_m-20^\circ}{3^\circ}\right) + 40(1+K_p)\left(1+\text{erf}\ \frac{\lambda_m-\lambda_b}{3^\circ}\right) \tag{12}$$

where erf stands for error function and λ_m is the geomagnetic invariant latitude. Equation (12) describes a drift of 50 m/sec at the geomagnetic equator, dropping to 20 m/sec at middle latitudes, and increasing with geomagnetic disturbance at latitudes above that, λ_b, of the high-latitude scintillation boundary. This description is in need of review, especially at high latitudes, and users with phase-sensitive applications involving geostationary satellites are advised to make use of the option provided for external specification of values for V_d. (Drift velocity is of no consequence for WBMOD application to systems susceptible only to intensity scintillation, and of little importance in even phase-sensitive low-orbit applications.) A rudimentary description of the effective (centroid) height, h, of scintillation-producing irregularities in the F layer also is included in WBMOD, as follows:

$$h = 500-75\left(1+\text{erf}\ \frac{\lambda_m-20^\circ}{3^\circ}\right)\ \text{km}. \tag{13}$$

That is, the equivalent phase-changing screen is taken to be at about the middle of the F layer, 350 km, except near the geomagnetic equator where nighttime scintillation seems to arise in a layer extended to considerable height.

Prior to the Wideband experiment, *in-situ* measurements of ionospheric irregularities were generally reported as indicating a one-dimensional power-spectral index of about 2, which would translate to a phase spectral index, p, of 3 and a value of 1.5 for ν. Our approach to scintillation modeling was predicated on being able to treat ν as a constant in contrast to the expected large range of variation in irregularity strength, C_sL. In view of Eq. (4), a constant ν translates to a constant p, which was a Wideband post-processing observable. Investigation of the occurrence of p values observed at VHF from Poker Flat did show a rather narrow distribution (Fremouw and Lansinger, 1979), but the peak was between 2.0 and 2.5 rather than near 3, as expected. Employing a value on the high end of the observed range tends to offset the effect of ignoring

diffraction wrought by employing Eq. (4), and we have coded the value 1.25 for ν (i.e., p = 2.50).

With simple models established for outer scale, drift velocity, height, and gradient sharpness (spectral index) of the irregularities, we are ready to consider the more variable and, therefore, important parameters: a, b, and δ for describing the three-dimensional configuration of the irregularities and $C_s L$ for describing their height-integrated strength. The three remaining shape parameters have been established for auroral-zone irregularities by analyzing the geometrical behavior of scintillation as observed by means of Wideband at Poker Flat. The procedure and results are described in Section III B. By far the most effort was put into modeling irregularity strength, by means of iterative comparisons with the data, as described in Section III C.

B. Three-Dimensional Configuration

Irregularities known *a priori* to have some degree of magnetic-field alignment are extended in a direction characterized by the dip angle, ψ. In a coordinate system thus tied to the geomagnetic field, the three-dimensional irregularity configuration is characterized by a, b, and δ. For isotropic irregularities, we would have a = b = 1, and geometrical control of scintillation would reduce to a path-length effect measured as sec ϑ. "Rodlike" irregularities displaying axial symmetry about the magnetic field (a >1,b=1) would produce enhanced scintillation only when the line of sight is nearly parallel to the field. Three-dimensionally anisotropic (a > 1, b > 1) irregularities produce enhanced scintillation near a line in the sky dictated by δ.

Pre-Wideband scintillation data clearly established field-alignment of the irregularities, with particularly large values of the along-field axial ratio, a, reported in the equatorial region (Koster, Katsriku, and Tete, 1966). It has now been established that the dominant irregularities immediately poleward of the nighttime high-latitude scintillation boundary are extended also in the geomagnetic east-west direction, as though layered like onion skins along L shells (Singleton, 1973; Martin and Aarons, 1977; Fremouw *et al*, 1977b; Rino, Livingston, and Matthews, 1978). More recently, it has been established that these so-called sheetlike irregularities are confined to the night side of the auroral irregularity zone (Fremouw, Lansinger, and Miller, 1980), with very few exceptions.

Alignment along L shells is described by $\delta = 0$, and we have coded such a constant value into WBMOD. (The value of δ is immaterial for axially symmetric irregularities, and the only three-dimensionally anisotropic irregularities presently known to exist are L-shell aligned.) We exploited the aspect-sensitivity of phase scintillation (Fremouw and Lansinger, 1981) to establish initial values of a and b, prior to iterative adjustments to be described in Section III C. Recent refinement of our technique suggests that there is some diurnal change in the value of a at the latitude of Poker Flat. This effect could result from a lesser degree of field alignment equatorward of the high-latitude scintillation boundary than poleward of it, however, since the boundary usually is located well poleward of Poker Flat during the day. Pending a possible future separation of latitudinal and diurnal variations of a, we have coded the following simple model:

$$a = 30 - a_h \left(1 + \mathrm{erf} \, \frac{\lambda_m - 20^\circ}{3^\circ} \right) \tag{14}$$

Equation (14) describes very elongated irregularities near the geomagnetic equator and a field-aligned axial ratio of $30-2a_h$ elsewhere. Our initial value of a_h was 10.75, which then was slightly modified by iterative tests to be described in Section III C.

It is commonly supposed that mid-latitude irregularities are axially symmetric (b = 1), but Wideband has shown that nighttime irregularities in the auroral precipitation zone are not (b > 1). The initial value of b established by our procedure is a kind of average nighttime value at the magnetic latitude of Poker Flat (65°), which often is near the scintillation boundary between the mid-latitude and auroral ionospheres at night. Unlike the situation with a, for which refinement of technique was necessary to uncover a possible diurnal variation, we found clear evidence for such a variation in b. Taking all daytime data from non-overhead pass corridors, we found no indication of a prominent geometrical enhancement. Again noting that such an effect might be caused by a latitudinal rather than a diurnal variation, we separated passes during which the dayside scintillation boundary was (1) equatorward and (2) poleward of Poker Flat, finding 100 of the former, 88 of which were from non-overhead corridors. In only five such passes was there evidence of geometrical enhancement outside the overhead corridor. We concluded, therefore, that irregularities on the day side of the auroral ionosphere are predominantly rodlike, and we coded the following model for b:

$$b = 1 + b_h \left[1 + \cos \frac{\pi (T_m - 2)}{12} \right] \left[1 + \mathrm{erf} \, \frac{\lambda_m - \lambda_b}{\lambda_n} \right] \tag{15}$$

where T_m = geomagnetic time, in hours.

Equation (15) describes rodlike irregularities over most of the earth, with the exception of the region poleward of the nightside scintillation boundary latitude, λ_b. Above the boundary, the value of b changes

smoothly from unity in the daytime to $4b_h + 1$ at night, peaking at that value 2 hours after geomagnetic midnight. The width of the latitudinal transition is described by λ_h, which will be discussed in Section III C. The initial value of b_h was set at 1 and then modified by iteration, as will be discussed in the next section.

C. Height-Integrated Irregularity Strength

The parameters discussed in the previous section are modeled in Subroutine MDLPRM of Program WBMOD. They leave only one of the eight irregularity parameters to be described. It is the most variable and, probably, the most important of the eight: the height-integrated strength, C_sL. The irregularity strength is modeled by means of Function CSL as follows:

$$\sqrt{C_sL} = E(\lambda_m, \lambda_g, T, D, \bar{R}) + M(\lambda_m, T) + H(\lambda_m, T_m, K_p, \bar{R}) \tag{16}$$

where λ_m = geomagnetic invariant latitude,

λ_g = geographic latitude,

T = local meridian time,

D = day of the year,

$\bar{R}$ = smoothed Zurich sunspot number,

T_m = geomagnetic time,

and K_p = planetary geomagnetic activity index. The three terms in Eq. (16) respectively describe the strength of equatorial, mid-latitude, and high-latitude irregularities. The first two have not been tested extensively against Wideband data, and we shall consider only H.

The high-latitude term is based on the observation that there often is a more-or-less abrupt boundary (Aarons, Mullen, and Whitney, 1969) between the mid-latitude region of relatively smooth ionosphere and the high-latitude scintillation region. It is located, typically, equatorward of discrete-arc auroras in the general vicinity of the diffuse auroral boundary. The underlying form of H stems from the supposition that the instantaneous boundary latitude is normally distributed about a mean value, λ_b, for a given set of T_m, K_p, and $\bar{R}$. This supposition, together with other considerations to be discussed shortly, yields the following form for H:

$$H = C_h\ (1+C_r\bar{R}) \left[1+\mathrm{erf}\left(\frac{\lambda_m-\lambda_b}{\lambda_h}\right)\right] \tag{17}$$

where the C's are constants to be established by iterative testing against scintillation data, and where the error function arises from integration over the normal distribution of instantaneous boundary location, which distribution has standard deviation λ_h (Fremouw and Bates, 1971).

The multiplicative dependence of H on $\bar{R}$ stems from (1) our observation at Poker Flat that scintillation increased with advancing phase of the solar cycle during the Wideband experiment (Fremouw and Lansinger, 1980) and (2) a consistent observation in the northern polar cap by Aarons (private communication). Moreover, we have found that scintillation activity is higher for a given K_p in years of high sunspot number than in years of lower $\bar{R}$ (Fremouw, Lansinger, and Miller, 1980), so the dependence of H on K_p and $\bar{R}$ may be modeled in separable fashion.

To establish the sunspot-number dependence of H, which is proportional to σ_ϕ through Eqs. (1), (7), and (16), we made a scatter plot of monthly averages of σ_ϕ against $\bar{R}$, as illustrated in Figure 2. Using nighttime data only, in order to minimize contamination from sub-boundary scintillation, we found the linear least-square fit shown in the figure. From the ratio of intercept to slope, C_r was evaluated as approximately 0.05. The linear fit describes the general upward trend of scintillation severity with increasing sunspot number rather well.

At the same time, there is an enticing quasi-cyclic departure of the observed behavior from the trend line. Since Basu (1975) found a marked seasonal dependence in scintillation activity in the Greenland sector, we investigated the possibility that the oscillatory behavior in Figure 2 may stem from a seasonal variation at Poker Flat. We found no statistically significant seasonal pattern, a point to which we shall return in Section IV. We should like to explore the possibility that the departure from the linear trend is related to reversals in the interplanetary magnetic field, but such an investigation is beyond the scope of the present endeavor.

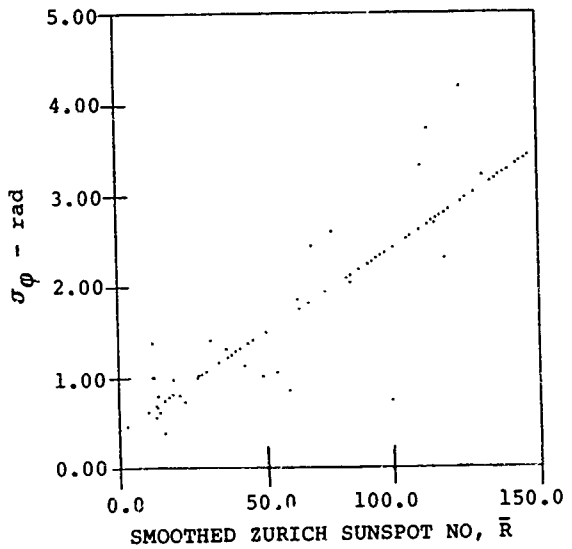


Figure 2. Sunspot-number dependence of nighttime, VHF phase scintillation index a. Poker Flat. Ordinate values represent monthly means.

It is well known that the auroral oval and a variety of boundaries essentially concentric with it migrate equatorward with increasing geomagnetic activity and poleward with decreasing disturbance. The scintillation boundary participates in this migration and lies at a higher latitude on the day side of the earth than on the night side, as do the other boundaries. These facts about the scintillation boundary, together with a reported near-independence of polar-cap scintillation from K_p control (Aarons, private communication), are described by Eq. (17) in conjunction with the following expression for the invariant latitude of the scintillation boundary:

$$\lambda_b = \lambda_1 - C_k K_p - C_{bt} \cos \frac{\pi (T_m - 2)}{12} \tag{18}$$

where λ_1 and the C's are constants to be evaluated.

The remaining characteristic of the "average" or "climatological" scintillation boundary to be described is its width, λ_1 (i.e., the latitudinal extent of the region over which, on average, scintillation severity changes from its low value in the sub-boundary "trough" to its high-latitude value). We have found the following description of average boundary width to be useful:

$$\lambda_h = C_{hb} \lambda_b \tag{19}$$

where C_{hb} was found by iterative testing to be 0.15.

Equation (18) describes a circular scintillation boundary centered $\lambda_1 - C_k K_p$ degrees toward the dayside of the earth from the geomagnetic pole. The line of symmetry passing through the pole, however, is not the geomagnetic-midnight meridian, but rather is shifted two hours after midnight. This shift is based on a finding of Basu and Basu (1981) from *in-situ* plasma-probe data. We found that imposing such a shift in Eqs. (15) and (18) improved our ability to obtain simultaneously satisfactory fits to some of the iterative-test data sets to be discussed next.

Inspection of Eqs. (14), (15), (17), (18), and (19) reveals the following eight constants to be established for quantitative description of the three-dimensional configuration and height-integrated strength of high-latitude scintillation-producing irregularities: a_h, b_h, C_h, C_r, λ_1, C_k, C_{bt}, and C_{hb}. As has been discussed, essentially deductive procedures were found for establishing at least starting values for four of the eight: a_h, b_h, C_r, and C_{hb}. Starting values for the remaining four were available from earlier work and from an experience-based intuition about behavior of the high-latitude ionosphere.

With starting values for the eight iterative constants established, the Wideband data population from Poker Flat was divided into 22 subsets, 14 of which were used for iterative model-building, with the remaining eight reserved for final testing. The division was made on the basis of pass geometry and K_p. Since Wideband is in a sun synchronous orbit, pass corridors established for defining geometry could be parameterized by means of pass

time. The first two corridors to be defined were (1) nighttime passes for which the pass-minimum value of the angle between the propagation vector and the geomagnetic field was 10° or less and (2) daytime passes meeting the same criterion. These two corridors turned out to be bounded in universal time (UT) by (1) 1018-1059 and (2) 1944-2022.

Local standard time at the station corresponds to UT - 10 hours. Following definition of the foregoing two overhead corridors, corridors well east (earlier in time) and well west (later in time) of the station were defined for iterative modeling. They are bounded in UT respectively by (night) 0854-0939 and 1154-1239 and by (day) 1739-1824 and 2039-2124. For both day and night, final-test corridors were defined between the overhead and east corridors and between the overhead and west corridors. They are designated as east-intermediate and west-intermediate, and data from them were not used in iterative model building but rather were used as an indication of model fidelity after all iterative constants were frozen.

Once the geometry/time corridors were established, the data in them were divided into the following three K_p ranges: 0 through 2+ (designated low), 3- through 5+ (mid), and 6- and above (high). For the low and mid categories of K_p, sufficient data exist for separation into all geometry/time corridors. There were too few high-K_p passes, however, to make statistically meaningful data subsets by corridor. Thus for high K_p, only two subsets were established (one for daytime and one for nighttime). Finally, then, there were 14 data subsets for iterative modeling (six each daytime and nighttime corridors for each of two K_p ranges plus two high-K_p sets) and eight for final testing (two each daytime and nighttime corridors for each of two K_p ranges).

For each geometry corridor, a representative pass was selected from near the center of the corresponding UT range. The pass time and the beginning and ending latitude and longitude for the representative pass were then input to WBMOD, and the code was run in ORBT mode. External software was employed to display the calculated value of σ_ϕ as a function of invariant latitude of the F-layer (350-km) penetration point on a graphics terminal. The values of σ_ϕ measured during passes in each data subset were then sorted into the same penetration-point latitude bins, and their average values also were displayed on the screen.

For the WBMOD calculations, K_p was set at the middle value (1, 4, or 7) of the range included in the data subset being used for comparison. For all iterative modeling runs, the sunspot number was fixed at 50, which was set after calculating 52 as the mean value encountered in the Wideband experiment (weighted by the number of data points available for each incremental value of $\bar{R}$). Sunspot-number dependence was modeled independently, as discussed in conjunction with Figure 2.

With the independent variables (geometry, time, K_p, and $\bar{R}$) set, WBMOD was run and the result compared graphically with the corresponding σ_ϕ measurements from a given data subset. The model constants were adjusted to provide better (subjectively judged) fits for each representative pass. Initially, such iterations were performed for the low-K_p and mid-K_p data sets for each of the six nighttime model-building corridors. With interim values for the constants set, a similar round of iterations was performed for the daytime cases. Several iterations often were run at a single modeling session, comparisons being made with a number of data subsets. Periodically, hard-copy plots of all 12 passes were made so as to permit simultaneous comparison of data and model results for all. Such comparisons yielded strategy for the next iteration session. The objective was to find a set of constants which simultaneously gave satisfactory fits for all 12 subsets plus the two high-K_p sets.

The final values iteratively established for the high-latitude model constants are as follows:

$a_h = 11$	$C_h = 4.3 \times 10^{11}$	$\lambda_l = 71°$	$C_{bt} = 5.5$
$b_h = 0.75$	$C_r = 0.0496$	$C_k = 1.5$	$C_{hb} = 0.13$

The resulting model outputs are compared with their data counterparts in Figures 3, 4, and 5.

Figure 3 contains the six low-K_p (bottom) and mid-K_p (top) nighttime representative passes (solid) and data subsets (broken). From left to right, the passes progress in time (and the orbital plane from east to west). The geometrical enhancements are prominent in both the calculations and the data sets, being located close to the latitude of the station at the point of minimum off-shell angle (subject to one-degree latitude resolution). The general increase in activity with increasing K_p is evident, and the differences in enhancement between overhead and off-meridian passes and between pre-midnight and post-midnight passes are reasonably well reproduced by the model.

In some instances, the model underestimates σ_ϕ to the north of the station. The calculated value there is depressed by a decrease in V_e as the line of sight scans along extended axes of either rodlike or sheetlike irregularities. It is tempting to introduce a further latitudinal change in axial ratios (decreasing at higher latitudes) in an attempt to improve the fits. There is insufficient information for reliably doing so, however, from the phase data only. Improvements may well be possible on the basis of S_4/σ_ϕ, but such an effort was beyond the scope of the present work.

Results for daytime passes in the overhead (left) and west (right) corridors are displayed in Figure 4. The

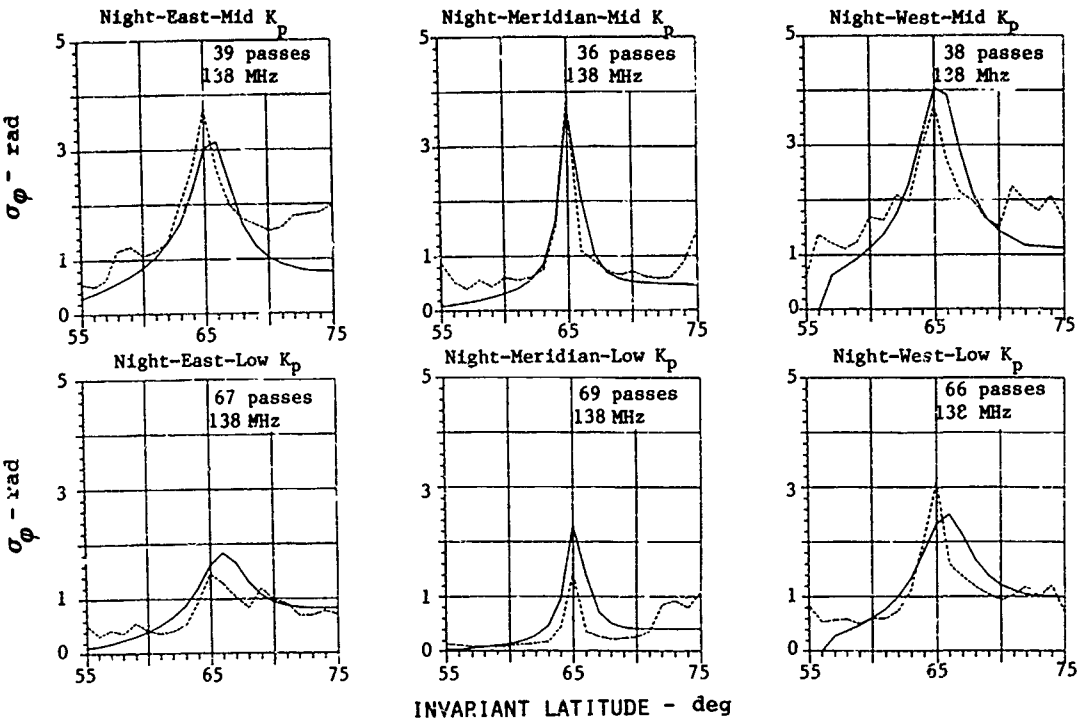


Figure 3. Comparison of nighttime, VHF phase scintillation index calculated (solid) by means of Program WBMOD and observed (broken) at Poker Flat, Alaska in the Wideband Satellite Experiment. Left: pre-midnight passes to the east of the station. Center: near-midnight passes essentially along the geomagnetic meridian. Right: post-midnight passes to the west of the station. Bottom: $0 \leq K_p \leq 2+$. Top: $3- \leq K_p \leq 5+$. No. of passes in each data set is indicated in the upper righthand corner of each grid. Calculated values are for a single representative pass.

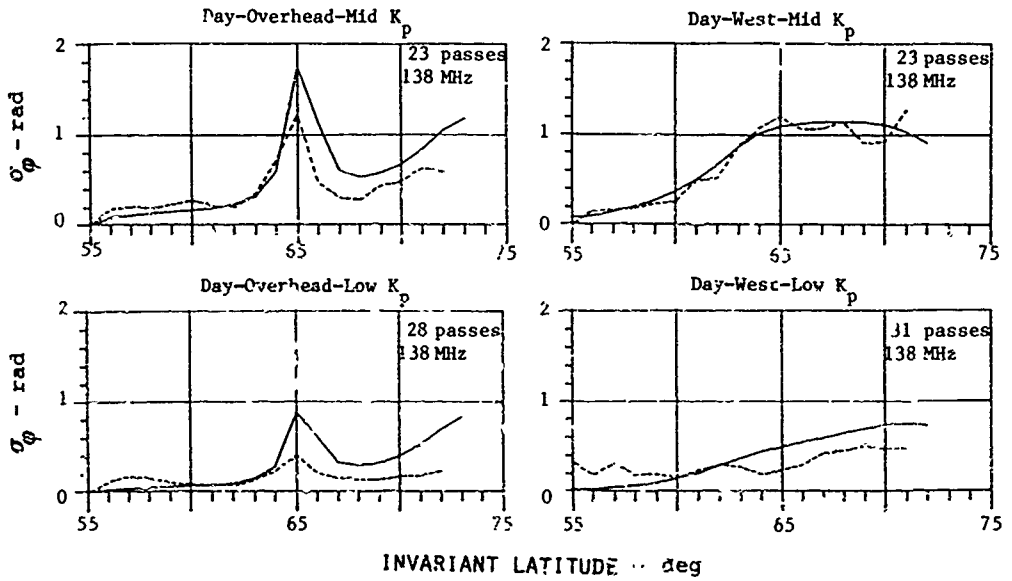


Figure 4. Comparison of WBMOD (solid) results and observed values (broken) of daytime VHF phase scintillation index. Left: overhead (morning) passes. Right: west (late morning) passes. Bottom: $0 \leq K_p \leq 2+$. Top: $3- \leq K_p \leq 5+$. No. of passes is indicated in the upper right-hand corner of each grid.

east corridor produced rather short passes and did not contribute appreciably to the iterative modeling. Again, the low-K_p sets are at the bottom and the mid-K_p sets are at the top. The total lack of a geometrical enhancement in the west corridor attests to the diurnal variation in b, describing L-shell alignment at night and axial symmetry (in view of the enhancement in the overhead corridor) in the daytime. The general level of scintillation is reasonably well reproduced, as it was at night.

The code overestimates daytime scintillation somewhat to the north of the station in several instances. The gratifying degree of fit in the west corridor under mid-K_p conditions, however, suggests reasonable fidelity in the latitudinal description of irregularity strength. This case is useful for that purpose because the transition region is accessible from Poker Flat, and because complications regarding axial ratios do not arise in this geometry.

Finally, the high-K_p results for nighttime and daytime are shown in Figure 5. For very disturbed conditions ($K_p \geq 6-$), there were too few Wideband passes to maintain separation of the data population into subgroups by geometry/time corridors. Geometrical and time considerations were taken into account in the model calculations in the following way. First, the number of high-K_p passes in each corridor was established, along with the average sunspot number and K_p value for that small data subset. The code was then run for the previously chosen representative pass geometry and time for that corridor, using the subset-average values of R and K_p. The resulting σ_ϕ values in each latitude bin then were averaged together after weighting by the number of data passes in each corridor. What appears in Figure 5 as solid curves are these average calculation results for nighttime (left) and daytime (right) independently, together with (broken curves) the observed averages from the corresponding data sets.

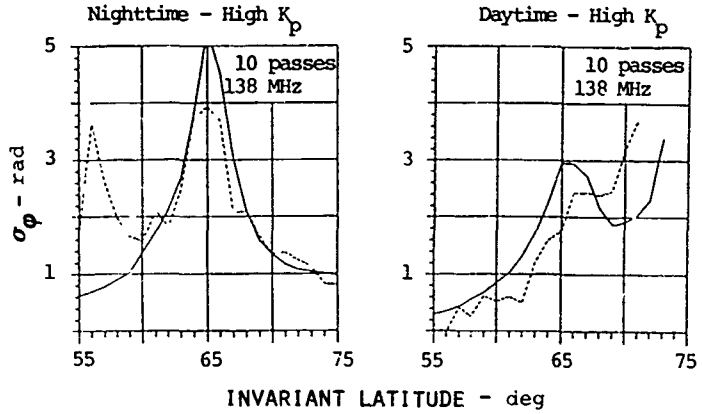


Figure 5. Comparison of model results (solid) and observed values (broken) of VHF phase scintillation index for geomagnetically very disturbed ($K_p \geq 6-$) nighttime (left) and daytime (right) conditions.

The results illustrated in Figure 5 were used in the few final iterations of the model constants, but far less so than the data contained in Figures 3 and 4. In particular, the rather satisfactory fit to the nighttime data in Figure 5 poleward of Poker Flat was obtained with very little high-K_p-based iteration. The overestimation of nighttime σ_ϕ in the geometric-enhancement region probably is overstated in Figure 5, as a result of suspected saturation of data-processing procedures under extremely disturbed conditions. The calculated peak may be a better representation than the observed-data curve would indicate.

The most striking and potentially significant deficiency in the nighttime model lies well equatorward of Poker Flat. In several instances, unexpectedly strong phase and intensity scintillations were observed under magnetically disturbed conditions at an ionospheric penetration latitude of about 56° invariant. Quite likely representing a signature of some form of convective instability at the plasmapause, this scintillation feature is totally unaccounted for in the present model.

The general level and latitudinal distribution of daytime scintillation under high-K_p conditions seems rather well described by the model. Figure 5 is consistent, however, with a possibility that there may be a diurnal variation in the field-aligned axial ratio, a, which is not described by the present model. As discussed in Section III B, systematic analysis has disclosed a smaller average value of a at the latitude of Poker Flat in the daytime than at night, but insufficient data were available to sort out reliably whether the effect stems from a diurnal variation or from different values of a poleward and equatorward of the scintillation boundary. Figure 5 is consistent with the former view. Unlike the results shown in Figures 3 and 4, those in Figure 5 include passes from intermediate corridors. They contribute to the geometrical enhancement indicated in the daytime model result shown in Figure 5, and they are sensitive to the modeled value of a. Data from a higher latitude station are needed to do a fully satisfactory job of modeling dayside auroral scintillation.

IV. CONCLUSION

Program WBMOD combines the most useful scattering theory available for calculating radiowave scintillation with the best available description of the electron-density irregularities responsible for ionospherically produced scintillation. It permits a user to specify his operating scenario and to obtain four quantities that characterize complex-signal scintillation for that scenario. In addition to the spectral index, p, for power-law phase scintillation, the code returns the spectral strength parameter, T, and the standard deviation, σ_ϕ, of phase. It returns these quantities and the intensity scintillation index, S_4, as functions of a changing independent variable chosen by the user.

The theory employed in WBMOD is based on the equivalent phase-screen representation of Booker, Ratcliffe, and Shinn (1950), formulated to account for three-dimensionally anisotropic irregularities (Singleton, 1970) described by a power-law spatial spectrum. The formulation employed was developed by Rino (1979) in the infinite outer-scale limit, but a means for dealing with the effect of a finite outer scale on phase scintillation has been incorporated in WBMOD. Similarly, a means has been provided for accomodating multiple-scatter effects on intensity scintillation that should suffice for practical applications.

The descriptive irregularity model is based on numerous observations (Fremouw and Bates, 1971; Fremouw and Rino, 1978), but most particularly on observations of phase scintillation performed in the DNA Wideband Satellite Experiment (Fremouw *et al*, 1978). The most significant caveat about use of WBMOD, however, is that it has been calibrated quantitatively against Wideband data from only a single station in the northern auroral zone (Poker Flat AK). As described in Section III, the descriptive model was developed by iterative comparison with most of the Wideband data population from Poker Flat, with a portion of the population reserved for final comparative tests.

In Figure 6, we present model comparisons with two data sets not used in iterative development. The data shown are from the west-intermediate nighttime (left) and daytime (right) corridors, collected under moderately disturbed geomagnetic conditions ($3- \leq K_p \leq 5+$). The model was run for the geometry and time of a representative pass in each corridor, using a K_p of 4 and a sunspot number of 50 (approximately the weighted-mean value for the Wideband experiment, as described in Section III C).

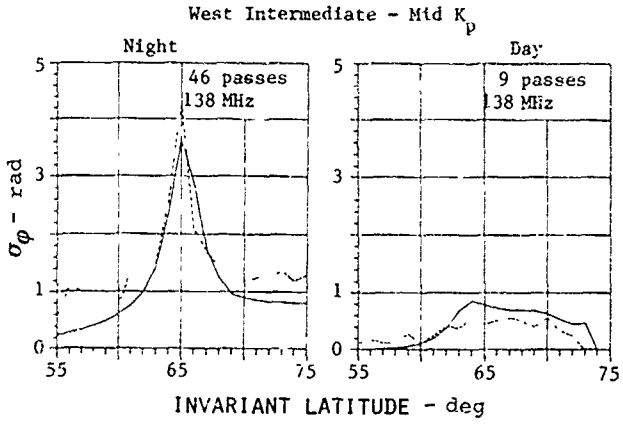


Figure 6. Comparison of model results (solid) with observed values (broken) of VHF phase scintillation index from two dissimilar data sets not used in iterative development of the model

The figure corroborates that the model describes the general level and the main features of auroral-zone phase scintillation with rather satisfactory fidelity (within a factor of two, say, for multi-pass data sets). We stress that the irregularity model contained in WBMOD is climatological in nature. Thus, the scintillation indices returned should be treated as expectation values and not as predictions for a specific observation.

As in Figures 3 and 4, there is some tendency in Figure 6 for the model to underestimate nighttime scintillation severity, especially at the lowest subauroral latitudes, and to overestimate daytime auroral-zone scintillation. Investigation of scintillation conditions at the plasmapause and in the subauroral ionospheric trough should remedy the former deficiency, which could be of operational significance under very disturbed conditions (Figure 5). The latter defect may be more difficult to deal with, but the shape of the daytime calculated curve in Figure 6 is consistent with the suggestion based on its counterpart in Figure 5 that the field-aligned axial ratio in the dayside auroral ionosphere is smaller than that included in the present model.

In closing, we re-emphasize that users of the code should judiciously select options according to their specific needs and according to the various model limitations described in this paper. The model ought to be most reliable for describing auroral-zone scintillation, especially in the Alaskan sector. It has been far less reliably compared with observations of equatorial scintillation. Moreover, it is much more reliable for phase-scintillation calculations in low-orbiting applications than in geostationary applications. Users desiring phase

information in the latter situation should consider providing their own estimates of ionospheric drift velocity, although the code will provide default values if desired. Finally, the model has been checked much more thoroughly against phase data than against measurements of intensity scintillation.

A new contractual effort is being undertaken to extend and improve the irregularity model in WBMOD. Emphasis is continuing to be put on phase scintillation at auroral latitudes, but the effort extends beyond that particular topic. First priority is on describing longitudinal differences in high-latitude scintillation and related seasonal effects (Basu, 1975). A concerted attempt also will be made to extend WBMOD's irregularity description to the polar caps and the plasmapause. In conjunction with extension of the model, its results will be compared with intensity scintillation data as well as with phase data.

ACKNOWLEDGEMENT

Program WBMOD was developed under contract DNA001-79-C-0372. Potential users desiring copies of the code should address requests to the Defense Nuclear Agency, Washington, D.C., 20305; attn: Major Leon Wittwer, RAAE.

REFERENCES

Aarons, J., J. P. Mullen, and H. E. Whitney (1969), *"The Scintillation Boundary,"* J. Geophys. Res. 74 (3), 884-889.

Basu, Sunanda (1975), *"Universal Time Seasonal Variations of Auroral Zone Magnetic Activity and VHF Scintillation,"* J. Geophys. Res., 80 (34), 4725-4728.

Basu, S and S. Basu (1981), *"Modelling of Phase and Amplitude Scintillations at High Latitudes with AE-D Irregularity Data,"* paper presented at National Radio Science Meeting (Winter URSI), Boulder, CO.

Booker, H. G., and J. A. Ratcliffe, and D. H. Shinn (1950), *"Diffraction from an Irregular Screen with Applications to Ionospheric Problems,"* Phil. Trans. Roy. Soc. Ser. A., 242, 579.

Briggs, B. H., and I. A. Parkin (1963), *"On the Variation of Radio Star and Satellite Scintillations with Zenith Angle,"* J. Atmos. Terr. Phys., 25 (6), 339-366.

Cronyn, W. M. (1970), *"The Analysis of Radio Scattering and Space-Probe Observations of Small-Scale Structure in the Interplanetary Medium,"* Astrophys. J., 161, 755-762.

Fremouw, E. J. (1980), *"Geometrical Control of the Ratio of Intensity and Phase Scintillation Indices,"* J. Atmos. and Terr. Phys., 42, 775-782.

Fremouw, E. J., D. J. Barnes, R. C. Livingston, M. D. Cousins, B. C. Fair, G. K. Durfey (1974), *"Wideband Satellite Experiment — Preparations for Launch,"* Final Report--Phase III, SRI Project 1972, Stanford Research Institute, Menlo Park, CA.

Fremouw, E. J. and H. F. Bates (1971), *"Worldwide Behavior of Average VHF-UHF Scintillation,"* Rad. Sci., 6, 863-869.

Fremouw, E. J. and J. M. Lansinger (1979), *"Continued Geophysical Analysis of Coherent Satellite Scintillation Data,"* PD-NW-79-213R, Annual Report, Contract F49620-78-C-0014, Physical Dynamics Inc., Bellevue, WA.

Fremouw, E. J. and J. M. Lansinger (1980), *"Computer Modeling of Auroral-Zone Ionospheric Scintillation,"* Bimonthly Progress Report No. 3, Contract No. DNA001-79-C-0372, Project No. PD-NW-79-179, Physical Dynamics Inc., Bellevue, WA.

Fremouw, E. J. and J. M. Lansinger (1981), *"A Computer Model for High-Latitude Phase Scintillation Based on Wideband Satellite Data from Poker Flat,"* PD-NW-81-240R, Draft Final Report on Contract DNA001-79-C-0372, Physical Dynamics, Inc. Bellevue, WA.

Fremouw, E. J., J. M. Lansinger, and D. A. Miller (1980), *"Further Geophysical Analysis of Coherent Satellite Scintillation Data,"* PD-NW-81-237R, Annual Report, Contract F59620-78-C-0014, Physical Dynamics, Inc., Bellevue, WA.

Fremouw, E. J., R. L. Leadabrand, R. C. Livingston, M. D. Cousins, C. L. Rino, B. C. Fair, and R. A. Long (1978), *"Early Results from the DNA Wideband Satellite Experiment —Complex-Signal Scintillation,"* Rad. Sci., 13 (1), 167-187.

Fremouw, E. J., R. C. Livingston, and D. A. Miller (1980), *"On the Statistics of Scintillating Signals,"* J. Atmos. Terr. Phys., 42, 717-731.

Fremouw, E. J. and C. L. Rino (1973), *"An Empirical Model for Average F-Layer Scintillation at VHF/UHF,"* Radio Science, 8, March, 213-222.

Fremouw, E. J. and C. L. Rino (1976), *"Modeling of Transionospheric Radio Propagation,"* RADC-TR-76-35, Rome Air Development Center, Griffiss AFB, NY.

Fremouw, E. J. and C. L. Rino (1978), *"A Signal-Statistical and Morphological Model of Ionospheric Scintillation,"* Proc. of AGARD Conference on Operational Modeling of the Aerospace Propagation Environment, Ottawa, Canada.

Fremouw, E. J., C. L. Rino, A. R. Hessing, and V. E. Hatfield (1977a), *"A Transionospheric Communication Channel Model,"* Quarterly Technical Report 7, Contract F30602-75-C-0236, Stanford Research Institute, Menlo Park, CA.

Fremouw, E. J., C. L. Rino, R. C. Livingston, and M. C. Cousins (1977b), *"A Persistent Subauroral Scintillation Enhancement Observed in Alaska,"* Geophys. Res. Ltrs., 4 (11), 539.

Koster, J. R., I. Katsriku, and M. Tete (1966), *"Studies of the Equatorial Ionosphere Using Transmission from Active Satellites,"* Annual Summary Report 1, Contract AF61(052)-800, University of Ghana-Legon, Accra, Ghana.

Livingston, R. C., C. L. Rino, J. P. McClure, and W. B. Hanson (1981), *"Spectral Characteristics of Medium-Scale Equatorial F-Region Irregularities,"* paper submitted to J. Geophys. Res.

Martin, E. and J. Aarons (1977), *"F Layer Scintillations and the Aurora,"* J. Geophys. Res., 82, 2717-2722.

Rino, C. L. (1979), *"A Power Law Phase Screen Model for Ionospheric Scintillation. I. Weak Scatter,"* Rad. Sci., 14 (6), 1135.

Rino, C. L. and E. J. Fremouw (1977), *"The Angle Dependence of Singly Scattered Wavefields,"* J. Atmos. and Terr. Phys., 39, 859.

Rino, C. L., E. J. Fremouw, R. C. Livingston, M. D. Cousins, and B. C. Fair (1977), *"Wideband Satellite Observations,"* DNA Report 4399F, Contract No. DNA001-75-C-0111, SRI Project 3793, SRI International, Menlo Park, CA.

Rino, C. L., R. C. Livingston, and S. J. Matthews (1978), *"Evidence for Sheetlike Auroral Ionospheric Irregularities,"* Geophs. Res. Ltrs., 5 (12), 1039.

Singleton, D. G. (1970), *"Dependence of Satellite Scintillations on Zenith Angle and Azimuth,"* J. Atmos. Terr. Phys., 32, 789-803.

Singleton, D. G. (1973), *"The Dependence of High-Latitude Ionospheric Scintillations on Zenith Angle and Azimuth,"* J. Atmos. Terr. Phys., 35, 2253-2265.

Trombka, B. T. and J. C. Cain (1974), *"Computation of the IGRF: I. Spherical Expansions,"* X-992-74-303, Goddard Space Flight Center (NASA), Greenbelt, MD.

ENHANCEMENTS AND SHARP DEPLETIONS OF TOTAL ELECTRON CONTENT IN THE NIGHTTIME EQUATORIAL IONOSPHERE

D.E. Donatelli and P.M. Walsh
Regis College Research Center
Weston, Massachusetts 02193

R.S. Allen and J.A. Klobuchar
Air Force Geophysics Laboratory
Hanscom Air Force Base, Massachusetts 01731

INTRODUCTION

Over the past decade there has been a synthesis in the understanding of nighttime processes in the equatorial ionsphere using contributions from such varied observational techniques as incoherent radar backscatter, radio scintillations, optical emissions and in-situ measurements (Basu and Kelly, 1979; Fejer and Kelly, 1980 and referencestherein). Considerations from aeronomy, plasma physics and numerical simulations (Ossakow et al, 1979) illustrate, to a larger extent, the unity of the diverse observations. Utilizing the fact that most of the contribution to TEC comes from the vicinity of the F-regionpeak, TEC observations are able to contribute to present understanding of the equatorial ionosphere.

The equatorial ionosphere has been continuously monitored for over two years from Ascension Island (Lon. 15° W., Lat. 8° S.) through measurements of total electron content (TEC) using the Faraday rotation of 136 Mhz signals from the geostationary satellite SIRIO. Two characteristic features of the nighttime ionosphere will be discussed here: a post-sunset enhancement which lasts for several hours, and follows a rapid sunset decrease unattributable to the usual ionospheric decay processes; and the sudden sharp changes in the electron content along the ray path that are often superimposed on these enhancements and, to a lesser degree, on the subsequent "normal" nighttime ionosphere.

Similar measurements have been used previously in the equatorial region. Hunter (1969) in Nairobi and Kosuer (1972) in Ghana have reported: rapid sunset decreases in Faraday rotation, primarily emphasized in the Ghana data; and post-sunset enhancements which may equal or exceed daytime levels, greatly emphasized in the Nairobi observations.

The sudden, sharp variations, to be defined here as superimposed "structures", are most often seen as depletions in TEC. Similar features are shown by Tsunoda and Towle (1979) to be associated with radar backscatter plumes. Kaushika and de Mendonca (1974) and Aarons and Whitney (1980) have shown that sharp variations in TEC are coincident with scintillation patches. Yeh et al., (1979) tie these structures to F-region depletions measured by in-situ probes and the plasma "bubbles" of theoretical studies. The variation in occurence of thse structures, both seasonally and diurnally, within the 27 months of observations are examined here.

OBSERVATIONS

The data base for this study consists of over two years of continous measurements of TEC, from September 22, 1978 through December 21, 1980. The ray path to the SIRIO satellite from Ascension Island is directly north at an elevation angle of 80°. A schematic representation of the geographic and magnetic relationships is shown in figure 1. It is seen that field-aligned, plasma density structures in the F-region that cross this ray path map to altitudes of 600 to 1000 km. above the magnetic equator. The observation period is near the maximum of the present solar cycle; topside density measurements in the evening sector over the Atlantic Ocean near the maximum of the last solar cycle at altitudes greater than 600 km. (Burke et al, 1979), suggest that TEC measurements should be highly structured at Ascension Island.

Figures 2 and 3 contain examples of nighttime TEC as a function of UT, illustrating daily and seasonal variability. Data from figure 2 (3) come from five (six) consecutive days near the June (December) solstice. Local time lags UT by one hour. Ground sunset occurs at about 1845UT (1915UT).

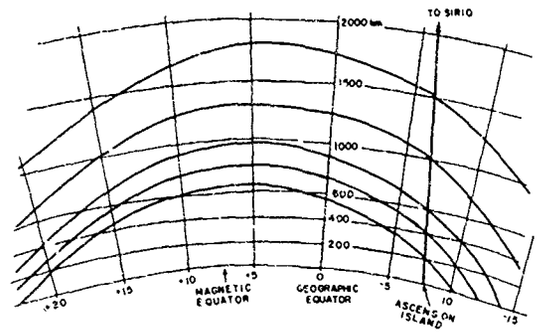


Figure 1. Schematic of Ascension Island- SIRIO ray path relationship to field lines over the magnetic equator.

A quiet, well-behaved ionosphere is expected during June at Ascension Island and found on the 10, 11, and 13 of June 1980. On some nights (June 10 and 13) the decay that begins during the sunset period persists as a gradual decrease over the entire night with only small hour-to-hour variations. On other nights (June 9, 11, and 12) there is a post-sunset increase in TEC peaking about 3 to 4 hours after sunset. Often this local maximum exceeds daytime TEC values. After the maximum, TEC usually decays to its expected pre-dawn minimum. Even during this period, however, there are examples of sudden, sharp variations in TEC. Depletions of $\sim$ 15%, accompanied by intense scintillation, appear on nights (June 9 and 12) with a particularly large post-sunset enhancement. These structures have distinct boundaries which suggest east-west dimensions of 300 to 500 km, as discussed below.

The examples of figure 3 are from December 1980. Unlike June, the atypical case is the night (December 18) of no superimposed structure or scintillation activity. On all other evenings there are post-sunset enhancements of large amplitude and 2 to 3 hours duration, the largest occurring on the 17th. This day had several structures in TEC which appear as depletions, and a large finger like enhancement. If this enhancement is part of the background the depletions are $\geq$ 30%. These structures appear as early as 2100 UT (December 13 and 17) or as late as 0100 UT (December 16). The first structure tends to have sharp boundaries in time. Structures appearing thereafter are less distinctly bounded, suggesting either a decay process or the existence of two or more structures along the ray path.

In the examples presented here, TEC could be measured continuously throughout the day including the rapidly changing evening periods. This is not always the case. On many evenings in September through March period, fluctuations in Faraday rotation occur so rapidly that the polarization angle, and thereby TEC, are indeterminate. This occurs most often with the first appearance of structure, infrequently in the late night period.

The nocturnal and seasonal variability are summarized in the histograms of figure 4. Each histogram is a plot of the percentage of nights in which structures were observed within each 12-minute time cell during the monthly period beginning with the 22nd day of each month. There is a sharp increase (decrease) in the frequency of occurrence near the September (March) equinox with a broad maximum (minimum) between equinoxes. There is a hint of solar cycle dependences in the increased frequency of occurrence for the September 1979 to March 1980 period. This behavior is similar to that for scintillation activity discussed by Aarons et al. (1980) for the 0° to 70°

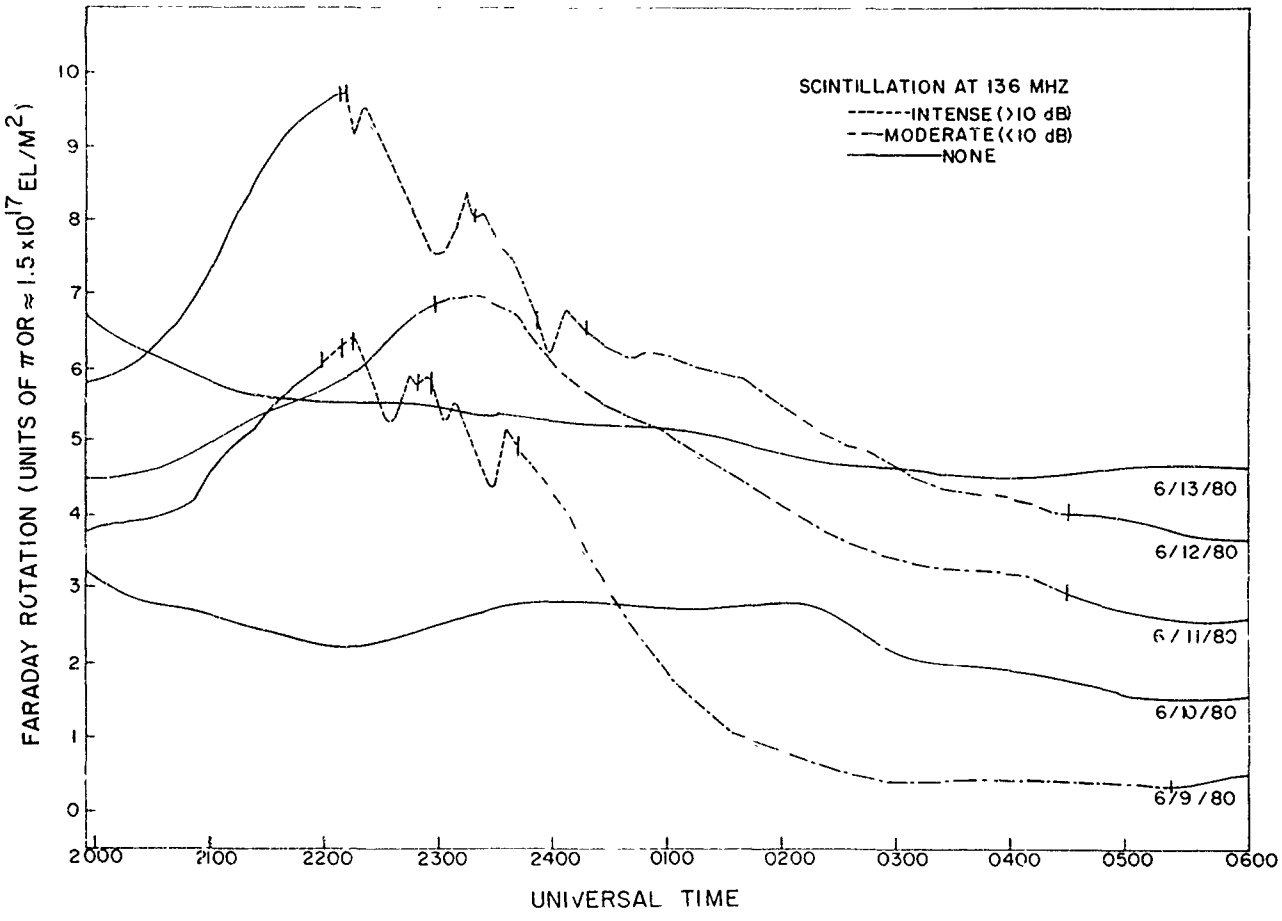


Figure 2. Measurements of nighttime TEC for five consecutive days in June 1980.

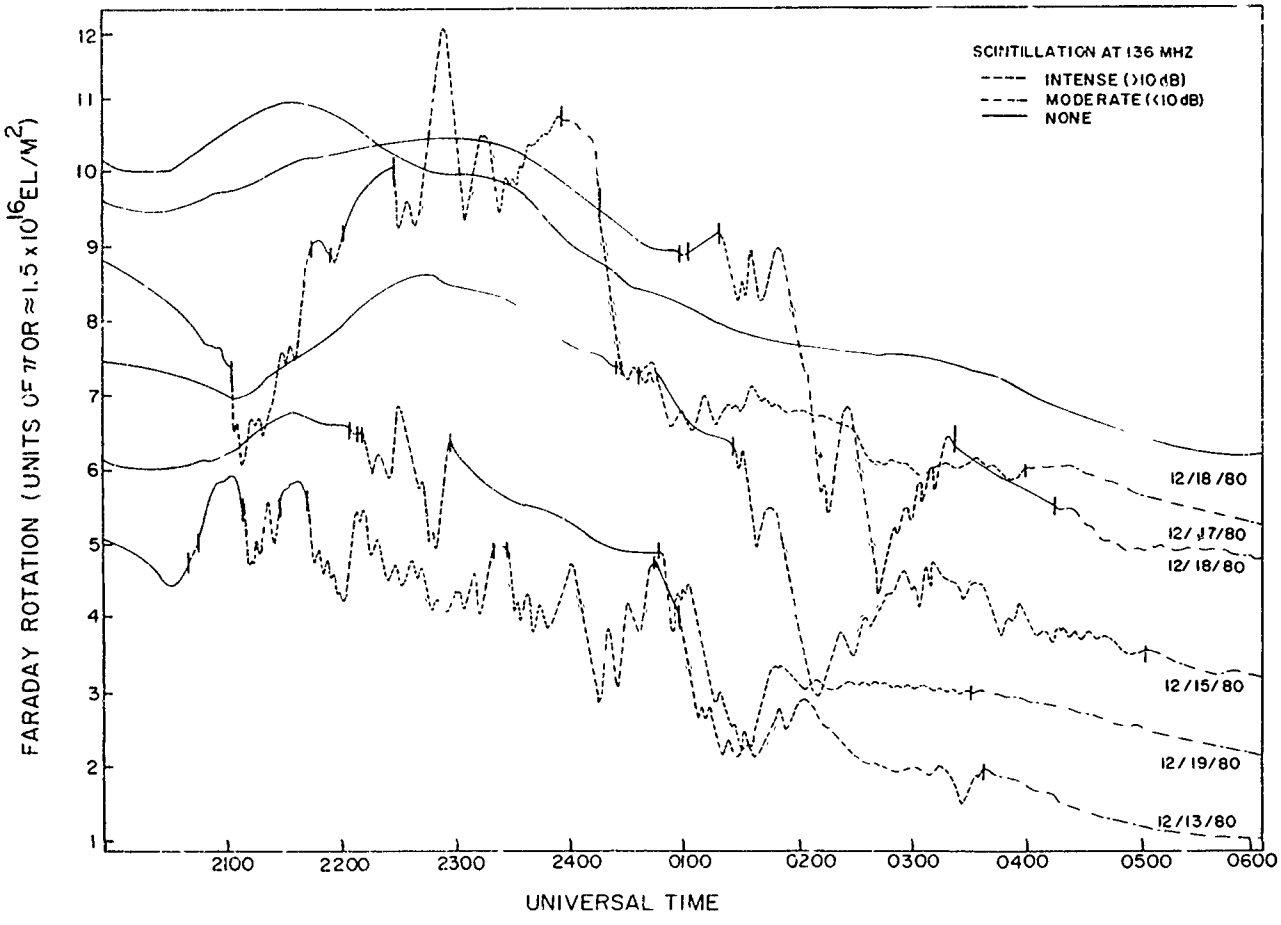


Figure 3. Measurements of nighttime TEC for six consecutive days in December 1980.

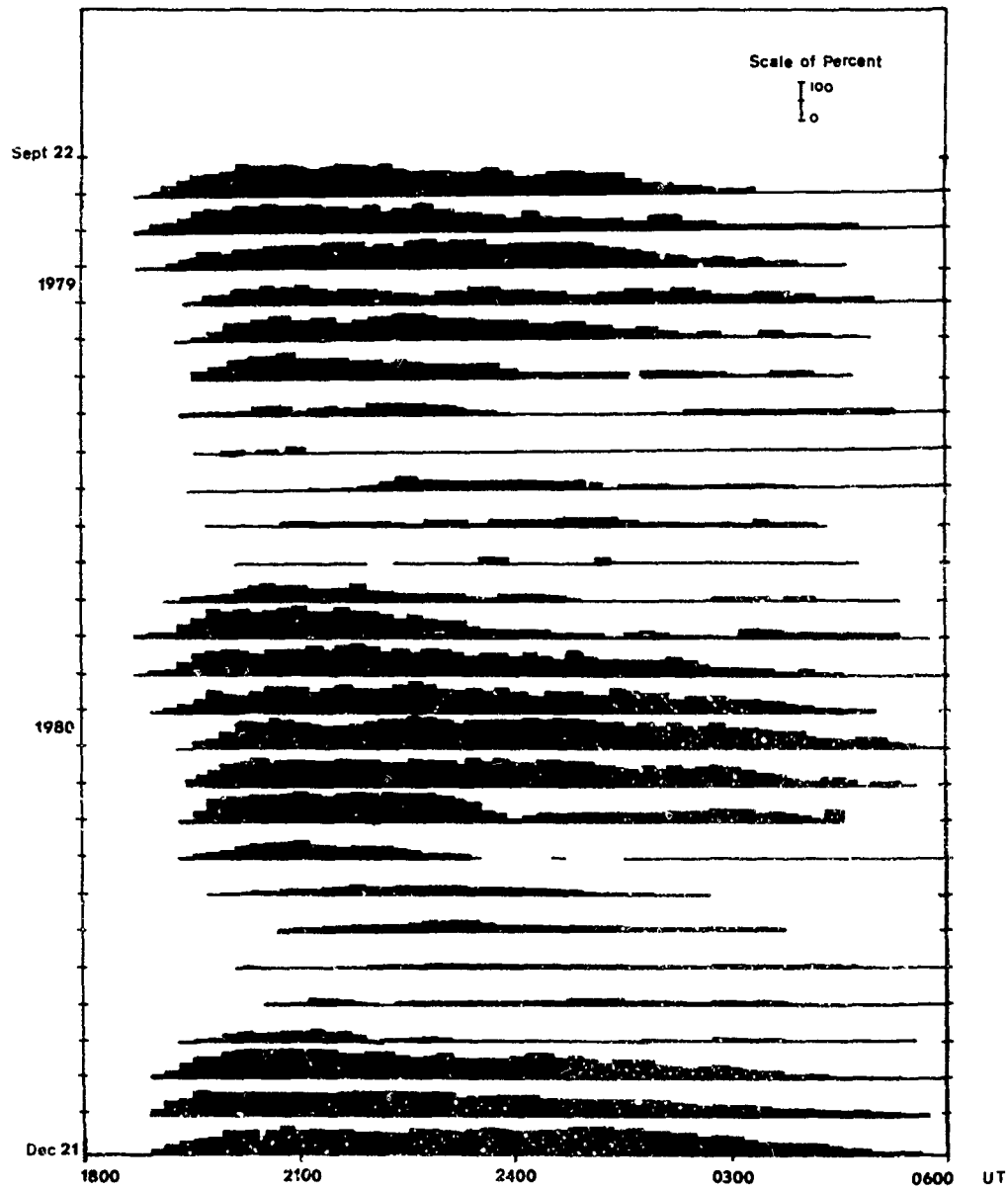


Figure 4. Histograms of percentage occurrence of structure in TEC over the nighttime for each of the 27 months of data from September 22, 1978 to December 21, 1980.

longitude sector.

These structures appear earliest in the September to December period and latest in June. The earliest structures are also those of shortest duration. Typically, they last one hour but vary from less than one minute to several hours in duration.

The nocturnal behavior shows a distinct seasonal control. A pronounced peaking in occurrence during early evening hours occurs at the equinoxes, but is observed at the solstices. This is particularly obvious near the December solstice which has an occurence frequency comparable to that near the equinoxes for the early evening hours. The early evening behavior tends to persist throughout the night, however.

DISCUSSION

There are basically two effects being considered: first, the large post-sunset enhancements; second, the sharp structures, usually depletions, associated with them. The first has been observed by Hunter(1969) and Koster(1972) and discussed by Koster and Beer(1972) and Koster(1972). They conclude that the F-region dynamo thoery(Rishbeth, 1971) offers the most plausible explanation for this effect. Since then vertical drift measurements at Jicamarca showing large enhancements in upward drift near sunset(Fejer et al 1979) give credence to Rishbeth's theory. The use of this theory by Heelis et al(1974) in modelling these drifts produced the effect observed at Jicamarca.

Rishbeth(1971, 1977) suggests that the neutral winds in the F-region produce polarization fields that short-circuit through the E-region in daytime, but build up at sunset through the sudden drop in E-region conductivity. This produces large upward drifts, resulting in a rapid lifting of the F-layer and massive transport of plasma along field lines. Near the magnetic equator this appears as a rapid sunset decline in TEC as noted by Koster(1972) with observations at dip latitude 4°. This effect is less pronounced in Hunter's(1969) observations at 13° dip latitude, and even less in the Ascension Island observations at 15° dip latitude. The subsequent enhancement, however, is more pronounced in Hunter's observations. The plasma must be transported along the field lines corresponding to the height above the magnetic equator to which the F-layer is lifted, indicating the layer is raised to altitudes $\geq$ 600 km. The topside density measurements reported by Young et al(1981) complement these observations. DMSP measurements in the post-sunset(19.5 LT) sector at 840 km. showed smooth density depletions in the immediate vicinity of the magnetic equator with enhancement in the $\pm$ 10° to 20° magnetic latitude range.

The second feature considered here is the sharp structure, often depletions,in TEC. Observational studies of the equatorial F-region suggest that these structures are plasma bubbles along the ray path. These bubbles, have been observed in the bottomside (Kelly et al, 1976), topside (Burke et al, 1979) and near the peak (McClure et al 1977) of the F layer. Presently accepted theory suggests that these begin as an E x B or collisional Rayleigh-Taylor instability in the bottomside of the F layer. The growth of the initially small perturbations depends on (1) the altitude of the initial perturbations, (2) the strength of the bottomside density gradient, (3) the altitude of the F peak, and (4) the strength of the post-sunset, eastward electric field (Ossakow et al, 1979 ; Anderson and Haerendel, 1979). Thus, large plasma depletions at and above the peak of the F layer most likely exist when the F layer is moving vertically upward.

It was noted above that most of the contributions to TEC at Ascension Island comes from flux tubes which cross the magnetic equator at altitudes between 600 and 1000 km. Young et al (1981) have found that in the Atlantic-African longitude sector, plasma bubbles appear near the magnetic equator at 840 km. on more than 80% of DMSP passes between 2000 and 2200 LT and never before 2000LT. This conicides with the local time at which the first structure of the evening appears in TEC at Ascension Island during the seasons of highest occurrence frequency.

The first structure of the evening is usually well defined, and as noted above, lasts for about one hour. This can be viewed as a structure of finite longitudinal extent drifting across the Ascension Island ray path. It has been established (Rishbeth, 1971) that post-sunset polarization electric fields cause the ionosphere to drift eastward at a faster speed than corotative. A typical value for this drift is 100 m/sec. Thus the structures appear to have longitudinal widths of about 300-400 km.

Structures that appear later in the evening usually have less well-defined boundaries. Two possible explanations are presented. First: Costa and Kelley(1978) suggest that plasma bubbles grow as large amplitude Rayleigh- Taylor waves. Sharp density gradients develop at the boundaries of the bubbles which give rise to small wavelength drift wave instabilities. The drift waves grow at the expense of the density gradient causing cross-field line diffusion, leading to decaying boundaries. Second: the ray path passes through more than one structure. This occurs if the wind speed varies with height. This would produce a shearing effect on the drifting structures, allowing portions of two or more structures to cross the ray path simultaneously.

The seasonal variation in the occurrence of these structures bears a similarity to the behavior of the vertical drift measurements of Fejer et al(1979). The most structure is observed in months for which the pre-reversal enhancement of vertical drift is greatest. This emphasizes the relationship of these structures to the evening enhancement in TEC, i.e. the enhancement as a pre-condition for the existence of structure. A clue to the difference in nocturnal behavior with season may be provided by McClure et al (1977). They observed that some "bubbles" drift more slowly than others, and in some cases, they move with the velocity of the background plasma. If the faster moving ones were to occur at the equinoxes, the slower, at the solstice, those occurring at the solstices would be longer lived.

ACKNOWLEDGEMENTS

This work was supported, in part, under the provisions of USAF contract # F19628-80-C-0016 with Regis College.

REFERENCES

Aarons, J., J.P. Mullen, J.R. Koster, R.F. da Silva, J.R. Medeiros, R.T. Medeiros, A. Bushby, J. Pantoja, and J. Lanat, Seasonal and geomagnetic control of equatorial scintillations in the 1° to 69° W longitudinal sector, J. Atmos, Terr. Phys., in press, 1980.

Aarons, J., and H.E. Whitney, Recent Observations of Equatorial and High Latitude Scintillations, AGARD Conference Proceedings # 284, Propagation Effects in Space/Earth Systems, 1980.

Anderson, D.N., and G. Haerendel, The motion of depleted plasma regions in the equatorial ionosphere J. Geophys. Res., 84, 4251, 1979.

Basu, S. and M.C. Kelley, A review of recent observations of equatorial scintillations and their relationship to current theories of F-region irregularity generation, Radio Sci., 14, 471, 1979.

Burke, W.J., D.E. Donatelli, R.C. Sagalyn, and M.C.Kelley, Observations of low density regions at high altitude in the topside equatorial ionosphere and their interpretation in terms of equatorial Spread F, Planet, Sp. Sci., 27, 593, 1979.

Costa, E., and M.C. Kelley, Linear theory for the collisionless drift wave instability with wavelenghts near the ion gyroradius, J. Geophys. Res., 83, 4365, 1978.

Fejer, B.G., D.T. Farley, R.F. Woodman, and C.Calderon, Dependence of equatorial F-region vertical drifts on season and solar cycle, J. Geophys. Res., 84, 5792, 1979.

Fejer, B.G., M.C. Kelley, Ionospheric Irregularities, Res. Geophys. Space Phys., 18, 401, 1980.

Heelis,R.H., P.C. Kendall, R.J. Moffet, D.W. Windle, and H. Rishbeth, Electric coupling of the E and F regions and its effect on F region drifts and winds, Planet. Space Sci., 22, 743, 1974.

Hunter, A.N. Faraday rotation of the 136 MHz Transmission from geostationary satellite "Canary Bird" observed at Nairobi, Radio Sci., 4, 811, 1969.

Kaushika, N.D. and F. de Mendonca, Nighttime fluctuations(scintillations) in Faraday rotation angle of VHF signals from geostationary satellites, Planet. Space Sci., 22, 1331, 1974.

Kelley, M.C., G. Haerendel, H. Kappler, A. Valenzuela, B.B.Balsley, D.A. Carter, W.L. Ecklund, C.W. Carlson, B. Hausler, and R. Torbert, Evidence for a Rayleigh-Taylor type instability and upwelling of depleted density regions during equatorial spread F, Geophys, Res, Lett., 3, 448, 1976.

Koster, J.R., Equatorial Scintillation, Planet, Space Sci., 20, 1999, 1972.

Koster, J.R.and T.Beer, An interpretation of ionospheric Faraday rotation observations at the equator, Department of Physics, University of Ghana, 1972.

McClure, J.P., W.B. Hanson, and J.F. Hoffman, Plasma bubbles and irregularities in the equatorial ionosphere, J. Geophys. Res., 82, 2650, 1977.

Ossakow, S.L., S.T. Zalesak, B.E. McDonald, and P.K. Chaturvedi ; Nonlinear equatorial spread F: Dependence on altitude of the F peak and bottomside background electron density gradient scale length, J. Geophys, Res., 84, 17, 1979.

Rishbeth. H., Polarization fields produced by winds in the equatorial F-regim, Planet. Space Sci., 19, 357, 1971

Rishbeth, H., Dynamics of the equatorial F-region J. Atmos. Terr. Phys., 39, 1159, 1977.

Tsunoda, R.T., and D.M. Towle, On the spatial relationship to 1-meter equatorial spread- F irregularities and depletions in the total electron content, Geophys. Res. Lett., 6, 873, 1979.

Young, E.R., W.J., Burke, F.J. Rich, R.C. Sagalyn, and M. Smiddy, The global distribution of equatorial spread F in the topside ionosphere during equinoctial periods, Proceedings of this Symposium, 1981.

Yeh, K.C., H. Soicher, and C.H. Liu, Observations of Equatorial Ionospheric Bubbles by the Radio Propagation Method, J. Geophys. Res., 84, 6589, 1979.

SOME RESULTS OF IONOSPHERIC TOTAL ELECTRON CONTENT AND SCINTILLATION OBSERVATIONS AT LUNPING

Yinn-Nien Huang
Telecommunication Laboratories, M.O.C.
Chung-Li P.O.Box 71, Taiwan, R.O.C.

I. INTRODUCTION

The subject of the present study is to analyze the characteristic variations of the ionospheric total electron content, slab thickness and scintillation activity observed at Lunping Observatory (geographic coordinate: 25.00°N; 121.17°E; geomagnetic coordinate: 13.8°N, 189.5°E). The total electron content data were reduced from the observed Faraday rotation angle of the VHF radio beacon signal transmitted from a geostationary satellite by use of a formula proposed by Titheridge (1972):

$$I_F = \Omega / kM_F \qquad (1)$$

where $k=2.36\times10^4/f^2$ in mks unit; Ω is the observed Faraday rotation angle; M_F is the Faraday factor determined at a fixed height of 420 km above the subionospheric point; and I_F is the total electron content integrated up to the altitude of 2000 km; f represents the frequency of the beacon signal in Hz. The total electron content (TEC) data observed at Lunping from March 4, 1977 to December 31, 1980 by measuring the Faraday rotation angle of the 136.1124 MHz beacon signal transmitted from the Japanese ETS-II geostationary satellite were used for analysis. The subionospheric point of ETS-II observed at Lunping is located at 23.0°N, 121.9°E geographic coordinate, which is near to the crest zone of the so called equatorial anomaly. The slab thickness of the ionosphere is defined by the following formula:

$$S = I_F / N_m \qquad (2)$$

where S is the slab thickness in km and N_m is the peak electron density of the ionosphere in electron/m^3. N_m is related to the F_2 layer critical frequency by the following formula:

$$N_m = (f_oF_2)^2/80.6 \qquad (3)$$

where f_oF_2 is the characteristic frequency of F_2 layer in MHz. f_oF_2 data observed at Chungli Ionosphere station (geographic coordinate: 24.95°N; 121.23°E) were used to caculate the slab thickness. The scintillation index, SI, defined by the following equation as proposed by Whitney et al. (1969) was used to represent the scintillation activity.

$$SI = P_{max} - P_{min} \qquad (4)$$

where P_{max} is the power amplitude of the third peak down from the maximum excursion and P_{min} is the power amplitude of the third level up from the minimum excursion.

II. DIURNAL, SEASONAL AND SOLAR CYCLE VARIATIONS OF THE TOTAL ELECTRON CONTENT

Figure 1 shows the monthly mean diurnal curves of TEC obtained for each month in 1978. The vertical bars show the standard deviations. The diurnal range is smallest in summer and largest in equinoctial months. Day to day

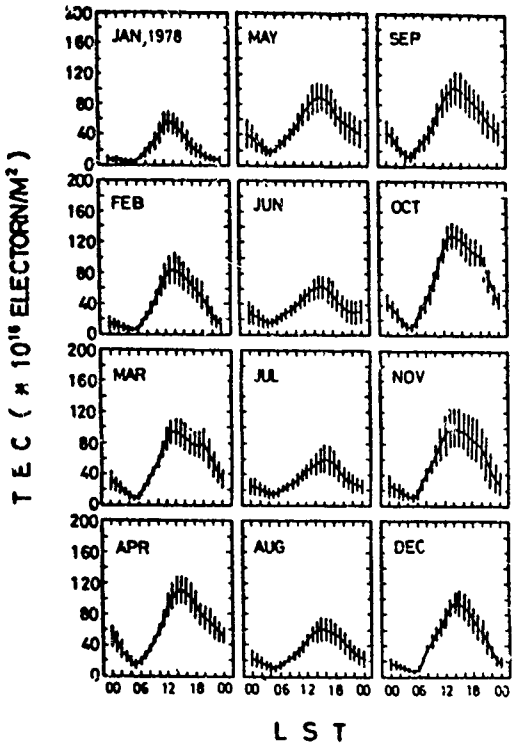


FIG1 Monthly mean diurnal curves of TEC for each month in 1978

variability as indicated by vartical bar is largest in after noon hours especially in equinoctial months. Although not show in the figure, it has been observed that the shape of the diurnal curve varies greatly with increasing solar activity (as indicated by sunspot number) in the months of equinoxes and winter. In high solar activity period, a subsidiary peak has been found to appear in the evening hours of equinoctial and winter months.

The full lines in Figure 2 show the variation of the monthly mean total electron content (denoted by I) at 00, 02, 04, 06, 08, 10, 12, 14, 16 and 20 hours in 120°E local standard time (LST=UT+ 8 hr). The apparent seasonal variation of TEC at different hours can be seen very clearly. The 12 month running average value of monthly mean sunspot number, R, denoted by $\overline{R}$. is also plotted at the top of the figure by a dashed line. The amplitude of the seasonal variation increases as the solar activity increases. The seasonal variation pattern also changes as the solar activity increases. Because of this solar activity dependence of the seasonal variation, we call it an apparent seasonal variation instead of merely a seasonal variation. The real seasonal variations will be given later. The following are the major findings about the apparent seasonal variations.

(i) The apparent seasonal variations are characterized by two maxima appearing in March/April and October/November and two minima appearing in November/December and June/July. However, there seems a clear tendency for the apparent seasonal variation to change from two-maxima and two-minima type to one-maximum and one-minimum type in the post-midnight hours of low solar activity period.

(ii) Winter anomaly, with electron content greater in winter than in summer at daytime is well developed through the whole observation period. The anomaly has been found even in evening hours (for example, 20 h).

To eliminate the seasonal variation, 12 month running average value of I, denoted by $\overline{I}$, were calculated and plotted as dotted curves in Figure 2. It is very clear that $\overline{I}$ increases with increasing $\overline{R}$, showing positive correlation between $\overline{I}$ and $\overline{R}$. The dependence of $\overline{I}$ on $\overline{R}$ can be seen more clearly in Figure 3. It is interesting to note that the rate of increase of $\overline{I}$ with increasing $\overline{R}$ changes suddenly at about $\overline{R}$ = 110. The larger the value of $\overline{I}$, the larger is the change of rate of increase.

The regression analysis was made between $\overline{I}$ and $\overline{R}$. The diurnal curves for the correlation coefficients and the slopes of the regression lines are plotted as a full line and dashed line, respectively in Figure 4. The correlation coefficients are found to be larger than 0.9 with a sharp minimum appearing around 5 hour. The diurnal variation of the slope of the regression line is characterized by a sharp maximum appearing around 21 hour, a subsidiary maximum appearing at around 17 hour and a minimum appearing around 5 hour. The appearance of maximum value at post-sunset hours is quite different from that for Sagamore Hill (Huang, 1978) and Hawaii(Huang, 1975) in which the maximum value appeared at around 13 hr. This difference may be due to the fountain effect which plays an important role on the variation of TEC at post-sunset hours (Huang, 1964) at the crest zone of the so called equatorial anomaly where Lunping locates.

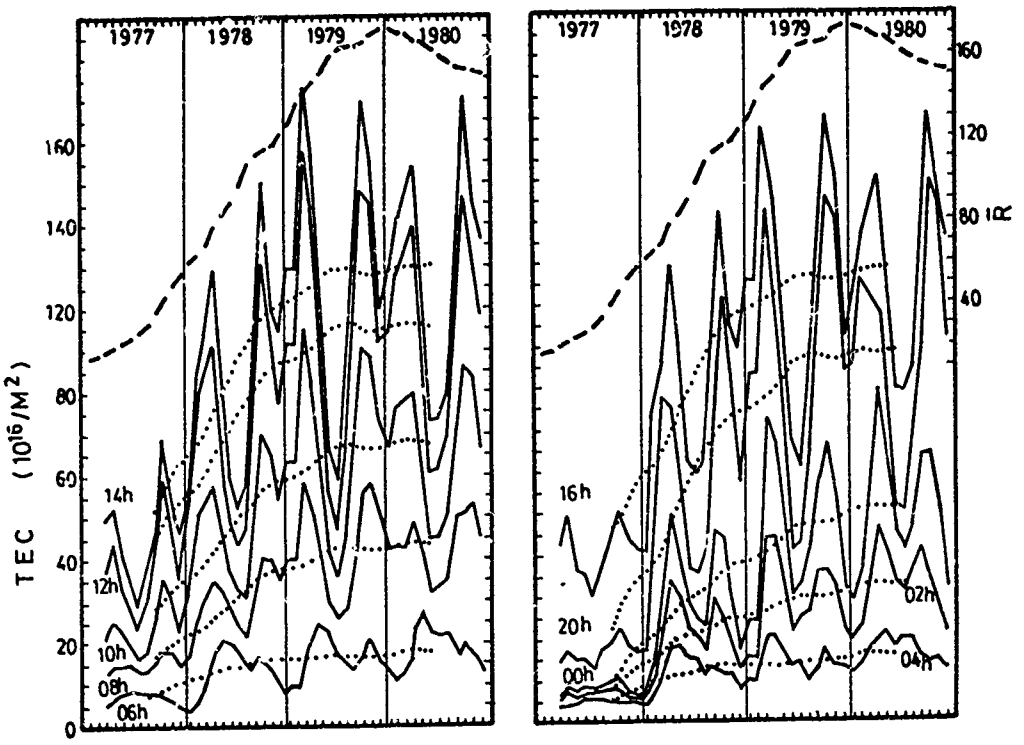


FIG2 Month to month variations of I and $\overline{I}$ at different hours as indicated. The smoothed sunspot number, $\overline{R}$, is shown at the top of the figure

The dotted line in Figure 4 shows the diurnal curve of the correlation coefficient between monthly mean hourly value of TEC, denoted by I, and monthly mean sunspot number, denoted by R. The correlation coefficients are much smaller as compared to those obtained between $\overline{I}$ and $\overline{R}$. This might be due to the fact that I includes a seasonal component while the corresponding R does not; and both $\overline{I}$ and $\overline{R}$ do not include seasonal components. It is also interesting to remark that the daytime correlation coefficients determined for Lunping are much smaller than those determined for Sagamore Hill (Huang, 1978) and Hawaii (Huang, 1975).

REAL SEASONAL VARIATIONS OF THE TOTAL ELECTRON CONTENT

It should be remarked that the apparent seasonal variations of TEC described above include the effect of the solar activity. This is the reason why we have defined them as the apparent seasonal variations to distinguish them from the real seasonal variations described below. In order to obtain real seasonal variation, we may first assume that the monthly mean hourly TEC value at n hour m month, denoted by I(n,m), and the smoothed sunspot number at m month, denoted by $\overline{R}(m)$, for each year are linearly related by the following equation:

$$I(n,m) = A(n,m) + B(n,m)\overline{R}(m) \qquad (5)$$

where A(n,m) and B(n,m) are constants to be determined by the standard method of regression analysis. Using equation (5), we may determine I(n,m) at any given value of $\overline{R}$. Using these I(n,m)s, the contour chart can be constructed and from which the real seasonal variation at any fixed solar activity of $\overline{R}$ can be observed.

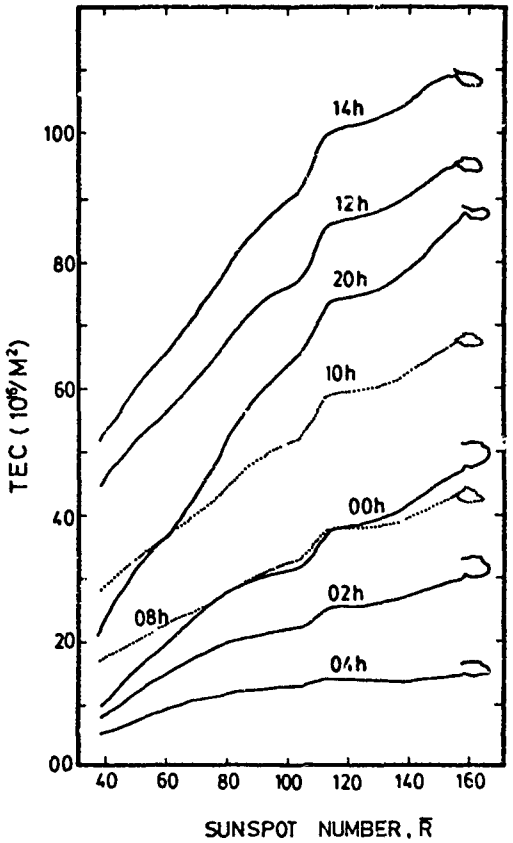


FIG3 Variation of $\bar{I}$ with $\bar{R}$ at different hours as indicated

is very strange to say that negative values appear in nighttime of October, November and December. The correlation coefficients for these hours are all greater than 0.98 showing very close linear relationship between I and $\bar{R}$. This result seems to infer that the linear relationship between I and $\bar{R}$ may change when $\bar{R}$ decreases to a certain small value; and one cannot use equation (5) to predict I(n,m) for $\bar{R}$ smaller than this value. In fact, it has been found that $\bar{R}$ should be greater than 43 in order for all I's to become positive.

Figure 7 shows the contour chart of I for $\bar{R}$=100. The following are the major characteristics: (i) The diurnal minimum line (chained line) runs nearly parallel with ground sunrise line (dashed line). (ii) The diurnal maximum line (double chained line), running parallel with ground sunset line(dashed line), locates at very different place as compared to that of diurnal maximum

In Figure 5 is given the contour chart of B(n,m) representing the rate of increase of I(n,m) per unit increase of $\bar{R}$(m). The horizontal axis represents the 120°E local standard time; and the dashed lines represent the sunrise and sunset times at ground. The chained line represents the diurnal minimum. The contour lines become almost straight lines for the period from 00 to 11 hr. showing no remarkable seasonal variations. The seasonal variation starts to appear at noon and becomes very significant at post-sunset hours, especially around 20 hr. During these hours, the seasonal variation of B(n,m) is characterized by two maxima in equinoctial months and two minima in winter and summer with winter minimum value much larger than summer one. This result shows that B(n,m) also shows winter anomaly.

Figure 6 shows the contour chart of I for the solar activity of $\bar{R}$ = 20 constructed by use of equation (5). It

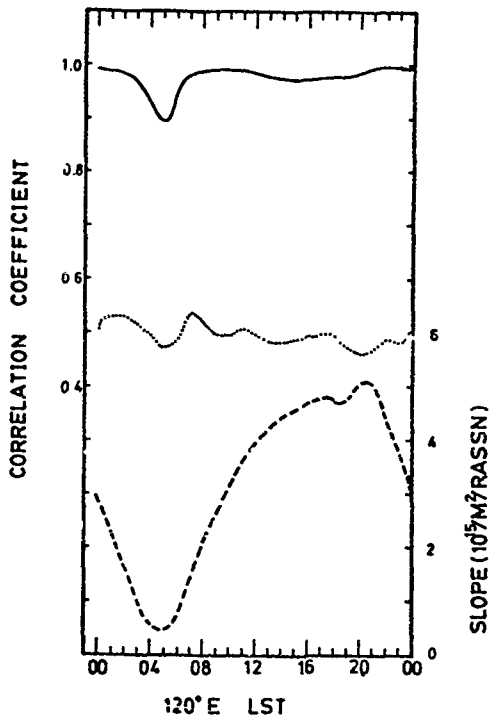


FIG4 Diurnal variations of the correlation coefficients between $\bar{I}$ and $\bar{R}$ (full line); the slope of the regression line (dashed line); and the correlation coefficients between I and R (dotted line)

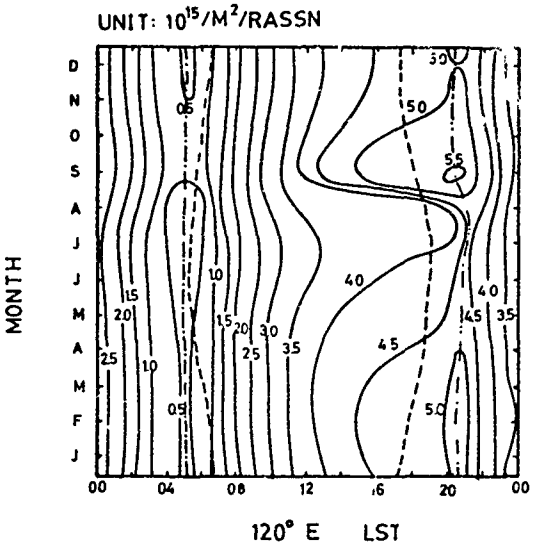


FIG5 Contour chart of the rate of increase of I with respect to $\bar{R}$. RASSN means 12 month running average value of monthly mean sunspot number

defined by:

$$D = \frac{I' - I}{I} \times 100 \ (\%) \qquad (6)$$

was used, where I' is the hourly value of electron content at local time t and I the corresponding monthly mean hourly value. Percentage deviations

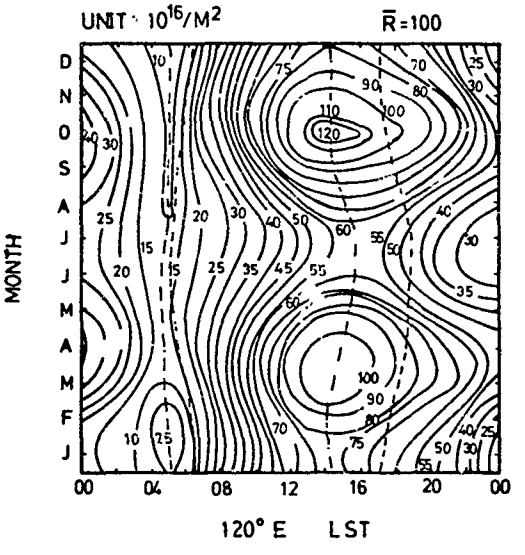


FIG7 Contour chart for I at $\bar{R}$=100

line of B(n,m). For Sagamore Hill and Hawaii, these two lines locate closely (Huang, 1975; 1978). (iii) Seasonally there are two maxima appearing in March /April and October and two minima appearing in June/July and December/ January. Winter anomaly is very significant in daytime. (iv) The October maximum is slightly larger than vernal one.

GEOMAGNETIC STORM EFFECTS

Geomagnetic storm effects on the variation of TEC were investigated by selecting 97 SC type geomagnetic storms which occured during the period from March 1977 to October 1980. As a measure of the effects of geomagnetic storm on the variations of the electron content at local time t, the quantity

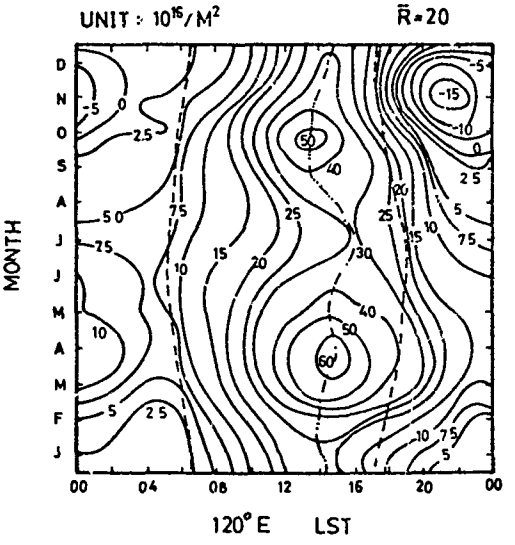


FIG6 Contour chart for I at $\bar{R}$=20

from the monthly mean instead of the simple deviations from the monthly mean were used to remove diurnal effects. The quantity D is called the disturbance variation and is a function of local time t and storm time T. As in the analysis of geomagnetic storms, it is assumed that D consists of three parts as shown below:

$$D(T,t) = D_{st}(T) + DS(T,t) + \mathcal{L} \qquad (7)$$

where D_{st} is the storm time variation, which is a function of storm time T; DS is the disturbance daily inequality and is a function of storm time T and local

time t, and Δ is an irregular variation which may be assumed to be small enough to be neglected if a large number of storms are used to obtain the mean variations. Since the DS variation is a periodic function with a period of one day or its submultiples and SSC's are distributed uniformly over the day, the storm time variation, $D_{st}(T)$, can be obtained by taking the mean of $D_i(T,t)$ over storm number i for each storm time T.

$$D_{st}(T) = \frac{1}{n}\sum_{i=1}^{n} D_i(T,t) \qquad (8)$$

where $D_i(T,t)$ denotes the disturbance variation obtained at storm time T and local time t for the storm with storm number i by use of equation (6).

The storm time variation of TEC was calculated for T=-12 hr up to T=72 hr; and the result is shown in Figure 8. The followings are the major characteristics. (i) Sudden increase of TEC starts at T=4 hr and reaches maximum value at T=8 hr then starts to decrease. (ii) Negative phase of D_{st} occurs on the second storm day and lasts until the first half part of the third storm day. (iii) Although the positive and negative phase of D_{st} can be identified as described above, there are quite a large amounts of irregular variations superposing on D_{st} variation and much more large number of storms are required to eliminate these irregular variations. (iv) Unusual negative phase is found in pre-storm time and a further collection of data is required to verify this pre-storm negative phase.

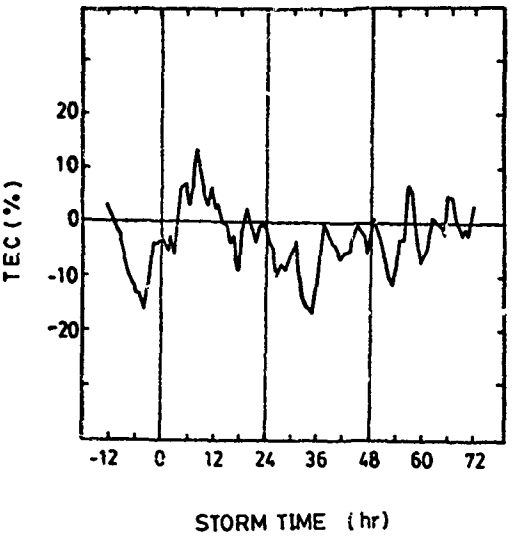


FIG8 Storm time variation of TEC

The DS variation can be obtained by subtracting D_{st} variation from D variation. However, due to large irregular variation of D_{st}, meaningful DS variation as well as the seasonal variation of D_{st} cannot be obtained before more SC magnetic storms become available.

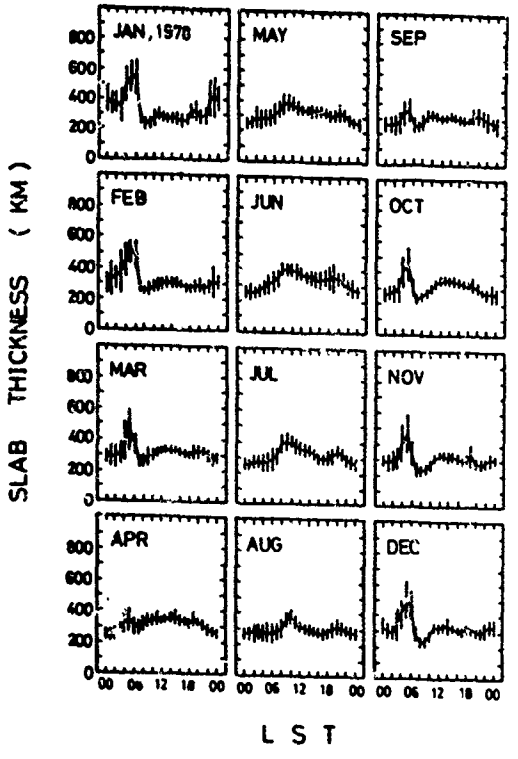


FIG9 Monthly mean diurnal curves of slab thickness for each month ir 1978

III. DIURNAL, SEASONAL AND SOLAR CYCLE VARIATIONS OF THE SLAB THICKNESS

Figure 9 shows the monthly mean diurnal curves of the slab thickness of the ionosphere obtained for each month in 1978. The vertical bars show the standard deviations and are the measures for day to day variability. It is larger in nighttime as compared to the daytime. The shape of the diurnal variation shows a remarkable seasonal change. In summer, there are one maximum appearing in pre-noon

hours and a subsidiary maximum appearing in post-sunset hours. In winter and equinoctial months, unusually large peak appearing in pre-sunrise hours. Although not shown in the figure, it has been observed that the occurrence of pre-sunrise peak depends on solar activity. In low solar activity of 1977, the pre-sunrise peak appears every month including summer months. However, in high solar activity of 1979 the pre-sunrise peak appears significantly only in March, February, January and December. The amplitude of pre-sunrise peak also depends on solar activity. It is larger in minimum solar activity as compared to the high solar activity.

The seasonal variations of slab thickness at given hour can be seen more clearly in Figure 10. The full lines show the month to month variations of the monthly mean slab thickness at even hours. The followings are the major characteristics: (i) In fore-noon hours, it is characterized by a maximum appearing in summer and a minimum in winter. (ii) From 22 hr up to pre-sunrise, the situation is reversed to have an unusually large peak appearing in December/January and a flat minimum in summer. (iii) During 12 hr to 21 hr, it is characterized by two maxima appearing in equinoxes and two minima appearing in summer and winter in high solar activity year of 1979. However, for other years, the two maxima and two minima shape is not well developed as compared to that in 1979. (iv) The amplitude of pre-sunrise peak decreases with increasing solar activity.

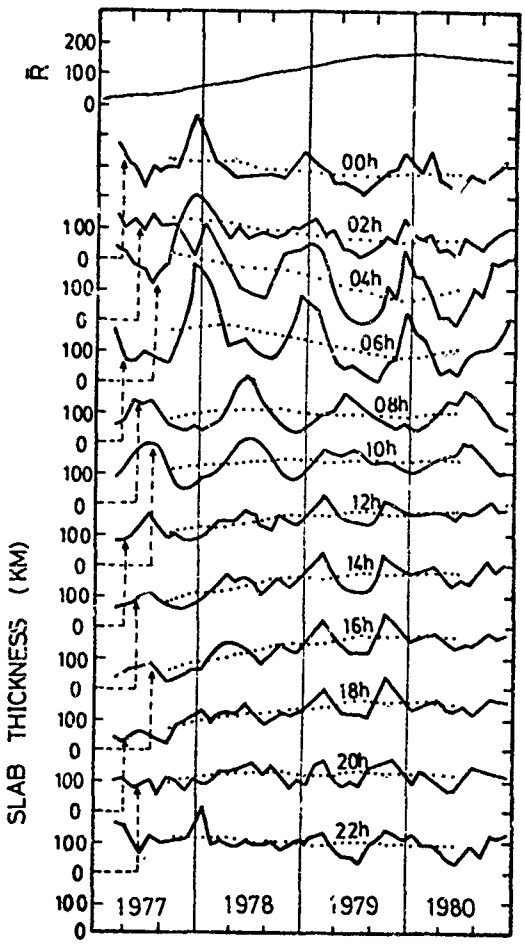


FIG 10 Month to month variations of S and $\bar{S}$ at different hours

The dotted lines in Figure 10 show the 12 month running average value, $\bar{S}$, of monthly mean hourly slab thickness, S. The smoothed sunspot number, $\bar{R}$, is also shown at the top of the figure. The relationship between $\bar{S}$ and $\bar{R}$ is found to be dependent on local time. Positive correlation between $\bar{S}$ and $\bar{R}$ is found in daytime. However, negative correlation is found in nighttime, especially for the hours when presunrise peak appears.

GEOMAGNETIC STORM EFFECT

Geomagnetic storm effects on the variation of slab thickness were investigated by chosing 90 SC type geomagnetic storms which occured during the period from March 1977 to June 1980. As a measure of the effects of geomagnetic storm on the variations of

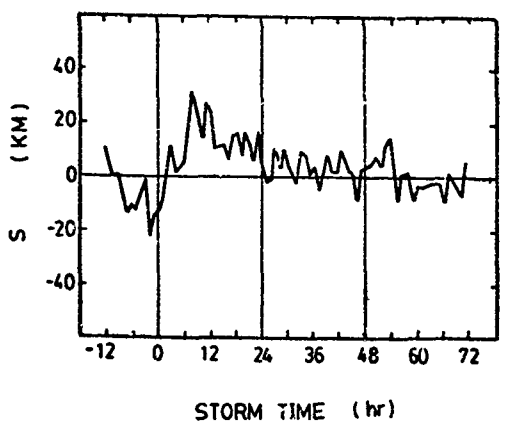


FIG 11 Storm time variation of slab thickness

the slab thickness at local time t, the quantity defined by:

$$D = S' - S \qquad (9)$$

was used, where S' is the hourly value of electron content at local time t and S the corresponding monthly mean hourly value. The storm time variation of slab thickness was calculated for T=-12 hr up to T=72 hr; and the result is shown in Figure 11. The followings are the major results: (i) Sudden increase of slab thickness starts at T=2 hr and reaches maximum value around T=5 hr, then starts to decrease. (ii) The positive phase lasts until the first quarter of the third storm day. (iii) Like D_{st} variation of TEC unusual negative phase is also found in pre-storm time and further study is required. (iv) Even though 70 SC geomagnetic storms have been used to obtain D_{st} variation of slab thickness, there are still quite a large amounts of irregular variations superposing on the D_{st} variation. (v) Due to large irregular variation in D_{st}, the DS variation of slab thickness has not been analyzed.

LUNAR DAILY VARIATIONS OF SLAB THICKNESS

The lunar daily variation L of slab thickness can be expressed by following formula (Chapman and Bartels, 1940):

$$L = \sum_{n=1}^{\infty} l_n \sin\left[t(n-2) + 2\tau + \lambda_n\right] \qquad (10)$$

where the subscript n represents the nth harmonic component; l_n and λ_n represent the amplitude and phase angle of the nth component of the lunar daily variation; t represents the solar time in hour, increasing from 00 to 24 from one local mid-night to next; and τ represents the lunar time in hour, increasing from 00 to 24 from one local lower transit of the moon to the next; t and τ are related by

$$\nu = t - \tau \qquad (11)$$

where ν is the phase of the moon measured by hour angle between the sun and the moon increasing from 00 at one new moon to 24 at the next. In terms of ν, equation (10) can be rewritten as

$$L = \sum_{n=1}^{\infty} l_n \sin\left(n\tau - 2\nu + \lambda_n\right) \qquad (12)$$

The most important component of the lunar daily variation has a period of half a lunar day and is expressible as

$$L_2 = l_2 \sin\left(2\tau - \lambda_2\right) \qquad (13)$$

Besides L_2, which is a purely lunar daily variation, L has a part $(L-L_2)$ which is dependent on both lunar and solar time and is called lunisolar component.

In the present report, the Chapman and Miller method (Chapman and Miller, 1940; Malin and Chapman, 1970) was used to determine the harmonic components and respective probable errors of the lunar daily variations up to n=4. Table 1

TABLE 1

	No. of	L_1			L_2			L_3			L_4		
	Days	l_1	e_1	λ_1	l_2	e_2	λ_2	l_3	e_3	λ_3	l_4	e_4	λ_4
All Data	1217	3.24	1.11	19	3.20	0.98	174	0.79	1.07	252	2.08	0.59	235
Equinoxes	426	(4.18)	2.78	53	3.77	1.79	147	(2.31)	1.56	224	2.19	0.94	174
Summer	429	5.57	1.99	15	4.06	1.28	155	(2.21)	1.61	69	3.45	1.07	282
Winter	360	(3.23)	2.30	305	5.09	1.62	232	(2.72)	1.52	264	(2.32)	1.38	214

summarizes the result. The letters l_n, e_n and λ_n represent the amplitude, probable error and phase angle of the nth harmonic component of the lunar daily variation of slab thickness. The amplitude can be considered to be statistically significant at 5% level, if it is larger than 2.08 times the probable error

(Leaton et al., 1962). The statistically insignificant values in Table 1 are in parenthesis. Unlike the lunar daily variations obtained for TEC at Lunping (Huang, 1979), only semi-diurnal lunar component, L_2, of the lunar daily variations of slab thickness at Lunping is statistically significant for all data inclusive and for all seasons separately. The result is shown by a harmonic dial in Figure 12. The letter W, E, S and A represent the semi-diurnal lunar components determined for winter, equinoxes, summer months and the complete year. The distance from the dial point to the origin gives the amplitude of the component. The position of the dial point with respect to the hours on the dial face gives the solar time at which the variation attain its maximum value.

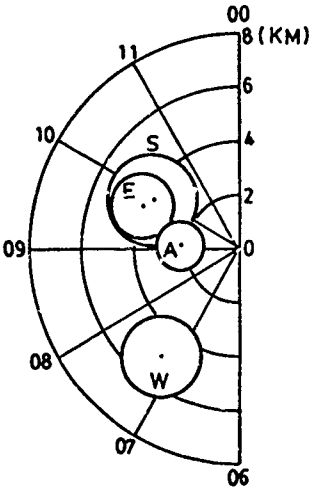


FIG 12 Harmonic dial of the semi-diurnal lunar component of slab thickness

IV. DIURNAL, SEASONAL AND SOLAR CYCLE VARIATIONS OF IONOSPHERIC SCINTILLATIONS

Figure 13 shows the diurnal variations of the monthly mean scintillation index for the month from March, 1977 to February 1978. The diurnal variation is characterized by a large maximum appearing around midnight and a small subsidiary maximum appearing around noon. The daytime subsidiary maximum is most profound during the summer months and almost disappears in winter months. The diurnal variation shows a remarkable seasonal variation. The nighttime maximum is largest in summer month and smallest in winter months. The pattern of seasonal variation changes with increasing solar activity as will be shown below. Figure 14 shows the month to month variation of monthly mean scintillation index obtained for the time intervals 00-02; 12-14 and 20-22 hr, representing post-midnight; daytime and post-sunset scintillation activity. The followings are the major characteristics: (i) The seasonal variation of the post-midnight scintillation activity is characterized by a maximum appearing in summer and a minimum in winter for the lower solar activity of 1977. However, as the solar activity increases (for example 1978 onward), the seasonal variation changes to have two maxima appearing in equinoctial months and two minima appearing in summer and winter. The summer minimum is slightly larger than the winter minimum. (ii) The seasonal variation of daytime scintillation activity is characterized by a single maximum appearing in summer. A small subsidiary maximum, though not very significant, is also found in winter. (iii) The post-sunset scintillation activity is characterized by two maxima appearing in equinoctial months and two minima appearing in winter and summer for each year

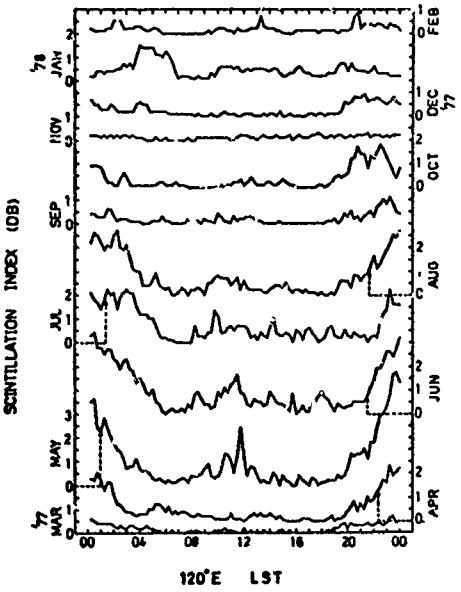


FIG 13 Diurnal variations of the monthly mean quarter-hourly scintillation index for each month from March 1977 to February 1978.

including 1977. The winter and summer minima have almost same values.

The dotted line in Figure 14 shows the 12 months running average value of the monthly mean scintillation index. The correlation between the smoothed scintillation index and sunspot number is quite irregular except for daytime scintillation in which a negative correlation between them can be seen.

Figure 15 shows the diurnal variations of the percentage occurrence of the scintillation with scintillation index greater or equal to certain limiting values of 0.5 dB, 4 dB and 8 dB as indicated by full, dotted and dashed lines, respectively. The diurnal and seasonal variations of the percentage occurrence of scintillation are similar to those of scintillation index. The daytime occurrence of scintillation is mainly due to the occurrence of intense Sporadic E; and the nighttime occurrence of scintillation is mainly due to irregularities in F region (Huang, 1978).

The storm effect on the scintillation activity was investigated by comparing day to day variations of magnetic ap index to those of scintillation index. No definite relationship between geomagnetic activity and scintillation activity has been found. Sometime, the scintillation activity increases with increasing geomagnetic activity showing a positive correlation between the two; however, for another time, no scintillation occurs even for large increase of geomagnetic activity showing almost no effect of the geomagnetic activity on the scintillation activity. Sometime, sever scintillation can occur even during geomagnetically quiet period. The geomagnetic effect on the scintil-

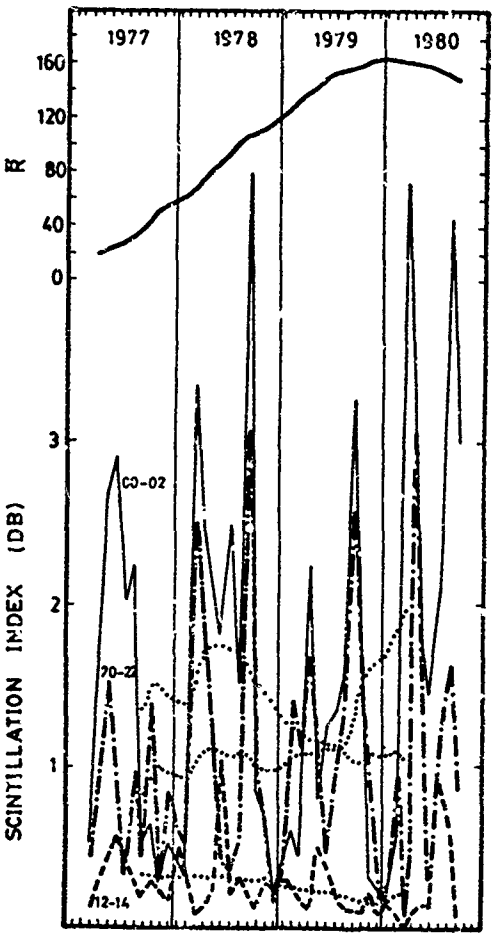


FIG 14 Month to month variations of monthly mean scintillation index for the given time interval as indicated

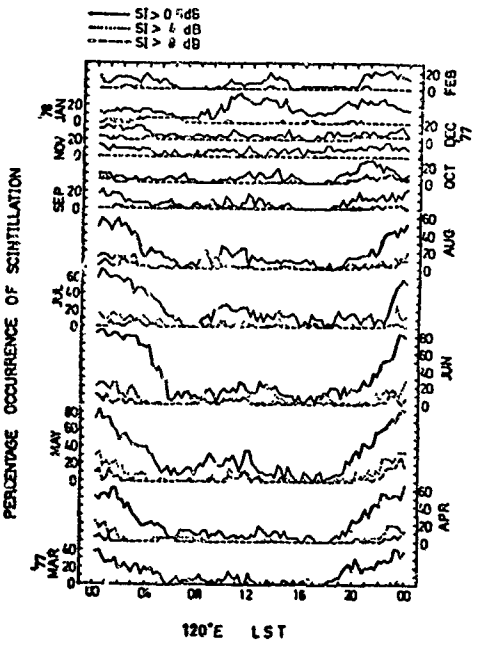


FIG 15 Diurnal variations of the percentage occurrence of scintillation index greater or equal to certain limiting values as indicated

lation activity seems very random and irregula .

ACKNOWLEDGEMENT

The geostationary satellite ETS-II was launched and station kept by the National Space Development Agency of Japan. The initial polarization angle of the VHF beacon signal of ETS-II was provided by Dr. Kenji Sinno of the Radio Research Laboratories, Japan. Without his assistance the absolute value of TEC could not be obtained.

REFERENCES

Chapman, S., and J. Bartels, Geomagnetism, Oxford at the Clarendon Press, London, 1940

Chapman, S., J. C. Gupta, and S. R. C. Malin, The sunspot cycle influence on the solar and lunar daily geomagnetic variations, Proc. Roy. Soc. London, A, 324, 1, 1971

Chapman, S., and J. C. P. Miller, The statistical determination of lunar daily variations in geomagnetic and meteorological elements, Mon. Notic. Roy. Astronom. Soc., Geophys. Suppl., 4, 649, 1940.

Huang, Y. N., Postsunset rise of f_0F2 and geomagnetic activity, J. Geophys. Res., Vol.69, 3633, 1964.

Huang, Y. N., Solar cycle variations of the total electron content at low latitude, Effect of the Ionosphere on Space Systems and Communications, 355, 1975.

Huang, Y. N., Solar cycle variation in the total electron content at Sagamore Hill, Jour. of Atmospheric and Terrestrial Physics, Vol.40, 733, 1978.

Huang, Y. N., Ionospheric Scintillations at Lunping, Journal of the Chinese Institute of Engineers, Vol.1, No.2, 81, 1978.

Huang, Y. N., Solar cycle and seasonal variations of the solar and lunar daily variations of total electron content at Lunping, Jour. of Geophysical Research, Vol. 84, 6595, 1979.

Leaton, B. R., S. R. Malin, and H. F. Finch, The solar and luni-solar daily variation of the geomagnetic field at Greenwich and Abinger, 1916-1957, Roy. Observ. Bull. 63, D273, 1962.

Titheridge, J. E., Determination of ionospheric electron content from the Faraday rotation of geostationary satellite signals, Planet. Space Sci., 20, 353, 1972.

Whitney, H. E., J. Aarons and C. Malik, A proposed index for measuring ionospheric scintillation, Planet, Space Sci., Vol.17, 1069, 1969.

SCINTILLATION OF SATELLITE SIGNALS AT GUAM FOR TWO ELEVATION ANGLES AND TWO FREQUENCIES

M. R. Paulson
Naval Ocean Systems Center
San Diego, CA 92152

ABSTRACT

Satellite signals at 257.55 and 1541.5 MHz were recorded at Guam from late February through December 1979 for two satellites, one at an elevation angle of about 10 degrees and the other at about 50 degrees. These records were used to determine daily occurrence and seasonal variation of equatorial scintillation at the two frequencies and the two elevation angles.

Seasonal variation in the occurrence of scintillation at 257.55 MHz showed a broad maximum from late February through October, with a minimum occurring from November through January. At 1541.5 MHz the occurrence showed two maxima, one in April and the other in September.

A periodic variation was observed in the occurrence of scintillation for both satellites, particularly during the first half of the year. It is proposed that this periodicity may be caused by the moon's gravity modulating the zonal winds, either increasing them or decreasing them, depending on the location of the moon with respect to the sun. This could then increase or decrease the scintillation occurrence and intensity by increasing or decreasing the turbulence at F-region heights.

A periodic variation observed in the geomagnetic activity may also be contributing to the periodic variation in the scintillation.

INTRODUCTION

Satellite signals at 257.55 and 1541.5 MHz were recorded by personnel at the NASA Satellite Tracking Station, Dandan, Guam from late February through December 1979. Signal amplitudes were recorded from two MariSat satellites, one at an elevation angle of about 10 degrees and the other at about 50 degrees. Additionally, signals at the 257.55 MHz frequency were recorded by personnel at the Naval Communications Area Master Station (NAVCAMS), Guam.

These data have been used to evaluate daily occurrence and seasonal variation in the occurrence of equatorial scintillation at the two frequencies and the two elevation angles. The data have also been used to look for relationships between scintillation and other solar and geophysical phenomena.

MEASUREMENT TECHNIQUE

The data taken at the NASA Satellite Tracking Station were recorded on strip-chart recorders 24 hours per day. These recorders were paralleled by analog magnetic tape recorders which were started at 1900 hours local time each night and allowed to run for the duration of a reel of tape, which was about 8 hours. The NAVCAMS data were recorded on strip-chart recorders 24 hours per day.

DATA ANALYSIS AND RESULTS

Daily Occurrence of Scintillation

The starting times and ending times for periods of uhf scintillation with fades greater than 6 dB were read from the strip charts for each of the satel-

lites. These times were summed on a daily basis to get the total number of hours each day that scintillation fading exceeded 6 dB. No attempt was made to distinguish between F-region and Sporadic-E scintillation for the low elevation angle (Indian Ocean) satellite. The results are shown for the Indian Ocean satellite, F2, in the top graph of figure 1 and for the Pacific Ocean satellite, F3, in the middle graph. The bottom graph shows the measured daily 2800 MHz solar flux taken from NOAA Solar Geophysical Data. Correlations between these solar flux measurements and the daily occurrence of scintillation for the two satellites gave values of about -0.1 to -0.2 when data for most of the year were used. These correlations are shown in figure 2.

Seasonal Variation of Scintillation

Sixty-one-day running averages of the data shown in figure 1 were calculated and plotted in figure 3 to investigate the seasonal variation of equatorial scintillation. These curves show a minimum of scintillation in December and January, with the occurrence increasing rapidly in February and March to a broad maximum through the local summer, and then decreasing again in October and November. This can be compared to what has been found in some other equatorial regions as reported by Aarons (1977) in a review of equatorial scintillation.

In Huancayo the occurrence showed a broad maximum from October through March with a minimum in the June-July period and only a slight decrease in activity in December (Mullen and Hawkins, 1975).

Livingston (1980) also found the occurrence to have a minimum in the June-July period at Ancon, Peru while at Kwajalein the seasonal variation had a minimum in the December-January period. The latter is similar to the observations at Guam. He reports that the seasonal variation in scintillation at Ancon and Kwajalein were quite similar, except that there was a 6 month shift between the two locations.

In Ghana, Koster (1972) found maxima around the equinoxes with minima during the solstices. The minimum around the June solstice was greater than that around the December solstice, however.

Basu et al (1976) have used in-situ measurements of ionospheric irregularities, measured during November and December, 1969, to estimate the occurrence of equatorial scintillation >4.5 dB as a function of latitude and longitude. Their results show extensive scintillation activity in the Huancayo area and somewhat less activity in the Ghana area and at the same time show very low scintillation activity in the Guam and Kwajalein areas. This is consistent with the occurrence-of-scintillation curves shown in figure 3.

Some asymmetry is evident in the scintillation curves shown in figure 3, with a higher occurrence in September than in the March-April period. This is particularly noticeable on the Indian Ocean satellite plot. This may be the result of variation in solar activity. In the bottom graph, which is a sixty-one-day running average of the 2800 MHz solar flux, the smoothed solar activity decreases from late January through the middle of July and then increases rapidly through the middle of October.

Another way of showing the seasonal variation of scintillation is plotted in figure 4. There the hours of scintillation occurrence when fades exceeded 6 dB are totaled for each month and shown as the percent of time that scintillation occurred that month. The results are quite similar to those shown in figure 3 with the occurrence being somewhat higher in September than in March or April. If the occurrence of scintillation is compared for the month of September for the two plots, they show scintillation occurring 22 percent of the time for the 10 degree elevation angle satellite compared to 17 percent for the 50 degree elevation angle satellite.

This same method of presentation was used for the L-band data to compare it to that at 257.55 MHz. Figure 5 shows the monthly percent of occurrence of L-band scintillation with fades greater than 6 dB for each of the satellites. These plots show maxima in April and September, with a minimum in the June-July period for the Indian Ocean satellite. The occurrence was so low for the Pacific Ocean satellite that the differences between March, April and May are probably not significant.

The occurrence during September was about 3 percent compared to more than 22 percent at the uhf on the 10 degree elevation angle satellite and less than 0.4 percent compared to 17 percent on the 50 degree elevation angle satellite. This gives some indication of the improvement that can be expected at the higher frequency.

Lunar Dependence of Scintillation

Various running averages of from 5 to 30 days were calculated for the scintillation data in figure 1 to try to identify other possible phenomena affecting the occurrence of scintillation. The top graph in figure 6 shows a plot for the Indian Ocean satellite where a running average of 11 days was used. A pronounced periodicity is seen in this curve, particularly for the first half of the year. A spectrum analysis of the data showed a line with a period near 30 days, suggesting that this periodicity might be connected with the phases of the moon. To investigate this, a sine wave plot is superimposed on the curve to show the phases of the moon, with the maximum corresponding to full moon and the minimum to new moon. The Pacific Ocean satellite data were treated in the same way, but since the results were similar they are not shown.

Cross correlations were calculated between phases of the moon and the scintillation curves for two different data segments. These are shown in figure 7 for the Indian Ocean satellite. For the data sample from Julian day 30 to day 150 the correlation for the Indian Ocean satellite was 0.75 occurring 3 days after full moon and for the Pacific Ocean satellite it was 0.68 occurring 4 days after full moon. For the data sample from Julian day 150 to day 300 these values were 0.34 and 8 days after full moon for the Indian Ocean satellite and 0.20 and 5 days after full moon for the Pacific Ocean satellite.

From the preceeding discussion it appears that the moon may have had an influence on the occurrence and intensity of the equatorial scintillation, particularly during the first half of the year. From July through October some additional factor may have become important, or the lunar influence may have become less. One possibility can be seen in the bottom graph of figure 6. From June on, a noticeable periodic variation in solar activity is present. This may have produced a variation in scintillation which, when combined with any lunar variation, could have confused the picture.

If the moon is causing variations in the occurrence of scintillation, the most probable way that it could do this is for the moon's gravity to modulate the zonal winds, either increasing them or decreasing them, depending on the location of the moon with respect to the sun. This could then increase or decrease the scintillation occurrence and intensity by increasing or decreasing the turbulence at F-region heights.

Scintillation Versus Geomagnetic Activity

Another possible explanation for the periodic variation in the scintillation can be seen in the middle graph of figure 6, which is a running average of the daily average of the 3-hour magnetic Kp indices obtained from NOAA Solar Geophysical Data. The January to June data show a periodic variation in activity which is very nearly out of phase with the variation in the scintillation activity. Cross correlations were calculated between the unaveraged geomagnetic data and the unaveraged scintillation data for the Indian Ocean satellite for two data segments. For the data sample from Julian day 30 to day 180 the correlation was -0.42 and for the data from day 180 to day 300 it was -0.18, both with zero time delay.

To further investigate the relationship between scintillation and geomagnetic activity, daily total hours of occurrence of scintillation were correlated with the daily sums of the 8 3-hour Kp indices. The data were taken in sample periods of 30 days at 5-day intervals and each correlation value plotted for the center of the sample period. The results are shown in figure 8.

These correlation curves start out near zero, or slightly positive. in January and then go quickly negative in February with values as large as -0.75 for the Indian Ocean satellite, F2, and -0.6 for the Pacific Ocean satellite, F3. Values for F3 remain quite negative through the middle of June, then go slightly positive in July before going negative again in September with values as large as -0.6. In November and December they become positive again with values up to 0.2 and 0.4.

The curve for F2 is quite similar to that for F3, but there are some differences. These differences may result from the fact that the two satellites had different look angles, the elevation angle to F2 being about 10 degrees and to F3 being about 50 degrees. Additionally, because of its low look angle, F2 was susceptible to the occurrence of sporadic-E scintillation.

It should be mentioned here that, with a data sample of 30, correlation values of about 0.4 are significant at the 1 percent level and values of about 0.3 at the 5 percent level.

SUMMARY

Several factors appear to be affecting the occurrence of equatorial scintillation at the same time. Because of this it is difficult to separate and identify the contribution of each.

The seasonal variation in uhf scintillation at Guam showed a minimum in December and January, then a rapid increase in February and March with a broad maximum through the local summer and a rapid decrease in late October and November. This is in good agreement with the findings at Kwajalein, but about 6 months out of phase with those in the American sector, where the minimum occurred in the June-July period.

The seasonal variation in the occurrence of L-band scintillation at Guam showed two maxima, one in April and one in September, with a decrease in occurrence during the June-July period. No L-band data were available in the December-January period.

A periodicity was evident in the occurrence of scintillation data, particularly during the first half of the year. It is proposed here that this variation may have been caused by the moon's gravity increasing or decreasing (depending on the moon's location with repect to the sun) the zonal winds at F-region heights and thus increasing or decreasing the scintillation by increasing or decreasing turbulence.

A second possible cause for the periodic variation in the occurrence of scintillation was a periodic variation in the geomagnetic activity occurring during the first half of the year. This variation was very nearly out of phase with the variation in the occurrence of scintillation.

The periodic variations in the scintillation may also result from the combined effects of geomagnetic activity and the lunar gravity. Additional work needs to be done to evaluate this. In particular, thermospheric winds need to be recorded simultaneously with equatorial scintillation measurements in the area where the measurements are being made. Then both scintillation intensity and thermospheric winds should be compared with geomagnetic activity and phases of the moon, as well as with each other.

REFERENCES

Aarons, J. (1976), Equatorial scintillation: A review, IEEE Trans. Antennas Propagat., AP-24, 729-736.

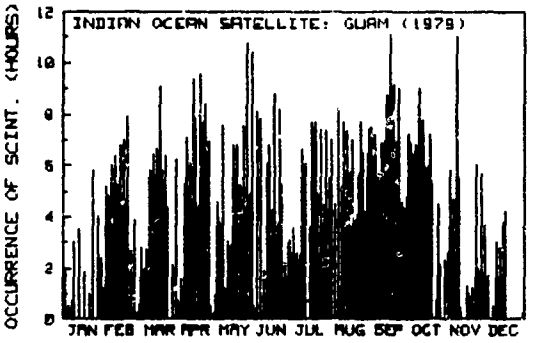
Basu, S., S. Basu, B. K. Khan (1976), Model of equatorial scintillations from in-situ measurements, Radio Science, 11, 821-832.

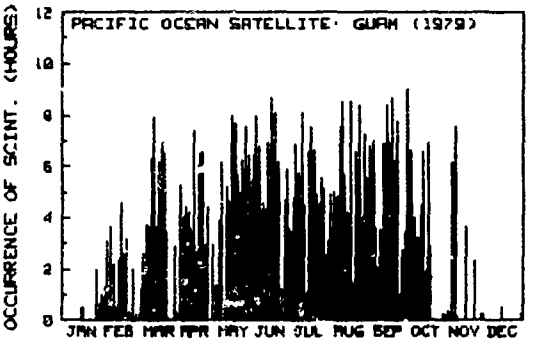
Koster, J. R. (1972), Equatorial Scintillation, Planet. Space Sci., 20, 1999-2014.

Livingston, R. C. (1980), Comparison of multifrequency equatorial scintillation, American and Pacific sectors, Radio Science, 15, 801-814.

Mullen, J. and G. Hawkins (1975), Equatorial scintillation at 136 MHz observed over half a sunspot cycle, in Effect of the Ionosphere on Space Systems and Communications, edited by John Goodman, NTIS CSCL 04/1 N75-30714#, Naval Research Lab., Washington, DC.

National Oceanic and Atmospheric Administration Environmental Data and Information Service, Solar-Geophysical Data, Prompt Reports, Boulder, CO.





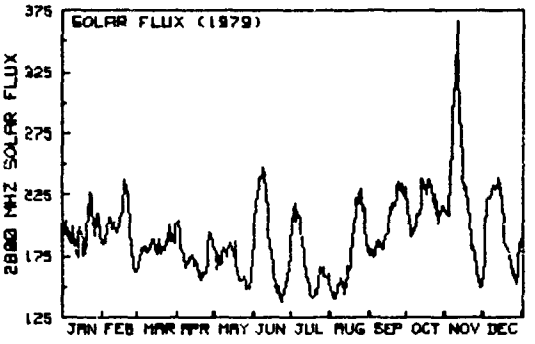


Figure 1. Total hours each day that uhf scintillation fading exceeded 6 dB. Bottom graph is the daily value of the 2800-MHz solar flux measurement.

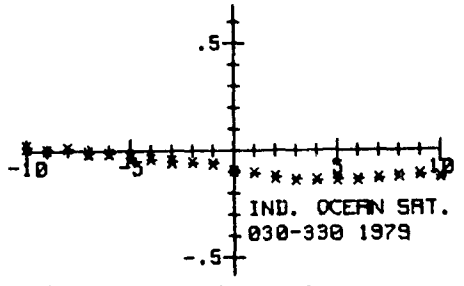


Figure 2. Cross correlations between the 2800-MHz solar flux data and the daily occurrence of scintillation for the two satellites.

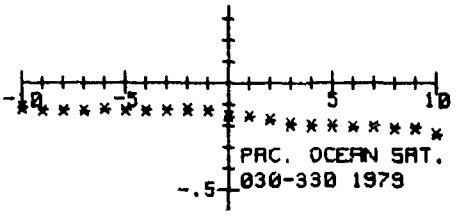
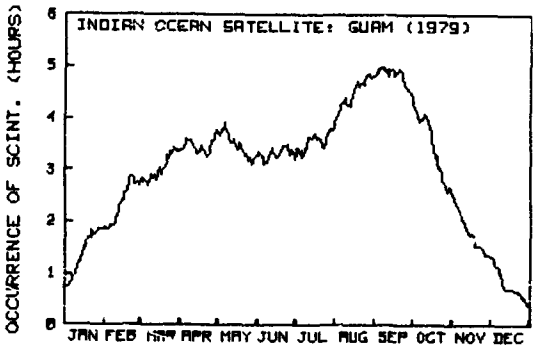
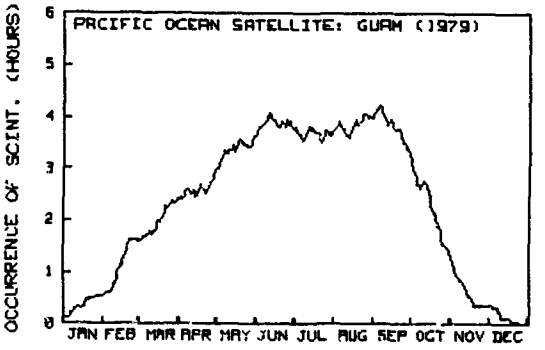


Figure 2. (Cont)





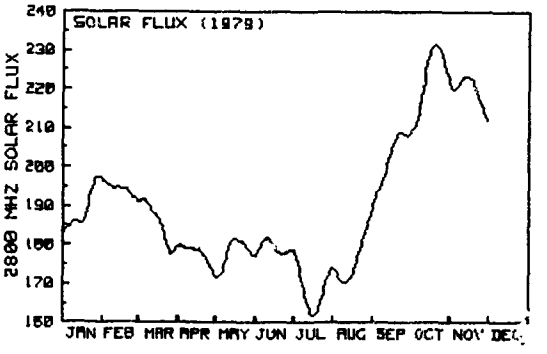


Figure 3. Sixty-one-day running averages of daily occurrences of uhf scintillation for the two satellites. Bottom graph is a 61-day running average of the 2800-MHz solar flux.

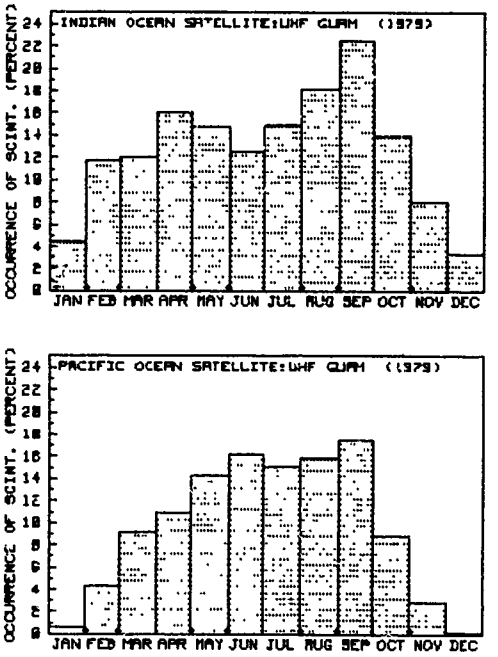


Figure 4. Monthly percent of time uhf scintillation occurred with fades greater than 6 dB for the two satellites.

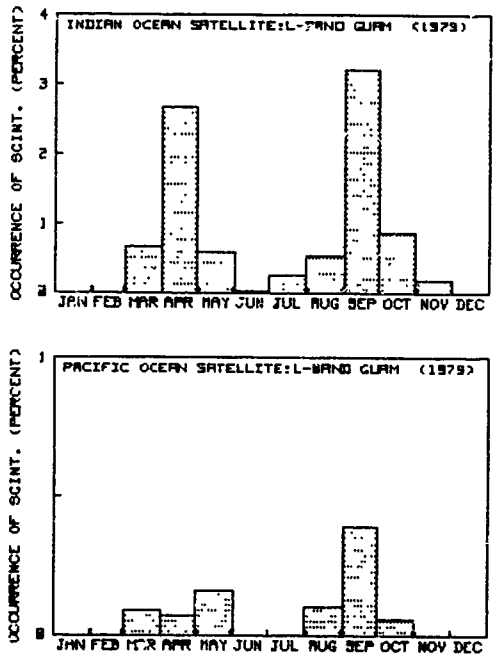


Figure 5. Monthly percent of time L-band scintillation occurred with fades greater than 6 dB for the two satellites.

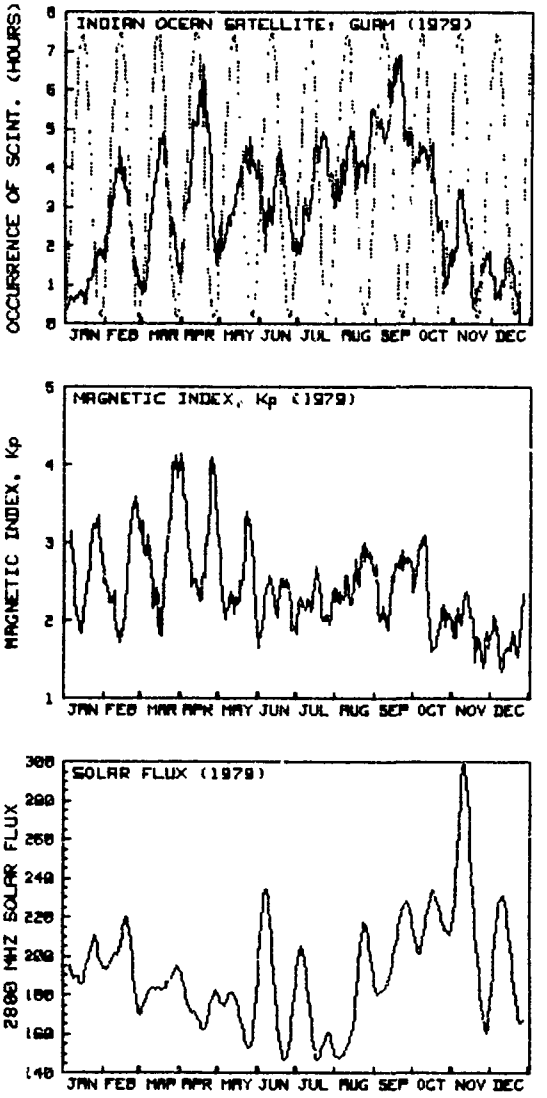


Figure 6. Eleven-day running average of the occurrence of scintillation data for the Indian Ocean satellite with the phases of the moon superimposed. Middle graph is a running average of the daily average 3-hour magnetic Kp index. Bottom graph is an 11-day running average of the 2800-MHz solar flux.

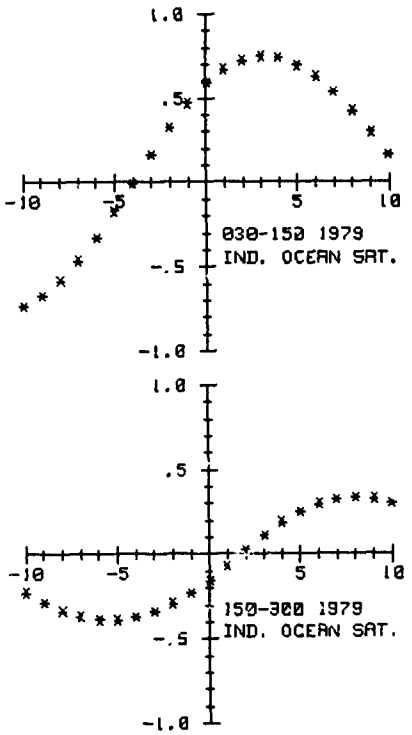


Figure 7. Cross correlations between the phases of the moon and the 11-day running averages for two data samples for the Indian Ocean satellite. The abscissa represents days before and after full moon. Correlation maxima occur after full moon.

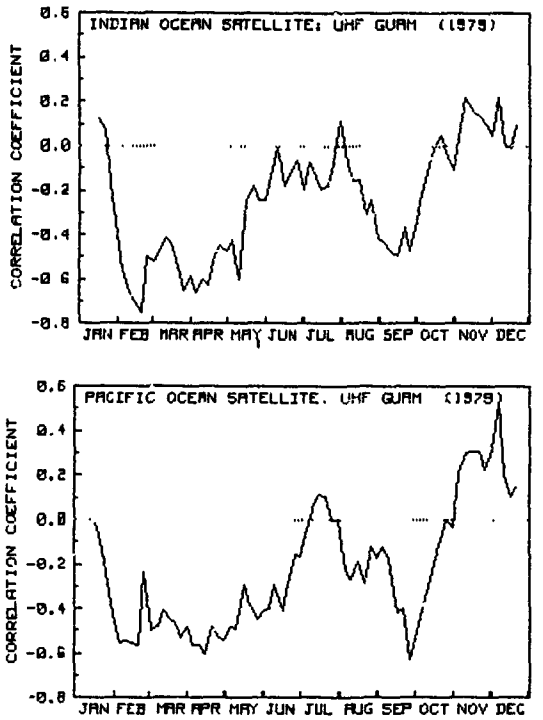


Figure 8. Cross correlations between total hours per night that scintillation fading exceeded 6 dB and the daily summed 3-hour Kp indices. Thirty-day data samples were used in the calculations.

C-BAND IONOSPHERIC SCINTILLATION MEASUREMENTS AT HONG KONG EARTH STATION DURING THE PEAK OF SOLAR ACTIVITIES IN SUNSPOT CYCLE 21*

D. J. Fang
COMSAT Laboratories
Clarksburg, Maryland 20734

ABSTRACT

For approximately 10 years, COMSAT Laboratories has employed INTELSAT earth stations around the world to conduct C-band ionospheric scintillation measurements. Early data established the existence of 4- and 6-GHz ionospheric scintillations and gross features such as dependence on local sunset conditions, diurnal and seasonal occurrence patterns, geomagnetic boundaries, and the order of magnitude of the frequency dependence of the scintillation phenomena. For weak scintillation events, power spectral density analyses confirmed an f^{-3} dependence, indicating that the conventional weak scattering theory for a thick screen with a power law electron density fluctuation spectrum of p = 4 would adequately explain the weak gigahertz ionospheric scintillation. Data collected from 1979 to 1980 at the Hong Kong Earth Station provide further insight into the nature of 4- and 6-GHz scintillations during a year of maximum solar activity. Some characteristics such as occurrence frequencies, durations, global scales, and spectral roll-off slopes are unique in that they have not been observed in previous years.

INTRODUCTION

Ionospheric scintillations are microwave signal fluctuations due to the existence of ionospheric irregularities of refractive indices [1-11]. When first occurring, intense signal fluctuations often erupt suddenly without a noticeable precursor. Once started, the scintillations can last for hours with intermittent changes in the magnitude of peak-to-peak fluctuations as well as in the rate of fluctuations.

COMSAT Laboratories has been using INTELSAT earth stations around the world and INTELSAT satellites in three oceanic regions to measure 4/6-GHz ionospheric scintillations for the past 10 years [8-11]. Although the INTELSAT commercial communications links are not designed for microwave research, the basic advantage of using them for microwave propagation experiments is that results such as cumulative statistics can be readily applied to estimate the performance of the communications system, which is the main goal of the measurements.

The gross features of 4/6-GHz ionospheric scintillations can be summarized as follows:

a. Scintillations occur in the geomagnetic equatorial region. While there are no clearly observable boundaries, it is assumed that boundaries are approximately 30° GMN and 30° GMS. They expand and contract as solar activities increase and decrease, respectively.

b. The frequency of occurrence of scintillation events has strong diurnal peaks. The probability that events will occur is greatest about 1 hour after local ionospheric sunset. Scintillations may last for hours, until midnight.

*This paper is based upon work performed at COMSAT Laboratories under the sponsorship of the International Telecommunications Satellite Organization (INTELSAT).

c. The frequency of occurrence exhibits seasonal variations. The peak activity period is around vernal equinox, and high activity is observed at autumnal equinox time.

d. An f^{-s} relationship with s between 1.5 to 2.0 appears to exist between the amplitude of scintillation events measured at 4 and 6 GHz.

e. The power spectral densities of the scintillation signals generally exhibit a power law frequency dependence for spectral frequencies greater than the Fresnel frequency. An f^{-3} asymptotic frequency dependence can be considered as a reasonable approximation for most weak scintillation events.

All of these features have annual variations, which are related to the 11-year sunspot cycles.

This paper presents the analysis results of the most recent ionospheric scintillation data collected at the Hong Kong Earth Station. The measurement period from March 1979 to March 1980 coincides with the solar maximum period of the current sunspot cycle 21 as shown in Figure 1. As a result, enormous amounts of high-quality scintillation data were collected. The data not only further substantiate the gross features of ionospheric scintillations, but also provide physical insight for refined characterizations of the phenomena.

EXPERIMENT CONFIGURATIONS

Two geostationary satellites, three earth stations, and six up/down-link signals are involved in the experiment. The two satellites, INTELSAT IV F8 Pacific Ocean Regional (POR) and INTELSAT IV-A F6 Indian Ocean Regional (IOR), are located at 174°E and 63°E, respectively. The three earth stations are Stanley, Hong Kong; Sentosa, Singapore; and Si Racha, Thailand. Detailed information is given in Table 1. Only the Hong Kong Earth Station (HK) was used for data collection. The separation between the IOR-HK and POR-HK links in the ionosphere at an altitude of 300 km is about 2,000 km, or about 1 hour geographical local time. The six signals consist of two beacons and four communications carriers, all being monitored at the Hong Kong Earth Station, as illustrated in Table 2. The lowest transmission path, from Si Racha to POR, has an elevation angle of approximately 8°. All elevation angles for other paths are over 10°, which is well above the angle (~5°) below which tropospheric and multipath effects may significantly degrade microwave signals [12-16].

Table 1. Earth Station Information

Earth Station	Geographic Location		Geomagnetic Latitude	Antenna (Satellite)	Elevation Angle (deg)
	Longitude	Latitude			
Stanley, Hong Kong	114° 13' 11"E	22° 11' 57"N	10.52°N	HK1 (POR) HK2 (IOR)	19.59 27.76
Sentosa, Singapore	103° 50' 12"E	01° 14' 50"N	10.36°S	SN1 (IOR) SN2 (POR)	42.78 11.31
Si Racha, Thailand	100° 56' 11"E	13° 06' 02"N	1.57°N	SR1 (POR) SR2 (IOR)	7.87 43.94

Reference 11 explains how to employ an INTELSAT earth station to collect ionospheric scintillation data. After more than 10 years of continuous measurements at selective INTELSAT earth stations, the method of monitoring carriers and data analysis procedures for ionospheric scintillation studies has been standardized. For each carrier or beacon, a buffered output proportional to the automatic gain control (AGC) voltage is made available on the respective IF amplifier. This output is fed into the recorder strip chart and FM magnetic tape recorder for data collection. The input to each IF amplifier has an attenuator inserted, which under normal communications conditions, is set to 5 dB; the amplifier gain is then adjusted to accommodate the attenuation. The data can thus be calibrated by changing the attenuation in 1-dB steps to provide a ± 5 dB range for a scintillation experiment. Accuracy of scintillation signal detection is within ± 0.5 dB. This range of accuracy is mainly due to the fact that even under clear sky conditions, the buffered IF output has small variations caused by noise.

Table 2. Six Signals Being Monitored at Hong Kong Earth Station for Ionospheric Scintillation Studies

Path	Frequency (MHz)		Nature
	Up-Link	Down-Link	
POR → HK1	None	3950 ± 2.5	POR Beacon
SN2 → POR → HK1	5977.5	3752.2	Communications Carrier
SR1 → POR → HK1	6227.5	4002.5	Communications Carrier
IOR → HK2	None	3950 ± 2.5	IOR Beacon
SN1 → IOR → HK2	6190.0	3965.0	Communications Carrier
SR2 → IOR → HK1	6103.75	3878.75	Communications Carrier

Data analyses for strip chart data follow the 15-minute P_{max} and P_{min} method devised by Whitney, Aarons, and Malik [17]. The results, such as cumulative scintillation statistics for a year as well as for a worst month, are essential for system applications. Magnetic tape data, however, are analyzed for detailed study of severe scintillation events largely for research interest. This paper reports mainly on magnetic tape data.

GENERAL PATTERNS OF SIGNAL FLUCTUATIONS OBSERVED FROM AN EARTH STATION ANTENNA RECEIVE SYSTEM

To identify the signature of ionospheric scintillation, first, it is necessary to recognize the general patterns of signal fluctuations at 4 GHz as observed from a standard INTELSAT 30-m antenna receive system. These patterns are illustrated in Figure 2, which shows six channels of strip chart records, including the following:

a. POR satellite beacon level, azimuthal tracking error voltage, and elevation tracking error voltage, all received at Taipei 1 Earth Station; and

b. IOR satellite beacon level, azimuthal tracking error voltage, and elevation tracking error voltage, all received at Taipei 2 Earth Station.

Figures 2a and 2b are typical patterns of ionospheric scintillations observed at the Taipei Earth Station during 1977 [11]. Mild (less than 4 dB), but clearly noticeable, scintillations occurred mostly along the IOR link. The events started suddenly, and peak-to-peak signal fluctuation reached its maximum value in a couple of minutes. The main scintillations lasted less than 2 hours, but mild fluctuations did not taper off until past local midnight.

Figure 3 shows the spectra for the six 10-minute records given in Figure 2a. Spectra A and F (before and after the intense signal fluctuation period, respectively) are typical of those produced by weak scintillations. The mean square fluctuations of amplitude increase with frequency at the low frequency end until they reach the maximum near the Fresnel frequency. Beyond this point, the spectra appear to roll off with the characteristics of a Gaussian form. However, during intense signal fluctuations (B, C, D, and E), the characteristics of roll-off change from Gaussian to a power law form. For the four records (B, C, D, and E), the roll-off slope remains approximately the same, yielding a frequency dependence of f^{-3}. The power spectral density for the scintillation event shown in Figure 2b is similar to the one presented in Figure 3; in particular, the same f^{-3} frequency dependence is noted.

Figure 2c is a pattern of signal fluctuations due to atmospheric turbulence and local wind. Strong signal fluctuations were observable from beacon and tracking channels at both IOR and POR antenna receive systems. The fluctuations are caused not only by effects such as turbulent refraction, lower atmospheric gradient, and local terrain multipath, but also by the angle-of-arrival change resulting from the dynamic response of the earth station automatic tracking systems [12-15]. For

many earth stations, such as Taipei and Hong Kong, those fluctuations are predominantly early morning or early afternoon events.

Figure 2d is typical for earth stations under precipitations. Electromagnetic waves are attenuated by absorption of water vapor and raindrops. The attenuation is not apparent in the figure because at 4 GHz the value of attenuation coefficients is very small [18,19]. Signal fluctuation is evident in beacon channels but not in tracking channels because under absorption conditions, no significant amount of phase or wave-front distortion will occur. The beacon fluctuations are attributed to the volatile nature of water vapors and raindrops in terms of particle sizes, shapes, drop-size distributions, terminal velocities, etc. [19].

Fluctuations of tropospheric origins shown in Figures 2c and 2d can be further differentiated from the ionospheric scintillations in Figures 2a and 2b by studying the power spectral densities in terms of Fresnel frequencies, roll-off slopes, etc. The turbulence structures in the lower atmosphere have scale sizes, elongations, and refractive index correlation spectra which are distinctively different from those in the ionosphere [5,7,11,14,16,20-24].

OUTSTANDINC FEATURES OF IONOSPHERIC SCINTILLATIONS IN THE SOLAR MAXIMUM YEAR

During the 12-month measurement period, ionospheric scintillations of over 1.0-dB peak-to-peak fluctuations at either IOR or POR links were observed at the Hong Kong Earth Station for about 100 evenings. In March 1979, scintillations occurred almost every evening, as shown in the monthly glance picture in Figure 4. The frequency of occurrence reduced to about 20 evenings in April 1979 and to 10 evenings in September 1979. This generally confirms the theory that scintillations have seasonal variations with activity peaks at vernal and autumnal equinoxes.

Except for the evening of March 21 when only the IOR link scintillated, the monthly glance picture clearly suggests that for most cases, scintillations were events on both links; at least in one link, scintillations lasted 3 to 5 hours. This behavior was not observed in the measurements made a few years earlier when sunspot numbers were much lower. For example, in Figures 2a and 2b, only the IOR link indicated signal fluctuations for about 2 hours, although residual fluctuations remained. Obviously, ionospheric disturbances that are responsible for gigahertz scintillations are more intense in strength, larger in spatial scale, faster in motion, and longer lasting in time scale in the solar maximum years than those encountered during years of low to medium solar activities. A similar conclusion has been reached by physicists studying various types of ionospheric disturbances, some of which are related to giganertz scintillations, such as equatorial spread-F and gravity waves [25-30].

Figures 5, 6, and 7 provide records of three ionospheric scintillation events in March 1979. In all three figures, fluctuation patterns of the two signals received from the IOR satellite are well correlated, irrespective of the up-link paths (whether the signals are from Singapore or from Thailand). This is also true for the fluctuation patterns of the two signals received from the POR satellite. This correlation pattern has proved to be valid mainly for data collected over the entire year, suggesting that ionospheric scintillations, as observed from an earth station, are predominantly down-link phenomena.

Figure 5 shows that the IOR link has scintillated before th POR link, which is a rare pattern for the year of maximum solar activities. Before 2200 local time, significant scintillation occurred in the IOR link, while the POR link had only low-level fluctuations. Furthermore, the IOR scintillation started suddenly without any precursor. All of these features are rather typical for years of low solar activities, as illustrated in Figures 2a and 2b. The IOR scintillation event lasted almost an hour and then faded quickly. A separate and obviously uncorrelated scintillation event occurred along the POR link at about 2200. The signal fluctuation enhanced itself gradually until it reached to a magnitude of about 8 dB peak-to-peak. The event lasted slightly more than an hour, and was followed by another minor event with peak-to-peak fluctuations of 3 dB.

Figures 6 and 7 are representatives of majority events with the following typical characteristics:

a. Scintillations occurred on both links. Although the onset times on the two links were not simultaneous, the scintillations overlapped for a significant amount of time, from 20 minutes to 2 hours. The fact the two links are separated at an F-max height (300 km) of about 2,000 km suggests that ionospheric disturbances which created irregularities for causing ionospheric scintillations may have scales of the order of at least 2,000 km.

b. Most scintillations did not erupt suddenly, but with a noticeable gradual increment of fluctuations as precursors. Frequently, the scintillations were intermittent fluctuation events for an entire evening; they were not confined to 2 to 3 hours as those observed in previous years. Individual fluctuation events lasting more than 2 hours along at least one link were common.

c. Due to the presence of a precursor, it is difficult to identify exactly the onset time of a scintillation event. However, the figures clearly show that the POR link scintillated much earlier than the IOR link and that the time delays of scintillations between the two links were on the order of 30 minutes to 2 hours. This is an indication that gigahertz scintillation is an ionospheric sunset phenomenon, and the time delays are associated with the ionospheric sunset times at altitudes of 200 to 400 km where F-region irregularities were known to be preset under disturbed ionospheric conditions.

d. A rather extensive effort was made to correlate the scintillation patterns between the two links with a proper shift of time scale to accommodate the time delays. In addition, the scintillation index, SI, given by the equation [17]

$$SI = \frac{P_{max} - P_{min}}{P_{max} + P_{min}} \times 100 \quad \text{(percent)}$$

was also plotted for the two links for comparison. Results indicated that detailed correlations between the links did not exist. Obviously, the scintillations along the two links were independent events. Recent observations [31,32] have suggested that while the disturbances may have a spatial scale as large as 2,000 km, a time scale of several hours, and travel with the ionosphere from east to west in line with local sunset, the specific irregularities responsible for gigahertz ionospheric scintillations created by the disturbances nevertheless exhibit much smaller spatial scales and are relatively short-lived.

Further details on ionospheric scintillations can be obtained by examining the signal fluctuation patterns in frequency domain. Since data collected at Hong Kong were AGC voltages corresponding to the intensities of the electromagnetic waves incident on the antenna, only power spectral density of the wave intensities can be analyzed.

Power spectral density curves for the event which occurred in the evening of March 18-19, 1979, are plotted in Figure 8 for three separate 10-minute segments of data starting at 2100, 2300, and 0000 local time. As expected, the mean square fluctuation first increases with frequency until it reaches maximum at the Fresnel frequency. Afterwards, it starts to roll off with a power law slope. An important feature to be noted is that the roll-off slope, which always follow a power law, is no longer f^{-n} with n approximately equal to 3 as concluded in scintillation studies for previous years [3-5,9,11,33]. That is, the slope and the Fresnel frequency change significantly as a function of time, just as the peak-to-peak fluctuation does. To illustrate this point, the Fresnel frequencies and roll-off slopes for the three representative events given earlier are tabulated in Table 3.

The difference in characteristics between the scintillations on the POR link and those on the IOR link can be clearly seen from Table 3. The IOR link, which is the westward link that in the evening penetrates the ionosphere in the direction of sun, has Fresnel frequencies generally less than 0.100 Hz and an f^{-n} roll-off with $n < 3.0$. On the other hand, along the eastward POR link that penetrates the dark ionosphere in the evening, a spectral broadening effect becomes evident with Fresnel frequencies consistently greater than 0.10 Hz. Furthermore, the roll-off slopes are signficantly steeper when the value n exceeds 3.0 all of the time. Ionospheric irregularities along the IOR and POR links are obviously not identical. They differ at least in physical parameters such as height, size, moving velocities, and random refractive index spectra which dictate the Fresnel

Table 3. Fresnel Frequency and Roll-Off Slope of Scintillation Events Given in Figures 5, 6, and 7

Date and Time	Up-link Carrier	POR + HK				IOR + HK			
		P-P Fluctuation (dB)	Scintillation Index (%)	Fresnel Frequency (Hz)	f^{-n} Roll-Off n	P-P Fluctuation (dB)	Scintillation Index (%)	Fresnel Frequency (Hz)	f^{-n} Roll-Off n
March 3, 1979									
2105	SN	1.5	18	0.200	3.2	4.0	43	0.100	2.8
	SR	0.6	-	-	-	2.2	25	0.120	2.9
2245	SN	6.0	60	0.150	3.8	0.3	-	-	-
	SR	6.5	63	0.180	3.9	0.3	-	-	-
March 18-19, 1979									
2015	SN	7.0	66	0.160	3.5	0.3	-	-	-
	SR	7.0	66	0.180	3.7	0.3	-	-	-
2100	SN	8.0	72	0.190	3.7	3.0	33	0.090	2.3
	SR	9.0	77	0.180	3.7	3.0	33	0.15	2.9
2200	SN	1.8	20	0.160	3.7	3.0	33	0.065	2.5
	SR	2.0	22	0.155	3.7	1.8	20	0.074	3.3
2300	SN	1.2	14	0.115	4.0	1.2	14	0.032	2.0
	SR	5.0	52	0.115	3.7	1.2	14	0.230	3.0
0000	SN	4.0	42	0.070	3.2	4.0	43	0.052	2.6
	SR	1.5	18	0.072	3.4	2.5	28	0.058	3.0
0100	SN	1.8	20	0.105	4.5	3.0	33	0.058	3.6
	SR	1.5	18	0.120	3.7	2.5	28	0.060	3.0
0200	SN	1.0	11	0.115	4.5	0.3	-	-	-
	SR	1.5	18	0.100	3.7	0.3	-	-	-
March 20, 1979									
2020	SN	6.5	63	0.160	3.5	0.3	-	-	-
	SR	6.5	63	0.175	3.5	0.3	-	-	-
2300	SN	1.5	18	0.118	4.6	6.0	60	0.084	3.4
	SR	1.8	21	0.128	3.8	4.5	48	0.088	3.6

frequencies and roll-off slopes. These parameters will be analyzed in detail in a subsequent paper.

CONCLUSIONS

The ionospheric scintillation measurement at Hong Kong from 1978 to 1979 coincided with the solar maximum period of the current solar cycle 21. During the 1-year measurement period, scintillations were observed for approximately 100 evenings, which is a much longer time than that recorded in previous years. The amount of meaningful data has provided a new view of the phenomena. Preliminary analyses of data, as reported here, revealed that many characteristics of gigahertz ionospheric scintillations are rather unique when compared with the results derived from similar measurements made in previous years when solar activities were relatively low. Furthermore, since measurement was made by observing signals from the POR and IOR satellite simultaneously, a rare opportunity is provided to examine ionospheric scintillation from a global point of view. Power spectral studies reveal that scintillations in evening hours as observed from the eastward link pointing to the dark differ in detail from those observed from the westward link pointing to the sun.

REFERENCES

[1] J. Aarons, H. E. Whitney, and R. S. Allen, "Global Morphology of Ionospheric Scintillations," Proc. IEEE, Vol. 59, 1971, pp. 159-172.

[2] E. J. Fremouw and C. L. Rino, "An Empirical Model for Average F-Layer Scintillation at VHF/UHF," Radio Science, Vol. 8, 1973, p. 213.

[3] B. H. Briggs and I. A. Parkins, "On the Variation of Radio Star and Satellite Scintillations with Zenith Angle," Journal of Atmospheric and Terrestrial Physics, Vol. 25, 1963, p. 339.

[4] C. L. Rufenach, "Ionospheric Scintillation by a Random Phase Screen-Spatial Approach," Radio Science, Vol. 10, 1975, pp. 155-165.

[5] R. K. Crane, "Spectra of Ionospheric Scintillation," Journal of Geophysical Research, Vol. 81, 1976, pp. 2041-2050.

[6] S. Basu et al., "Preliminary Comparison of VHF Maps of F Region Irregularities with Scintillations in the Equatorial Region," Journal of Atmospheric and Terrestrial Physics, Vol. 39, 1977, pp. 1251-1261.

[7] E. J. Fremoun, R. C. Livingston, and D. A. Miller, "On the Statistics of Scintillating Signals," Journal of Atmospheric and Terrestrial Physics, Vol. 42, 1980, pp. 717-731.

[8] R. R. Taur, "Ionospheric Scintillation at Frequencies Above 1 GHz," COMSAT Technical Review, Vol. 4, 1974, pp. 461-476.

[9] R. R. Taur, "Simultaneous 1.5-and 4-GHz Ionospheric Scintillation Measurements," Radio Science, Vol. 11, 1976, pp. 1029-1036.

[10] D. J. Fang, D. J. Kennedy, and C. Devieux, "Ionospheric Scintillations at 4/6 GHz and Their System Impact," EASCON-78 paper, IEEE Pub. No. CH1352-4/78/C000-0385, 1978.

[11] D. J. Fang, "4/6 GHz Ionospheric Scintillation Measurements, AGARD Conference Proceedings No. 284, Propagation Effects in Space/Earth Paths, 1980, pp. 33-1 to 33-12.

[12] P. Beckmann and A. Spizzichino, The Scattering of Electromagnetic Waves from Rough Surfaces, New York: MacMillan, 1963.

[13] D. J. Fang, "Scattering from a Perfectly Conducting Sinusoidal Surface," IEEE Transactions on Antennas and Propagation, Vol. AP-20, 1972, pp. 388-390.

[14] R. K. Crane, "Low Elevation Angle Measurement Limitations Imposed by the Troposphere," Lincoln Laboratory Report ESD-TR-76-115, May 1976.

[15] D. J. Fang and J. M. Harris, "Precipitation Attenuation Studies Based on Measurements on ATS-6 20/30 GHz Beacon Signals at Clarksburg, Maryland," IEEE Transactions on Antennas and Propagation, Vol. AP-27, January 1979, pp. 1-11.

[16] D. J. Fang, F. T. Tseng, and T. O. Calvit, "A Low Elevation Angle Propagation Measurement of 1.5 GHz Satellite Signals in the Gulf of Mexico," Accepted for publication, IEEE Transactions on Antennas and Propagation, 1981.

[17] H. E. Whitney, J. Aarons, and C. Malik, "A Proposed Index for Measuring Ionospheric Scintillations," Planetary Space Science, Vol. 17, 1969, pp. 1069-1073.

[18] D. J. Fang and J. Jih, "A Model for Microwave Propagation Along an Earth-Satellite Path," COMSAT Technical Review, Vol. 6., No. 2, Fall 1976, pp. 379-411.

[19] D. J. Fang, "The Effects of Raindrop Size Distributions on Satellite Communications Systems: A New Assessment," to be published, draft copy available upon request, 1981.

[20] V. I. Tatarskii, "The Effects of the Turbulent Atmosphere on Wave Propagation," IPST Catalog 5319, NTIS, Springfield, Virginia, 1971.

[21] G. R. Ochs, S. F. Clifford, and T. I. Wang, "Laser Wind Sensing: The Effects of Saturation of Scintillation," Applied Optics, Vol. 15, 1976, pp. 403-408.

[22] E. J. Fremoun and D. J. Fang, "An All-Sky-Camara Format for Radar Auroral Studies," Radio Science, Vol. 10, 1975, pp. 891-904.

[23] D. J. Fang, "The Formation of High-Latitude Plasma Irregularities by Wind-Shear Action in the Upper Atmosphere," Journal of the Atmospheric Sciences, Vol. 22, 1970, pp. 316-321.

[24] R. K. Crane, "Refractive Effects in the Nature Atmosphere," Methods of Experimental Physics, Vol. 13B, Section 2.5, M. L. Meeks, ed., New York: Academic Press, 1976.

[25] J. Aarons, J. Buchau, S. Basu, and J. P. McClure, "The Localized Origin of Equatorial F Region Irregularity Patches, Journal of Geophysical Research, 83, 1978, pp. 1659-1664.

[26] H. G. Booker, "The Role of Acoustic Gravity Waves in the Generation of Spread-F and Ionospheric Scintillation," paper presented at URSI General Assembly, Helsinki, August 1-8, 1978.

[27] J. D. Huba, P. K. Chaturvedi, and S. L. Ossakow, "High Frequency Drift Waves with Wavelengths Below the Ion Gyroradius in Equatorial Spread-F," Geophysical Research Letters, 5, 1978, pp. 695-698.

[28] M. K. Hudson, "Spread-F Bubbles: Nonlinear Rayleigh-Taylor Mode in Two Dimension, Journal of Geophysical Research, 83, 1978, pp. 3189-3194.

[29] J. Rottger, "The Macroscale Structure of Equatorial Irregularities, Journal of Atmospheric and Terrestrial Physics, 38, 1976, pp. 97-102.

[30] J. D. Whitehead, "Ionization Disturbances Caused by Gravity Waves in the Presence of an Electrostatic Field and Background Wind," Journal of Geophysical Research, 1971, pp. 238-241.

[31] J. Aarons, J. P. Mullen, and H. E. Whitney, "The Dynamics of Equatorial Irregularity Patch Formation, Motion and Decay," Journal of Geophysical Research, Vol. 85, 1980, pp. 139-149.

[32] J. Aarons, H. E. Whitney, and Eileen MacKenzie, "Microwave Equatorial Scintillation Intensity During the Current Solar Maximum," to be published, Radio Science, 1981.

[33] D. L. Dyson, J. P. McClure, and W. B. Hanson, "In Situ Measurements of the Spectral Characteristics of F Region Ionospheric Irregularities," Journal of Geophysical Research, Vol. 79, 1974, pp. 1497-1502.

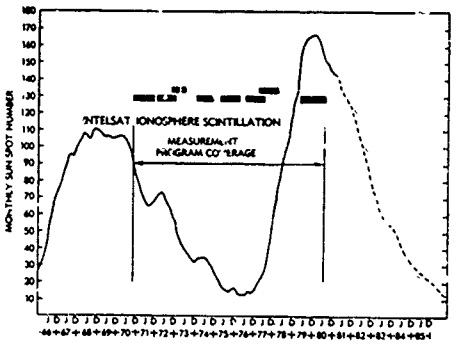


Figure 1. Monthly Sunspot Numbers From 1966 to 1985 and INTELSAT Ionospheric Scintillation Measurement Program Period From 1970 to Present

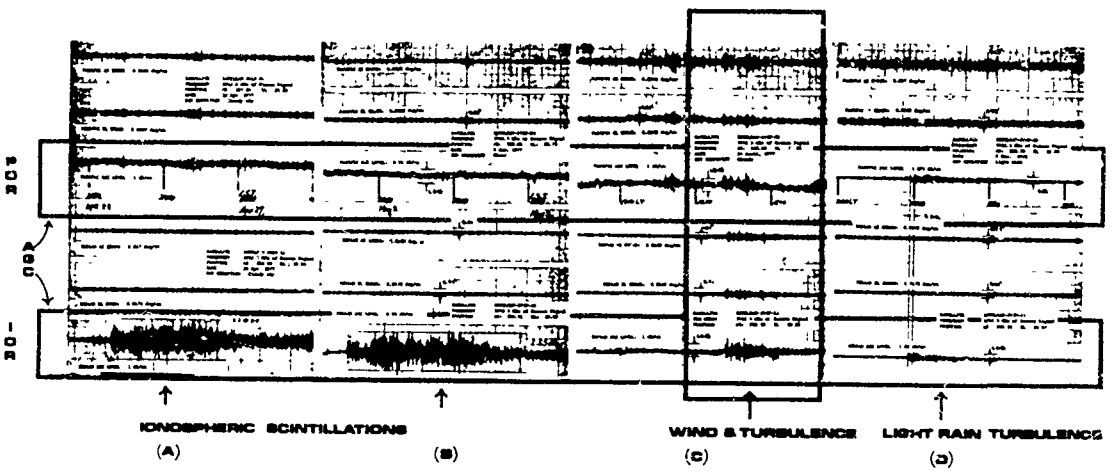


Figure 2. General Patterns of Signal Fluctuations as Observed From an Earth Station Antenna Receive System

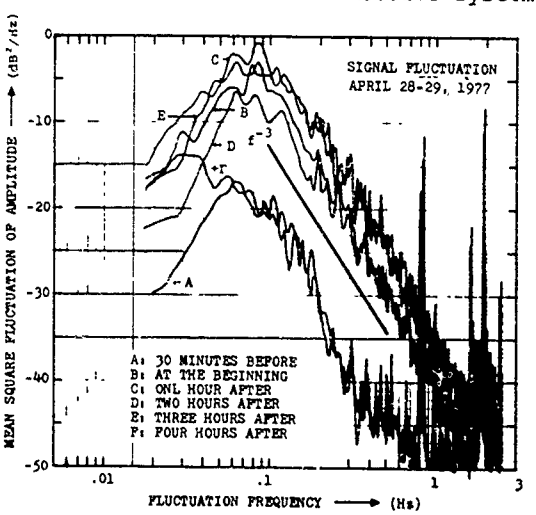


Figure 3. Power Spectral Density for Six 10-Minute Sample Periods of the Ionospheric Scintillation Event Shown in Fig

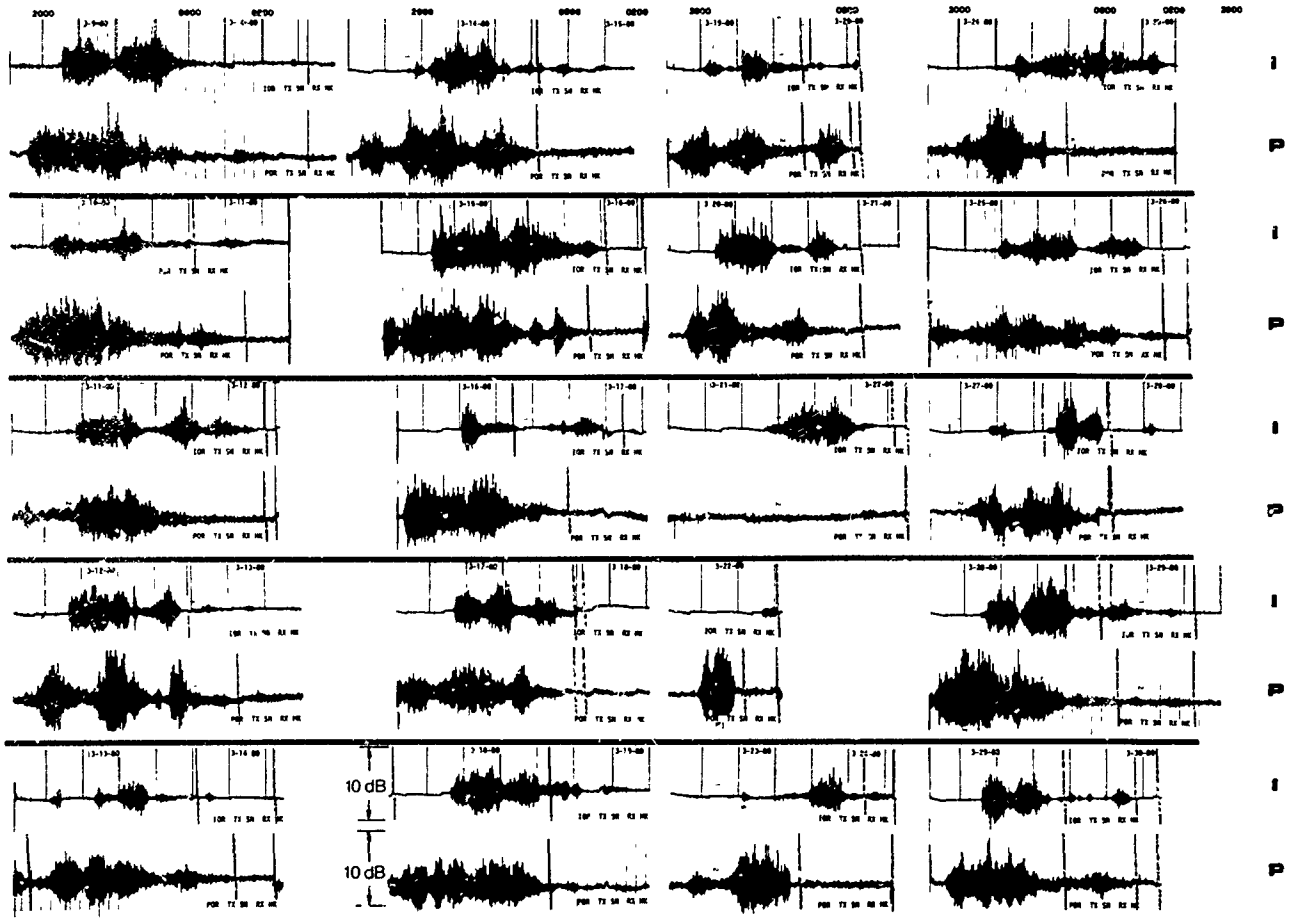


Figure 4. Monthly Glance of Ionospheric Scintillations in March 1979

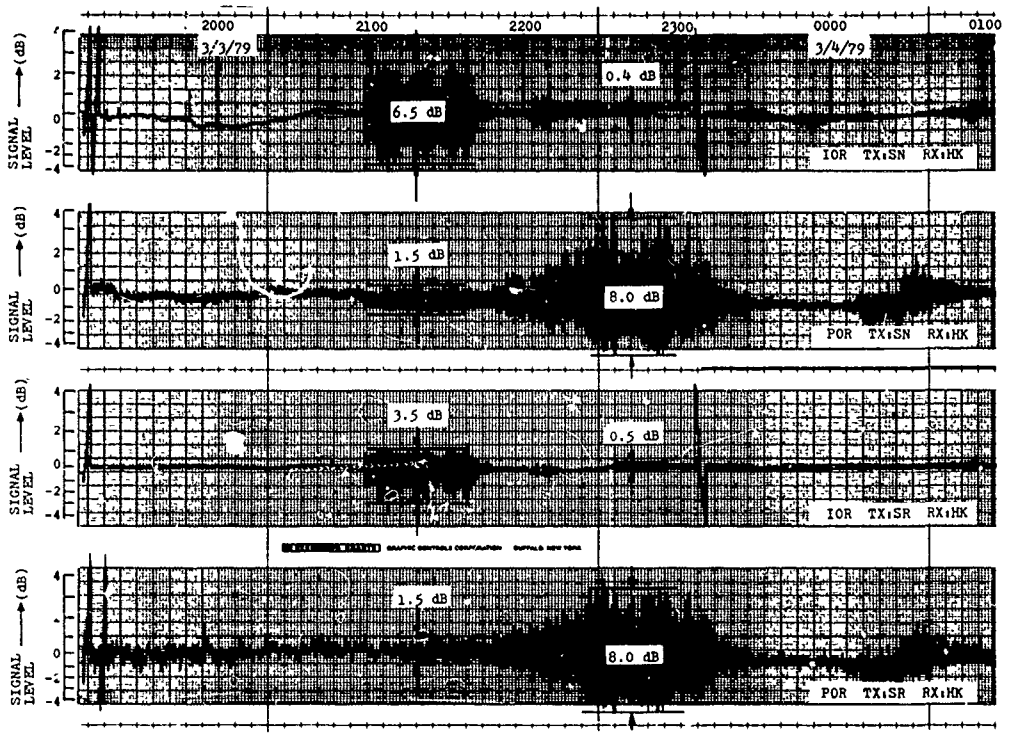


Figure 5. Records of Ionospheric Scintillation on March 3-4, 1979

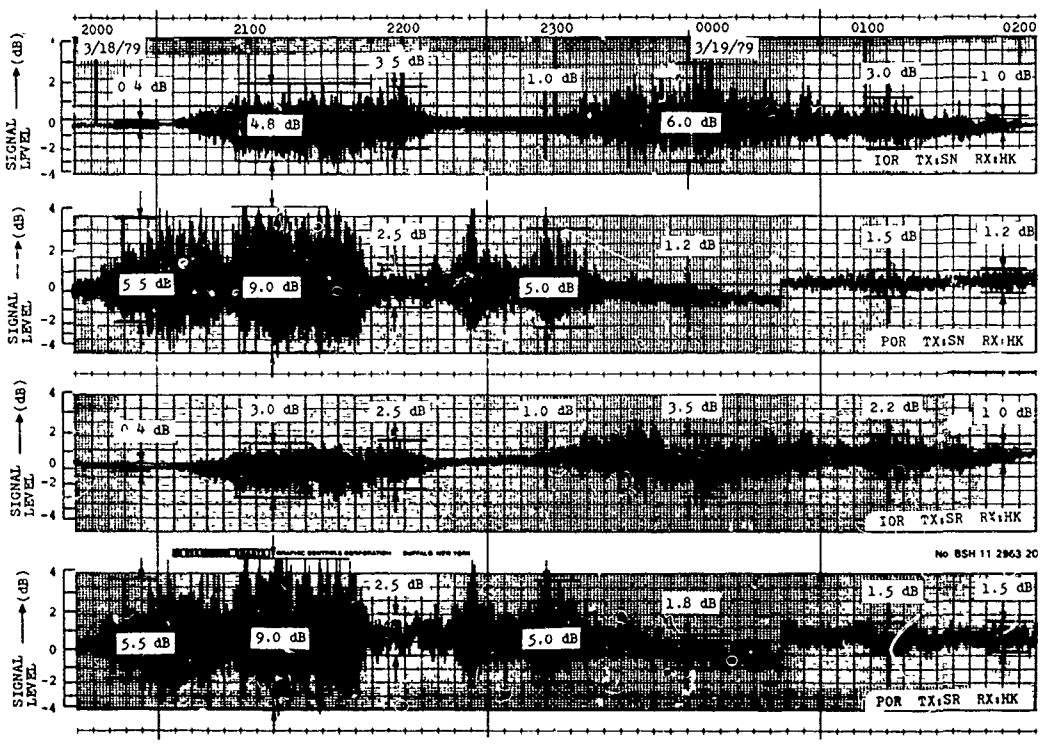


Figure 6. Records of Ionospheric Scintillation on March 18-19, 1979

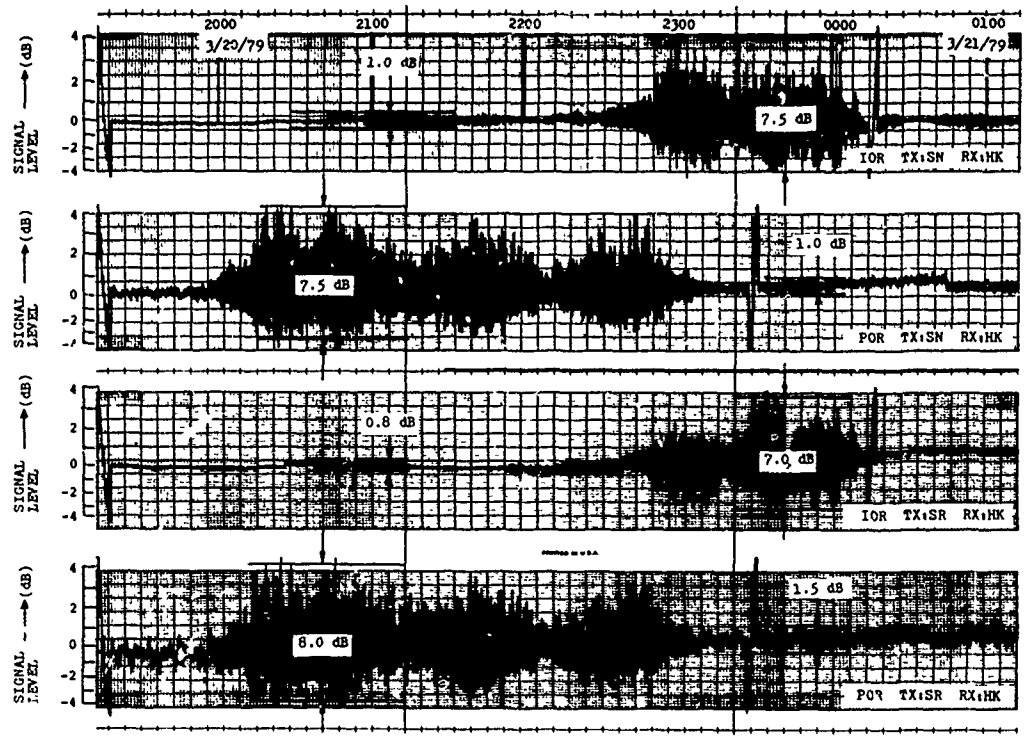


Figure 7. Records of Ionospheric Scintillation on March 20-21, 1979

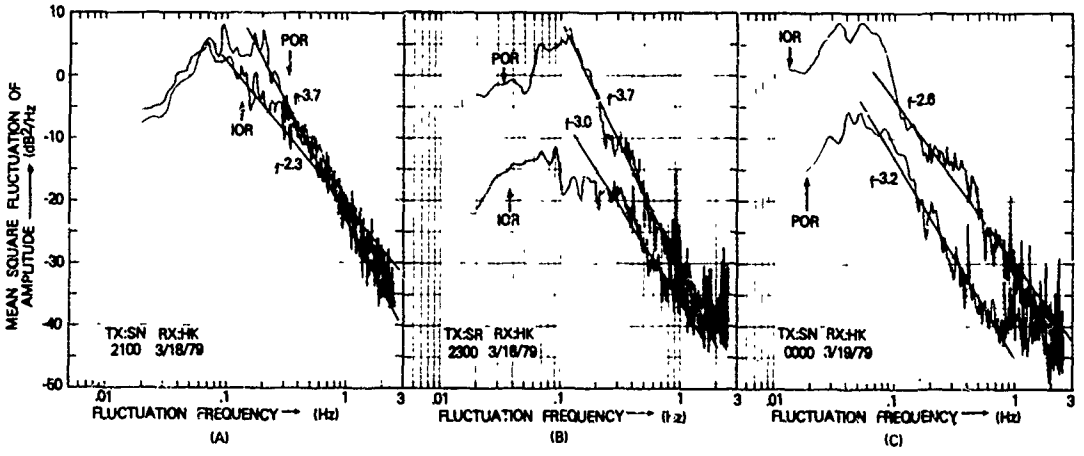


Figure 8. Power Spectral Density for the Three 10-Minute Sample Periods During the Ionospheric Scintillation Event Shown in Figure 6

MICROWAVE EQUATORIAL SCINTILLATION INTENSITY DURING SOLAR MAXIMUM

Jules Aarons
Herbert E. Whitney
Air Force Geophysics Laboratory
Hanscom Air Force Base, Bedford, MA 01731

Eileen MacKenzie
Santimay Basu
Emmanuel College
Boston, MA 02115

ABSTRACT

A comparison of scintillation levels at 1.5 GHz made from the Appleton anomaly region of the magnetic equator and from the region close to the magnetic equator (termed the electrojet latitudes) showed increased F region irregularity intensity over the anomaly region during years of high sunspot number. Peak to peak fading greater than 27 dB was noted from Ascension Island (through a dip latitude of 17°) in the anomaly region while only 7-9 dB from Natal, Brazil and Huancayo, Peru were noted, the last two paths close to the magnetic equator.

The hypothesis advanced is that the dominant factor responsible for the intense gigahertz scintillation is the traversal of the propagation path through the anomaly region. During years of high sunspot numbers, the high levels of ΔN constituting the F region irregularity structure are due to (1) very high electron density in the anomaly region (compared to the electrojet region) and (2) the late appearance of these high electron densities (to 2200 Local Time) in the anomaly region. The patches or plumes of irregularities seen in the post sunset time period then produce high ΔN; scintillation excursions are proportional to this parameter.

The postulation of vertical irregularity sheets in the patches was examined to determine the possibility of this being an important factor in the difference between electrojet and anomaly scintillation levels. Older gigahertz data from the sunspot maximum years of 1969-1970 were reanalyzed and more recent observations from other studies were also reviewed. It was found that through the anomaly region, high scintillation indices were noted at a variety of azimuths of the propagation path rather than just along a path closely aligned with the magnetic meridian. A more complete evaluation of the geometrical factor, which must be of considerable importance in determining the absolute value of the scintillation intensity, awaits further observations.

INTRODUCTION

In recent years a variety of methods has been used to probe and measure F layer irregularities in the equatorial region (Aarons, 1977; Basu and Kelley, 1979). One method is to record transionospheric scintillation of satellite signals in the VHF to microwave frequency range.

This paper will contrast levels of scintillation activity at 1541 MHz which were relatively low when a satellite beacon was observed through latitudes close to the magnetic equator (Natal, Brazil and Huancayo, Peru) and very high when the same satellite beacon was observed nearly overhead at Ascension Island through a region near the peak of the Appleton equatorial anomaly during the peak of solar activity in 1979 and 1980. The equatorial or Appleton anomaly shows during noontime a distinct trough of electron density in the bottomside and topside ionosphere at the magnetic dip equator with crests of ionization at ~15°-20° north and south dip latitudes. We shall discuss later that the F region ionization near Ascension Island exhibits a latitudinal variation in the post-sunset hours during the solar maximum period similar to the daytime Appleton anomaly.

MEASUREMENTS

AFGL, in cooperation with various groups, records scintillation activity of the MARISAT beacon, transmitting at UHF (257 MHz) and at L band (1541 MHz) at three sites near the magnetic equator. The sites are listed in Table 1 with the coordinates of the 400 km sub-ionospheric intersection points, the dip latitude of this point and the propagation angle (i.e., the angle between the path to the satellite and the lines of force of the earth's magnetic field) and the ground elevation and azimuth angles of the path to the satellite. Figure 1 is a map with geographic latitudes and dip latitudes plotted.

MAGNETIC EQUATOR OBSERVATIONS

In a recent paper, Basu et al. (1980) have analyzed the L band observations made from Huancayo, Peru over a period of 3 years. During that time no fluctuation greater than 8 dB peak-to-peak was recorded. In the month of December 1979, a period of intensive and common measurements at the three sites, scintillation activity maximum was 6.8 dB at Huancayo. Viewing of MARISAT from this site was done almost perpendicular to the field lines, i.e., at an azimuth of 83° thru a dip latitude of 2.5°N as shown in Table 1.

In cooperation with the Federal University of the Northern Rio Grande similar observations have been made recently at Natal, Brazil at 65° elevation angle across the field lines, i.e., at an azimuth of 75°. Scintillations greater than 8 dB were not detected during the December period.

In December 1979, a 10 foot antenna was put into operation at Ascension Island for both L band and UHF. Over several hours on many nights some scintillations exceeded 27 dB peak to peak (Figure 2) at 1541 MHz. These fades reached the noise level of the equipment. Calibrated and consistent measurements have been made since that time which show hours of such intense activity during the December to March period. Figure 3 gives the percentage of occurrence of scintillation activity at 1541 MHz greater than 20 dB for Kp = 0-3 and Kp = 3^{+}-9 during January-February 1980. The highest percentage occurrence >20 dB for Kp = 0-3 was ∿42% with ∿23% for Kp = 3^{+}-9. It might be noted that during these months, in accord with Aaarons et al. (1980b), increased magnetic activity inhibits scintillation. Ascension Island was viewing Marisat close to the magnetic meridian plane at an elevation of 80° during these observations.

Statistics of the L-band scintillations observed at Huancayo, during February-March 1980, show a 16% occurrence of scintillation >3 dB in the time period 00-04 UT (∿19-23 LT) with no scintillation >6.8 dB even during those times. The data is clear, i.e., the scintillation activity is considerably higher for Ascension than for the two sites near the magnetic equator during the same periods of time. While the intersection longitudes of Ascension and Natal are somewhat separated, they behave similarly (Aarons et al., 1980b).

Although the data contrasts are clear, a problem exists in interpreting the observations; the Ascension Island observations were taken with the propagation path near the magnetic meridian while those at Natal and Huancayo were not. However, the Ascension Island 400 km intersection is in the anomaly region (a dip latitude of 17°) while the Natal and Huancayo intersections are close to the magnetic equator. We shall attempt to weigh the relative important of these two factors.

THE IRREGULARITY STRUCTURE AND ITS GEOMETRY

The configuration of the patch containing the small scale irregularities which has emerged from optical and radio measurements and from in-situ probing is that of an elongated structure extending several thousand kilometers north and south of the magnetic equator (Weber et al., 1980; Aarons et al., 1980a; Basu et al., 1980b).

Within the highly elongated patch, it is possible to contemplate various models for the irregularities producing the scintillations. If a 10:10:1 irregularity model of vertical sheets is assumed (10 magnetic N-S, 10 vertical in the magnetic meridian plane, and 1 unit magnetic E-W), then a pronounced maxima will result from geometrical considerations along the magnetic meridian (Sunanda Basu, private communication, 1980). This effect will enhance or minimize the amplitude scintillation value depending on the magnetic azimuth of the ray path.

Using a program developed by Mikkelsen et al. (1978), the scintillation intensity (S_4) for the propagation path from Ascension for a vertical sheet (10:10:1) is a factor of .98 relative to observations made overhead. The factor for the same postulated configuration using the geometry appropriate to Natal is .14 (an increase of S_4 ASC/S_4 NATAL of 7). Therefore the 10:10:1 sheets might account for a large portion of the ratio of Ascension to Natal levels. The geometrical situation

for Huancayo is very similar to that for Natal.

HIGH SUNSPOT NUMBER AND SCINTILLATION INTENSITY

Measurements at 137 MHz and 250 MHz have shown that the occurrence of scintillation is greater in the same season for years of high versus years of low sunspot activity (Aarons et al., 1978). However, accurate levels of strong scintillation cannot be measured at these two frequencies since both strong scattering and saturation of the signal-to-noise levels of the equipment took place frequently. Therefore, we shall confine our remarks to frequencies above 1 GHz including data at 2.2-4 GHz where signals are not saturated i.e. did not show fading down to noise. Peak to peak fading of 3 dB at 4 GHz is equivalent to fading of 18 dB at 1.5 GHz if a frequency dependence of $f^{-1.5}$ holds (Whitney, 1974).

Fang et al. (1978) states that "Gigahertz ionospheric scintillations follow the sunspot cycle. For years with sunspot number less than 30, scintillations at 4 GHz with magnitudes over 2 dB were never observed at INTELSAT earth stations." For the purposes of relating various measurements, we have plotted sunspot number in Figure 4 for recent years.

For the sake of outlining our hypothesis we have divided the equatorial region into "electrojet" latitudes and "anomaly" latitudes. Electrojet latitudes are taken to be in the range of 5°N - 5°S magnetic dip latitude, while anomaly latitudes are on either side of the electrojet latitudes, falling off beyond 15°-20° dip latitude. Anomaly peaks may be relatively narrow in latitude. We postulate that the differences in scintillation intensity are related to the following:

1. The equatorial anomaly has higher electron density values in high sunspot number years than in years of low solar activity,
2. The occurrence of maximum electron density for anomaly latitudes is near sunset in the years of high sunspot number, and in the afternoon in years of low solar activity.

To illustrate these two points we have reproduced diagrams from Rawer (1963) for MUF (3000) F_2 in the Europe-African zone for the month of December for years of low sunspot number (Figure 5) and for years of high sunspot number (Figure 6). These contour maps are derived from observations rather than a model. They indicate that during years of high sunspot number the peak of the MUF (3000) F_2 in the southern hemisphere is 36 MHz; in low sunspot years the peak is 28 MHz The time of peak MUF is 18-22 LT during high sunspot number years, 13-17 LT during low solar activity years. MUF 3000 (F_2) is related to foF_2. Ascension Island belongs between the American sector and the Europe-Africa sector but the ionospheric path to the satellite is through the Appleton anomaly.

Observations of total electron content at Ascension Island (J.A. Klobuchar, private communication), December 1979 - February 1980, with the propagation path corresponding to that to MARISAT showed maxima at 1830 LT and at 2100-2200 LT corresponding well with the times of maxima in Figure 6.

During the December period, observations were made with AE-E using the retarding potential analyzer (W.B. Hanson, private communication, 1980). The value of N was greater by an order of magnitude in the anomaly region than near the magnetic equator. The ratio, $(\Delta N/N)$, was the same order of magnitude in both regions. Scintillation intensity which is proportional to ΔN would then be expected to be greater in the anomaly region

EARLIER OBSERVATIONS

Since the Ascension Island data could not resolve the question of the importance of the geometrical versus the latitudinal variables it was thought worthwhile to reexamine older data to try to resolve the question.

Between November 1969 and June 1970, when a high sunspot number of ~120 was reported, observations were made of an S band (2.2 GHz) transponder on the moon (Christiansen, 1971) from Ascension Island with both Earth and moon based signals in this frequency range. Scintillations as large as 20 and 25 dB on the two-way path, ground to transponder and return, were noted. In addition, Canary Island observations from high dip latitudes on the other side of the magnetic equator were also disturbed. Guam observations (at a different longitude) showed shorter periods of lower intensity scintillations during this time. No other stations on the magnetic equator made observations.

The Christiansen data were reanalyzed to determine if high scintillation levels were predominantly a function of azimuth (the vertical sheet model predicts maximum index at an azimuth of ~340°) or if high levels were noted at all azimuths of the moon's path as observed from Ascension Island. Figure 7 is a plot of azimuths when scintillation was greater than 10 dB for the 2200 MHz two-way path. Approximately 45 days of data were available.

The results are fairly conclusive. High intensity scintillation occurs at all azimuths. The

levels of scintillation were high at many azimuths (elevation angles were greater than 25° in all of the data illustrated). Similar results were noted for the Canary Island observations of 2.2 GHz signals.

More recently, similar results of high scintillation activity were noted by Fang (1980) at Hong Kong. He recorded 9 dB fluctuations on the 4 GHz COMSAT downlink to Hong Kong at azimuths ∿105° (the Pacific Ocean satellite) and 252° (the Indian Ocean satellite) for periods of time up to an hour. Thus, high values were noted almost perpendicular to the magnetic meridian.

DISCUSSION AND CONCLUSIONS

From theoretical studies, (Basu and Kelley, 1979 and references therein) it seems most likely that the generation of the irregularity bubbles commences at or near the magnetic equator. Once formed, the irregularities in the patch below 600 km would not be observed at Ascension. The field lines of the irregularities below 600 km terminate at the 200 km level at latitudes lower than those of the Ascension propagation path. One effect of increased solar sunspot number is to raise the height of the more intense irregularities (Aarons et al., 1980a) as seen on the Jicamarca backscatter.

Increased solar activity does produce higher electron densities in the F region as shown by the long-term sounder observations, the TEC and the in-situ data. If the effect of irregularity formation acts on a percentage basis, then the high sunspot years should show increased scintillation levels at gigahertz frequencies at electrojet and anomaly latitudes. This has been seen at electrojet latitudes (Huancayo) where a comparison of Sept.-Oct. 1977 L-band data with Sept.-Oct. 1979 data showed higher intensities in 1979, the higher solar activity year. In addition, Fang's statement indicates that the increase is a general increase.

While the scintillation intensities during years of high sunspot number are higher in the anomaly regions than those in the electrojet sector, the contribution of geometry of the irregularity model to intensity levels has not been established.

ACKNOWLEDGEMENTS

The authors would like to thank the following: the Instituto Geofisico del Peru at Huancayo, the Federal University of the Northern Rio Grande at Natal, and the Eastern Space and Missile Center at Ascension. We would also like to thank Sunanda Basu for many conversations relative to the role of geometry and latitude in gigahertz scintillations levels and J.A. Klobuchar for his initial observations of the L-band transmissions.

REFERENCES

Aarons, J. (1977), Equatorial scintillations: A review, IEEE Trans. Ant. and Propagat., AP-25, No. 5, 729-736.

Aarons, J., E. MacKenzie and K. Bhavnani (1978), Equatorial and high latitude empirical models of scintillation levels. AGARD Conference Proceedings #238 on Operational Modelling of the Aerospace Propagation Environment.

Aarons, J., J.P. Mullen, H.E. Whitney and E.M. MacKenzie (1980a), The dynamics of equatorial irregularity patch formation, motion and decay, J. Geophys. Res., 85, No. A1, 139-149.

Aarons, J., J.P. Mullen, J.R. Koster, R.F. daSilva, J.R. Medeiros, R.T. Medeiros, A. Bushby, J. Pantoja, J. Lanat and M.R. Paulson (1980b), Seasonal and geomagnetic control of equatorial scintillations in two longitude sectors, Accepted for publication in J. Atmos. Terr. Phys.

Basu, S. and M.C. Kelley (1979), A review of recent observations of equatorial scintillations and their relationship to current theories of F-region irregularity generation, Radio Sci., 14, No. 3, 471-485.

Basu, S., S. Basu, J.P. Mullen and A. Bushby (1980a), Long-term 1.5 GHz amplitude scintillation measurements at the magnetic equator, Geophys. Res. Lett., 7, No. 4, 259-262.

Basu, S., J.P. McClure, S. Basu, W.B. Hanson and J. Aarons (1980b), Coordinated study of equatorial scintillation, in-situ and radar observations of nighttime F-reg on irregularities, J. Geophys. Res., 85, No. A10, 5119-5130.

Christiansen, R.M. (1971), Preliminary report of S-band propagation disturbance during ALSEP mission support, Report X-861-71-239, Goddard Space Flight Center, Greenbelt, MD.

Fang, D.J., D.J. Kennedy and C. Devieux (1978), Ionospheric scintillations at 4/6 GHz and their system impact, Paper presented at IEEE Electronics and Aerospace Systems Convention, Arlington, VA.

Fang, D.J. (1980), 4/6 GHz ionospheric scintillation measurements, AGARD Conference Proceedings #284 on Propagation Effects in Space/Earth Paths.

Mikkelsen, I.S., J. Aarons and E. Martin (1978), Geometrical considerations of 136 MHz amplitude scintillation in the auroral oval, J. Atmos. Terr. Phys., 40, 479-483.

Rawer, K. (1963), "F2 layer ionization" in Advances in Upper Atmosphere Research, ed. B. Landmark, Pergamon Press, Oxford, England.

Weber, E.J., J. Buchau and J.G. Moore (1980), Airborne studies of equatorial F layer irregularities, J. Geophys. Res., 85, No. A9, 4631-4641.

Whitney, H.E. (1974), Notes on the relationship of scintillation index to probability distributions and their uses for system design, AFCRL-TR-74-0004, Air Force Cambridge Research Laboratories, Bedford, MA.

Table 1

Geometry of Primary Observations

Station	Elevation Angle	Azimuth Angle	Prop. Angle	400 km Sub-Ion. Lat.	Long.	Approx. Dip Latitude
Huancayo, Peru	20°	83°	85°	11°S	68.5°W	2.5°N
Natal, Brazil	65°	75°	84°	5.6°S	33.7°W	5.5°S
Ascension Island	80°	356°	52°	7.4°S	14.4°W	17°S

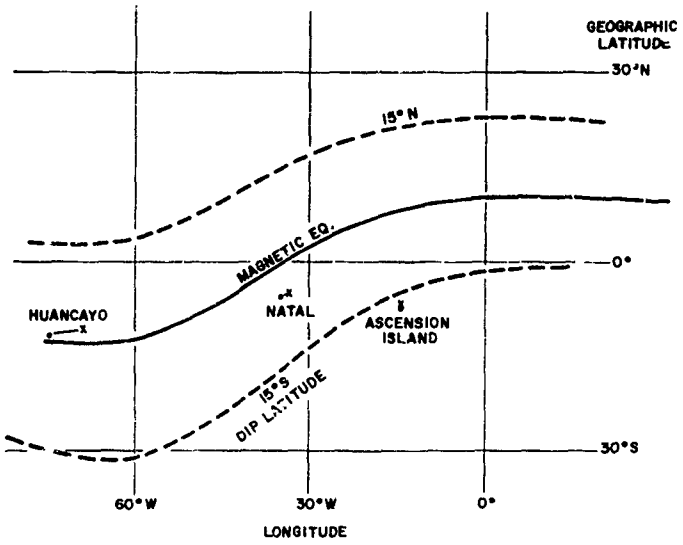


Fig. 1. Map of equatorial region in the 0-70°W longitude region using the 1975 epoch of the Defense Mapping Agency magnetic inclination map for magnetic parameters.

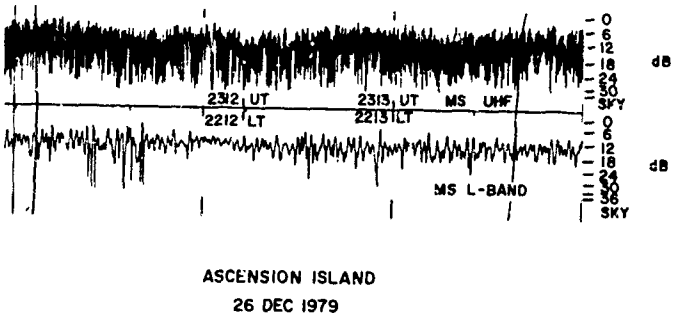


Fig. 2. Sample of both UHF and L-band data recorded at Ascension Island during December 1979 - January 1980. Note excursions ≥ 27 dB on both UHF and L-band channels.

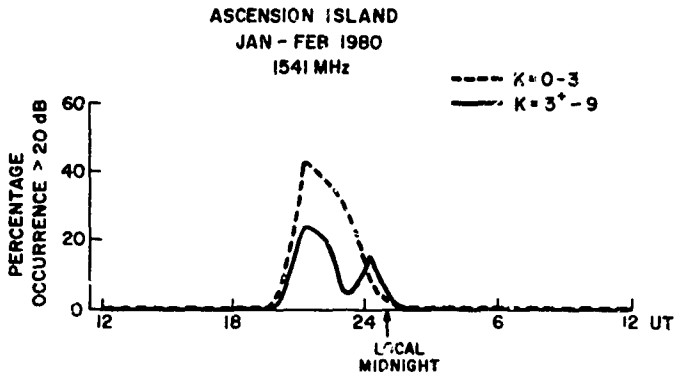


Fig. 3. Percentage occurrence of L-band scintillations ≥ 20 dB at Ascension Island during January-February 1980. Observations are segmented into quiet (Kp = 0-3) and disturbed (Kp = 3^{+}-9) magnetic conditions.

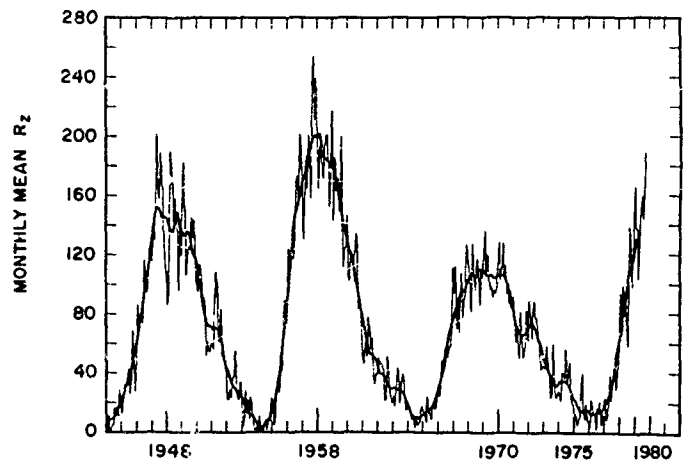


Fig. 4. Plot of mean Zurich sunspot numbers from 1944-1980.

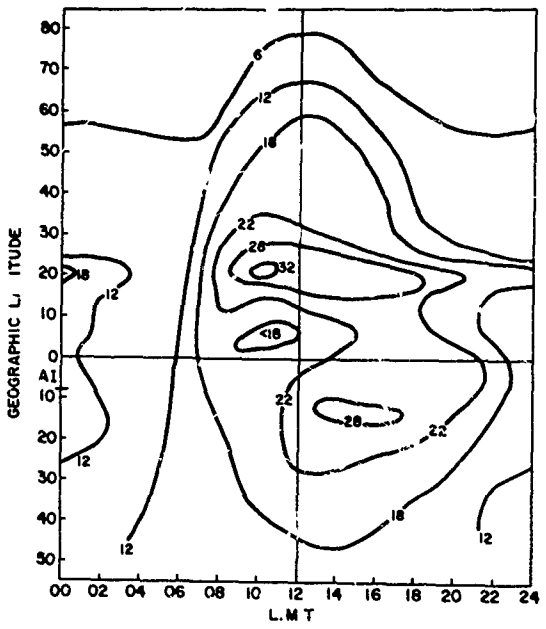


Fig. 5. Map of MUF (3000) F_2 (December) in Europe-Africa zone derived from observations during years of low solar flux (Rawer, 1963).

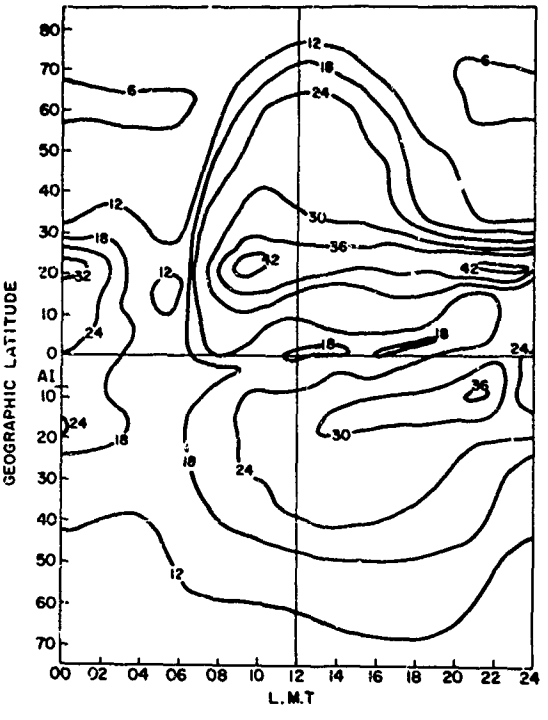


Fig. 6. Map of MUF (3000) F_2 (December) in Europe-Africa zone derived from observations during years of high solar flux (Rawer, 1963).

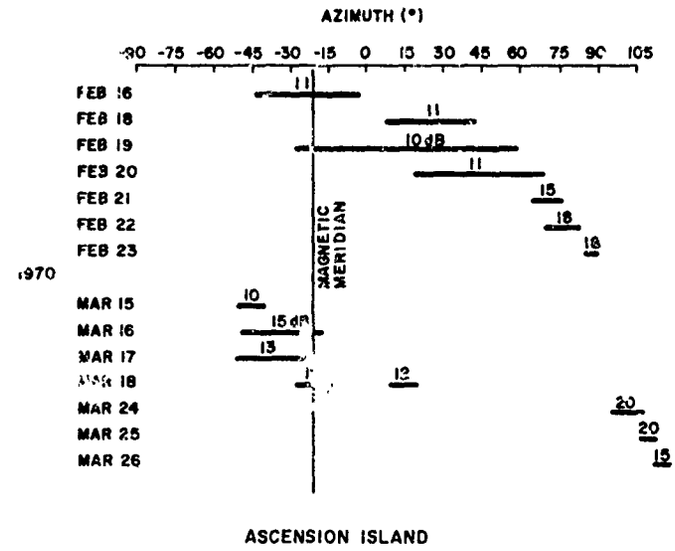


Fig. 7. Azimuths of 2.2 GHz lunar observations of Christiansen, 1971 with maximum dB excursions for e[illegible] grouping.

IONOSPHERIC SCINTILLATION OBSERVATIONS AT NATAL

K.C. Yeh
Department of Electrical Engineering
University of Illinois
Urbana, IL 61801 U.S.A.

J.P. Mullen
Air Force Geophysics Laboratory
Hanscom Air Force Base, MA 01731 U.S.A.

J.R. Medeiros, R.F. da Silva and R.T. Medeiros
Grupo de Astronomia e Fisica da Ionosfera
DFTE-Centro de Ciencias Exatas
Universidade Federal do Rio Grande do Norte
59.000 - Natal - RN - Brazil

ABSTRACT

This report stresses ionospheric scintillation results obtained at the Ionospheric Radio Observatory, which is staffed by the Department of Theoretical and Experimental Physics, Center of Exact and Natural Sciences, Federal University of the Northern Rio Grande. The results are presented within the overall framework of 2½ years' observations of scintillation at Natal, an equatorial station (32.23°W, 5.85°S).

The Observatory recorded signals of MARISAT 1 (1976-17A) and LES-9 (1967-23B). the former at 257.55 MHz (65°el) and the latter at 249.6 MHz (48.5°-54°el). During the nighttime hours scintillation on both signals commonly exceeded 29 decibels peak-to-peak. Scintillation measurements which were performed with spaced receivers on a 278 meter E-W baseline indicated eastward irregularity velocities of approximately 30 - 150 meters per second during the nighttime hours. The fading rates observed ranged from 1 per second to 6 per minute. For weak scintillation the ratio of fading rate and drift speed remained constant for a range of about 0.1 sec^{-1} to 0.4 sec^{-1}. For stronger scintillation the ratio increased markedly due to increased fading rates caused by multiple scatter effects.

I. INTRODUCTION

In July 1974 experimental observations and measurements of ionospheric scintillations near the magnetic equator were initiated on the campus of the Federal University of the Northern Rio Grande, Natal, Brazil (geographic coordinates 35.23°W, 5.85°S, dip -9.6°). In the beginning observations were made on the 254 MHz signals transmitted by the LES-6 (1968-81D) satellite which, since January 1977, has been switched to the 257.55 MHz signals transmitted by the MARISAT-1 (1976-17A) satellite. The receiving system consists of a 30-foot parabolic reflector, a standard receiver and recording device. The overall system response time is 0.1 second.

During the period 27 February to 13 March 1978 an intensive equatorial ionospheric scintillation program was performed. The program included participation of personnel associated with the Federal University of the Northern Rio Grande (UFRN), the Geophysical Institute of Peru (IGP), the University of Texas at Dallas (UTD), the University of Illinois at Urbana-Champaign (UI), Emmanuel College (EC), the Air Force Geophysics Laboratory (AFGL), and the Space and Missiles Test Center (SAMTEC). The purpose of the experiment was to study the generation, persistence and decay of the large scale irregularity structures which were responsible for diffraction of radio waves transmitted through the ionosphere. The observational sites operated included one in Brazil (Natal), three in Peru (Ancon, Jicamarca and Huancayo) and one in Ascension Island. In addition, the AFGL Airborn Ionospheric Observatory flew selected flight patterns over Peru and Ascension Island. A wealth of data has been obtained during this two-week period and some of these data has already been reported (Aarons et al., 1980) and more will be forthcoming. In this report we will concentrate on observations and preliminary results gleaned from the UFRN station at Natal only.

In presenting our results, we will first discuss the long term behavior and properties associated with large geophysical scales. These results are followed by discussions of those with ever increasingly smaller scales both in time and space. The paper is then concluded with a discussion and comparison with other scintillation results and with implications of these results on generating mechanisms.

2. EXPERIMENTAL RESULTS

The equatorial scintillation activities are known to have a strong diurnal and seasonal dependence. Such activities at Natal are depicted in Fig. 1. Because of the proximity of subionospheric intersections (5.35°S, 35.4°W for LES-6 and 5.3°S, 33.7°W for MARISAT-1) and closeness of transmission frequencies (254 MHz for LES-6 and 257.55 MHz for MARISAT-1), scintillation data from both satellites are combined to generate the occurrence contours. For purpose of plotting Fig. 1, the scintillation index is defined by the formula:

$$SI(\%) = (P_{max}-P_{min})/(P_{max}+P_{min}) \times 100 \qquad (1)$$

where P_{max} is the maximum power observed on the chart record and P_{min} is the minimum power. A discussion of the scintillation index SI as defined by (1) and its relation to other definitions of scintilllation index such as peak-to-peak fluctuation in dB and S_4 index has been made by Whitney (1974). The contours of Fig. 1 show a high probability of occurrence of scintillation greater than 60% or equivalently a peak-to-peak fluctuation of more than 6 dB at approximately 250 MHz near the Nov.-Dec. months and a minimum probability at the June-July months. No appreciable scintillation at this frequency was found during the daylight hours. The diurnal maximum was found at 2100-2300 LMT, which is not very different from that found at Huancayo and reviewed by Aarons (1977).

In addition to observing radio transmissions from MARISAT-1, transmissions from LES-9 (1976-23B) were also received. Of the 14 nights, scintillation occurred on ten of them. We have chosen one of these for detailed examination. Fig. 2 compares the scintillation for the nights of 9-10 March 1978 observed on the signals of MARISAT-1 and LES-9 (1976-23B). The observational angles for MARISAT-1 were constant, being 65° in elevation and 76° in azimuth. The observational angles for LES-9 were changing, varying between 48.5°-54° in elevation and 31° -299° in azimuth on the night of 1-2 March and being 41°-44° in elevation and 318° in azimuth on the night of 2-3 March. The scintillation shown in Fig. 2 was quite intense, exceeding 30 dB on MARISAT-1 and 16 dB on LES-9. The difference in these maximum scintillation levels can be attributed in large part to differences in the dynamic range of the two receiving systems; the MARISAT-1 receiver system utilized a 30 ft. parabolic reflector having a nominal gain of 25 dB whereas the LES-9 equipment used single yagi with a gain of 12 dB. As seen from Fig. 2 the intense scintillation continued for more than four hours. There was little discernible evidence of multiple onsets and decays that would be indicative of multiple discrete irregularity clouds either in Fig. 2 or the raw data. However, the initial onset time (approximately 45 minutes after sunset) generally agrees with the long term results shown in Fig. 1. The sub-ionospheric points for a 400 km intersection are (5.34°S, 33.71°W) for MARISAT-1 and (3.68°S, 38.11°W) for LES-9. This spacing corresponds to an east-west distance of approximately 530 km and thus furnishes an opportunity to examine the relative timing of the onsets of strong (i.e., greater than 6 dB or 60%) scintillation. When scaled from the original data the onset is found to be 2155 UT for MARISAT-1 and 2159 UT for LES-9. Because the east-west distance (250 km) is too great to be transversed by a reasonable irregularity drift in the measured time (a physically unreasonable speed of 2.2 km/s is required), it must be hypothesized that the onsets represented the generation of two separate irregularity clouds. In this example the easterly location along the MARISAT-1 path encountered the irregularity cloud 4 minutes earlier than the westerly location along the LES-9 path. Fig. 3 shows six other instances in which scintillation began earlier in the easterly direction. Also shown in Fig. 3 are two cases (March 1-2 and 2-3) in which there was no scintillation seen on MARISAT-1 but measureable scintillation on LES-9. On March 7-8 scintillation occurred earlier in the west.

In mid-1977 a second receiving system was established by the University of Illinois 278 meters away to perform multifrequency and multistation observations of MARISAT-1. During the program this system recorded the 257.55 MHz signal to MARISAT-1. Interconnection of this system with the AFGL equipment enabled drift measurements to be made. A helical antenna with a gain of 16 dB is used as compared with a 25 dB gain provided by the 30-foot parabolic antenna. In addition to recording amplitudes on a paper chart, some data are recorded digitally on magnetic tapes with a 20 Hz sampling rate. Figs. 4(a) and 4(b) depict a small section (102 seconds) of the two-station data reproduced from the digital tape. A close inspection of the data shows a high correlation with a slight time lag between the two channels. In fact the computed normalized cross correlation function (see Fig. 5) exhibits a maximum of 0.905 at a 2.00 second lag. Since the cross correlation coefficient is so close to unity the spaced-station data can be interpreted to mean an eastward drift of 139 m/s. This value is in close agreement with the value obtained by Mullen et al. (1977) using the same drift measurement technique at Huancayo, Peru and by Woodman (1972) using the radar Doppler technique. The power spectra computed for these two channels are very similar as depicted in Figs. 6(a) and (b). Both spectra show a fairly flat spectrum in the low frequency regime extending to a roll-off fre-

quency known as the Fresnel frequency ν_F, beyond which the spectrum approaches asympotically to ν^{-n} where n is equal to 3.4 for Fig. 6(a) and 3.6 for Fig. 6(b). Taking an average value of n = 3.5, this would imply a power-law electron density irregularity spectrum of the form $\kappa^{-4.5}$. The Fresnel frequency ν_F for both spectra is found to be 0.16 Hz. Theoretically, it has been found that the major contribution to the scintillation phenomenon comes from those irregularities whose spatial frequency is equal to the Fresnel wave number κ_F (Wernik and Liu, 1974). For the geometry of interest it is given by

$$\kappa_F = \sqrt{\pi k/z} \tag{2}$$

where k is the wave number of the radio wave and z is the slant distance to the middle of the irregularity region. Since the cross correlation coefficient is 0.905, we may use the frozen-in concept (Tatarskii, 1971) to relate the spatial behavior to the temporal behavior. The application of this concept to the present problem (Umeki et al., 1977) yields a relation

$$\kappa_F = (2\pi\nu_F/v_d) \tag{3}$$

where ν_F is the Fresnel frequency (whose value can be scaled from the power spectrum of the scintillating amplitude) and v_d is the drift velocity (whose value can be measured from the spaced-station data such as described here). Eliminate κ_F between eqs. (2) and (3) and solve for z to obtain

$$z = \frac{1}{2\lambda}\left(\frac{v_d}{\nu_F}\right)^2 \tag{4}$$

where λ is the wavelength of the radio wave. Inserting the experimentally measured values for v_d(=159 m/s) and ν_F(=0.16 Hz) in eq. (4), the slant range to the middle of the irregularity region is computed to be 324 km. This corresponds to a vertical height of 294 km (as the elevation angle from the ground station to the satellite is 65°), which compared favorably with measured heights by other techniques (Woodman and LaHoz, 1976; McClure et al., 1977; Aarons et al., 1980).

Fig. 7 shows the behavior of 257.55 MHz scintillation during the evening of March 9, 1978. Evidence of scintillation first appeared at 2154 UT. This scintillation rose quickly in intensity and was fully developed in about 2 minutes at which time the peak-to-peak fluctuations were 15 dB, which corresponded to a S_4 scintillation index of 0.63. About 10 minutes later the scintillation reached its maximum value of 20 dB (corresponding to S_4=0.80). It stayed at this high value for about 40 minutes, after which a gradual decay phase set in. This decay continued for more than five hours until the scintillation completely disappeared at 0440 UT March 10, 1978. During the same time the fading rate had a behavior shown in the middle panel of Fig. 7. Because of the slow recording speed before 0020, which made scaling the fading rate impossible, the fading rate data were available only after 0020 UT. In this report, we define the fading rate as the average number of signal peaks per second in a record of one minute duration centered about the time of interest. It is seen in Fig. 7 that the fading rate decreased continuously from nearly 1 sec^{-1} shortly after midnight to 0.1 sec^{-1} in three hours. During the same time the time delay of the fading patterns observed at the spaced stations increased steadily from less than 2 seconds to more than 6 seconds as depicted by the curve on the bottom panel of Fig. 7.

The apparent anticorrelation between the fading rate and the time delay as evidenced in Fig. 7 is interesting and needs further investigation. Since the baseline distance and orientation are known, it is possible to compute the component of the drift velocity along the baseline. The results are shown in Fig. 8. The eastward drift was approximately 150 m/s shortly after the midnight and it decreased steadily to a value of approximately 30 m/s at 0400 UT. The average trend of the fading rate obtained by replotting the data shown in Fig. 7 is also depicted in Fig. 8. Both the velocity curve (marked V) and the fading rate curve (marked F) showed a decreasing trend after about 0100 UT. This is expected; since under the weak scatter limit and the frozen-in hypothesis, the scintillation theory predicts the fading rate to be directly proportional to the irregularity drift speed. However, before 0100 UT when the scintillation was strong, the ratio of fading rate and drift speed becomes much higher. This is attributed to the multiple scattering effects which give rise to small scale structures in the scattered field. Theoretical computations have shown that effects caused by multiple scattering will actually decrease the correlation intervals (Yeh. et al., 1975). In other words, for a given drift speed, the fading rate will increase with the increasing scintillation. As the observed scintillation was decreasing through the period of interest, it may explain why the average rate of decrease of the fading rate was quite different from that of the velocity in the beginning, even though they were very similar near the end of the scintillation period as shown in Fig. 8.

3. DISCUSSION AND CONCLUSION

In section 2 we have described our scintillation observations at Natal, Brazil. In this section we will discuss implications of our observational results and compare them with other observations.

The diurnal and seasonal occurrence behavior depicted by contours of Fig. 1 shows a single maximum centered around October-November months and 2300-0100 UT (or approximately 2100-2300 LMT). This is not very different from the results obtained for Huancayo, Peru (Aarons, 1977) in months and local time for maximum activities, although Huancayo has a tendency to show a secondary maximum in the month of February. Notice the absence of any activity in the months of May through August in Fig. 1. This is also the case in Huancayo. However, in the African sector as represented by Accra, Ghana, even though the monthly and local time maxima are not very different from the South American sector, low level to medium level scintillation activities extend to all months of the year (Aarons, 1977). One possible explanation might be the different geometries encountered; Accra observed the satellite at 12° elevation whereas Huancayo saw it at 70°. Using the DNA Wideband satellite data Livingston (1980) shows that at Kwajalein in the Pacific sector, the scintillation activities actually peak in the months of July and August, that is during the months when South America is nearly void of any scintillation activity and Africa has only moderate scintillation activity.

In all sectors around the equator, the observations of scintillation are usually confined to the local nighttime. Since the intense scintillation is a post-sunset phenomenon, an earlier onset in scintillation at an eastern point may be expected. On a 1100 km east-west spacing at ionospheric heights (Bandyopadhyay and Aarons, 1970), scintillation was commonly seen earlier in the east, but nevertheless in 20% of the time it was observed earlier in the west. The east-west ionospheric spacing of 520 km used in our experiment strongly suggest that the onset of scintillation at these points is usually independent. More recently multi-satellite observations with various east-west spacing indicate that immediately following the formation of a new patch, the western boundary of the patch may expand westward by 300 km (Aarons et al., 1980). This suggests that whatever mechanism that is responsible for triggering the instability it is small (smaller than 300 km) in spatial extent initially. As soon as the instability mechanism is activated, the formation of the irregularity goes through a growth phase during which the patch size increases to several hundred kilometers in the east-west direction. The patch is then subject to electrodynamic drift which, during the post-sunset period, is generally in the eastward direction near the magnetic equator.

The scintillation data collected on a 278 meter spacing have been used to calculate the drift. The resulting velocity shown in Fig. 8 is eastward, decreasing from 150 m/s at 0000 UT to 30 m/s at 0400 UT. This diurnal trend and the absolute magnitude are in general agreement with other results (Woodman, 1972; Mullen et al., 1977). According to the weak scatter theory, the velocity and the fading rate should be proportional. This is supported by the experimental results of Fig. 8. Departure from this proportionality relation is expected when the scintillation is strong since multiple scatter contributes additional decorrelation effects (Yeh et al., 1975; Rino, 1979). This is also supported by Fig. 8. The computed power spectra of the amplitude scintillation data are similar to those obtained by other observers (Crane, 1976; Umeki et al., 1977; Rino and Fremouw, 1973; Rufenach, 1972; Meyer et al., 1979). From the computed power spectra such as those shown in Fig. 6 the Fresnel frequency can be determined. If one applies the frozen-in concept (Tatarskii, 1971) the Fresnel wave number K_F can be determined by Eq. (3) provided the drift speed v_d is also known such as by using the spaced-station data. Now the weak scintillation theory (Wernik and Liu, 1974) can be used to relate K_F to the slant range and thus also the height to the region of irregularities. The computed height in one instance is 294 km, which is quite reasonable. This lends credence to the frozen-in hypothesis and the correctness of the scintillation theory, at least under the weak scatter condition.

ACKNOWLEDGEMENT

The work reported in this paper was partially supported by the U.S. Army Research Office under grants DAAG 29-76-G-0286 and DAAG 29-78-G-0101 and partially supported by the U.S. Air Force under grant AFOSR 78-3613. We wish to thank S. Basu for comments on the manuscript. Computations of the spectra and the cross correlation function were made by S.J. Meros.

REFERENCES

Aarons, J., Scintillation studies of the ionosphere. In AGARDograph No. 166 on Total Electron Content and Scintillation Studies of the Ionosphere, edited by J. Aarons, 1973.

Aarons, J., Equatorial scintillation: A review, IEEE Trans. Ant. Propagation. AP-25, 729-736, 1977.

Aarons, J., J. Buchau, S. Basu and J.P. McClure, The localized origin of equatorial F region irregularity patches, J. Geophys. Res., 83, 1659, 1978.

Aarons, J., J.P. Mullen, H.E. Whitney and E.M. MacKenzie, The Dynamics of equatorial irregularity formation, motion and decay. J. Geophys. Res., 85, 139-149, 1980.

Bandyopadhyay, P. and J. Aarons, The Equatorial F-Layer Irregularity Extent as Seen from Huancayo, Peru, Radio Science, 5, 931-938, 1970.

Crane, R.L., Spectra of ionospheric scintillation, J. Geophys. Res., 81, 2041-2050, 1976.

Koster, J.R., Phase and amplitude scintillation at the equator, AFGL-TR-78-0289, 31 Oct. 78.

Livingston, R.C., Comparison of multifrequency equatorial scintillation: American and Pacific sectors. To appear in Radio Science, 1980.

Meyer, W.J., R.J. Gjeldum, C.H. Liu and K.C. Yeh, A study of ionospheric scintillations of phase and quadrature components, J. Geophys. Res., 84, 2039-2048, 1979.

Mullen, J.P., H.E. Whitney, S. Basu, A. Bushby, J. Lanat and J. Pantoja, Statistics of VHF and L-band scintillation at Huancayo, Peru, J. Atmos. Terr. Phys., 39, 1251-1261, 1977.

McClure, J.P., W.B. Hanson and J.H. Hoffman, Plasma bubbles and irregularities in the equatorial ionosphere, J. Geophys. Res., 82, 2650-1977.

Rino, C.L. and E.J. Fremouw, Statistics for ionospherically diffracted VHF/UHF signals, Radio Sci., 8, 223, 1973.

Rino, C.L., A power-law phase screen model for ionospheric scintillation, 2. Strong scatter, Radio Sci., 14, 1147-1155, 1979.

Rufenach, C.L., Power-law wavenumber spectrum deduced from ionospheric scintillation observations, J. Geophys. Res., 77, 4761, 1972.

Tatarskii, V.I., The effects of the turbulent atmosphere on wave propagation, Nat. Tech. Infor. Ser., Springfield, VA , 1971.

Umeki, R., C.H. Liu and K.C. Yeh, Multifrequency spectra of ionospheric amplitude scintillations, J. Geophys. Res., 82, 2752-2760, 1977.

Wernik, A.W. and C.H. Liu, Ionospheric irregularities causing scintillation of GHz frequency radio signals, J. Atmos. Terr. Phys., 36, 871-879, 1974.

Whitney, H.E., Notes on the relationship of scintillation index to probability distributions and their uses for systems design, AFGL-TR-74-0004, 3 January 1974.

Woodman, R.F., East-west ionospheric drifts at the magnetic equator, Space Res., 12, 969-974, 1972.

Woodman, R.F. and C. LaHoz, Radar observations of F region equatorial irregularities, J. Geophys. Res., 81, 5447, 1976.

Yeh, K.C., C.H. Liu and M.Y. Youakim, A theoretical study of the ionospheric scintillation behavior caused by multiple scattering, Radio Science, 10, 97-106, 1975,

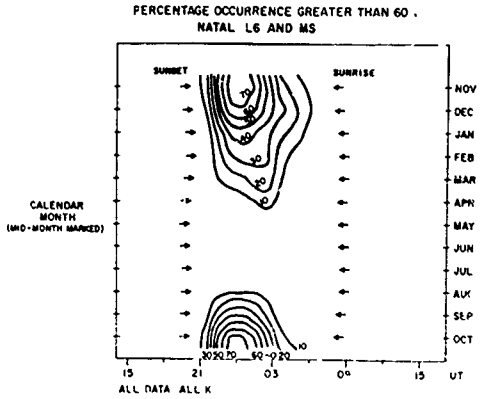


Fig. 1. Percentage occurrence contours showing scintillation index SI having a value greater than 60%. Data from MARISAT-1 at 257.55 MHz and data from LES-6 at 254.14 MHz are combined to generate this figure.

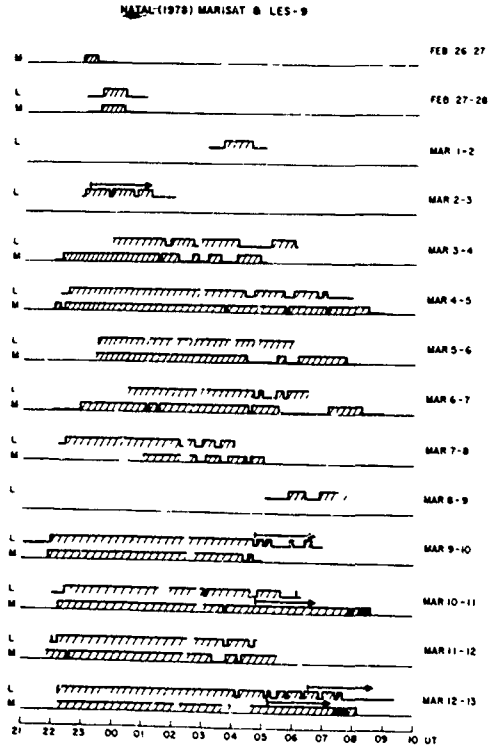


Fig. 2. Comparison of scintillation index in dB for signals transmitted by MARISAT-1 and LES-9. The difference in the maximum scintillation level is believed to come from differences in the dynamic range of two receiver systems (See text).

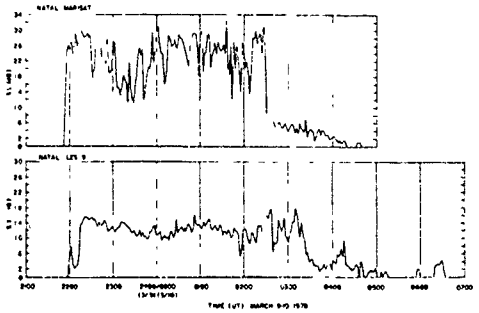


Fig. 3. Comparison of LES-9 and MARISAT-1 scintillation activities during the two-week period. Arrows indicate weak scintillation; otherwise, the blocks represent scintillation of 10 dB or more. Breaks indicate lack of data.

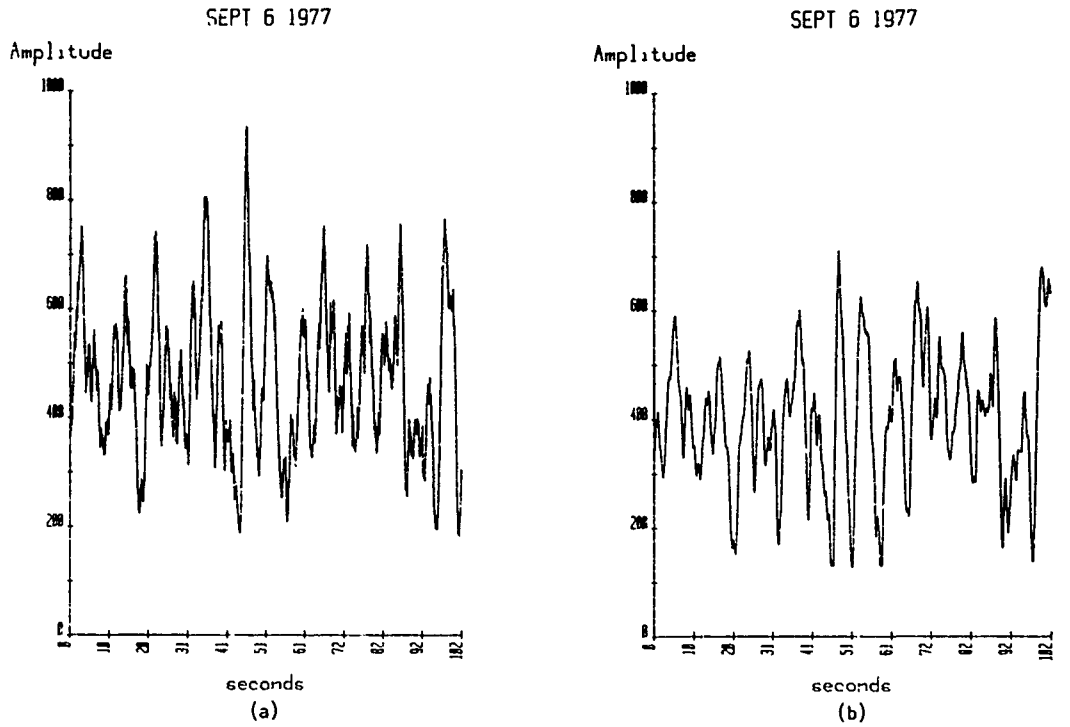


Fig. 4. Amplitude scintillation of 257.55 MHz signals transmitted by the geostationary satellite MARISAT-1 (1976-17A) and received at Federal University of the Northern Rio Grande, Natal. Notice the close correlation of signals between these two channels. The data shown in (a) was recorded on a base line of 278 meters to the west of that shown in (b).

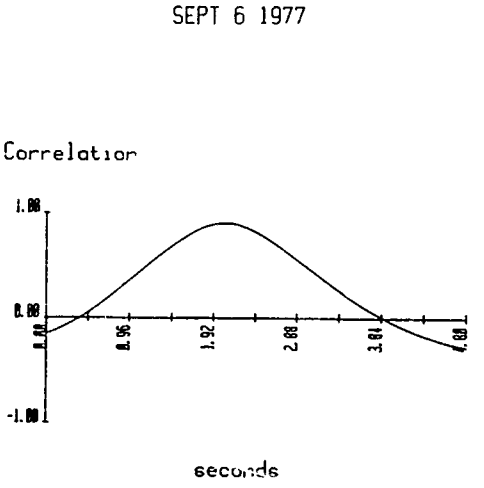


Fig. 5. The cross correlation function of the two-station data shown in Fig. 4.

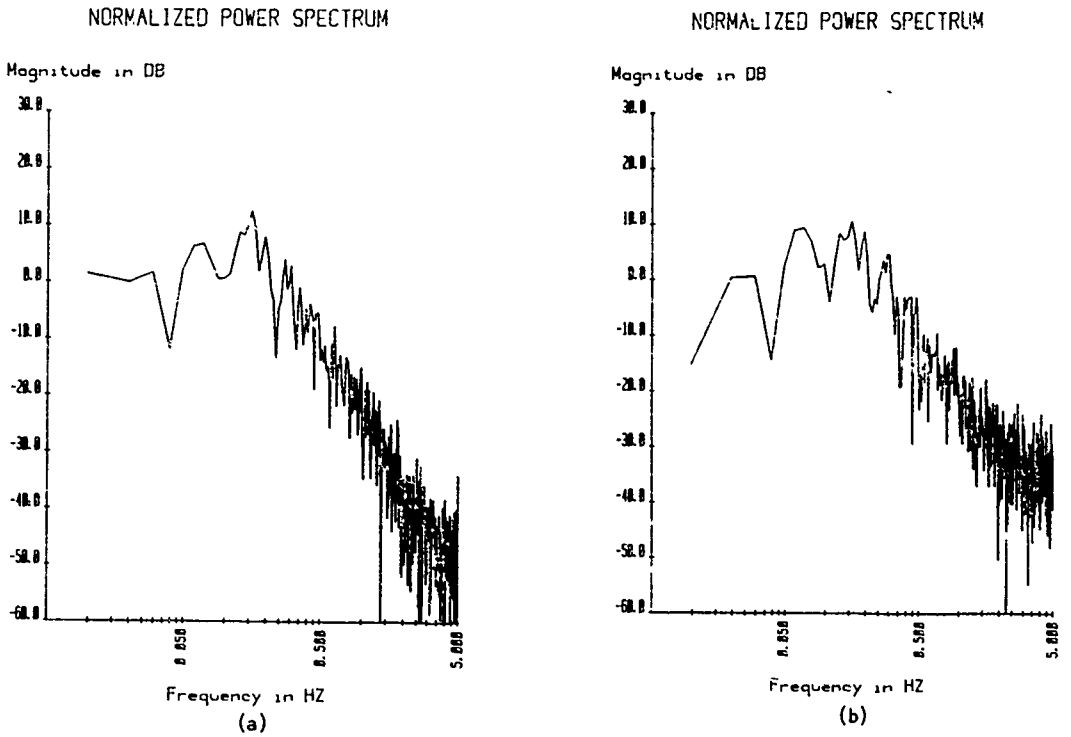


Fig. 6. The amplitude power spectra for the data shown in Fig. 4(a) and Fig. 4(b) respectively

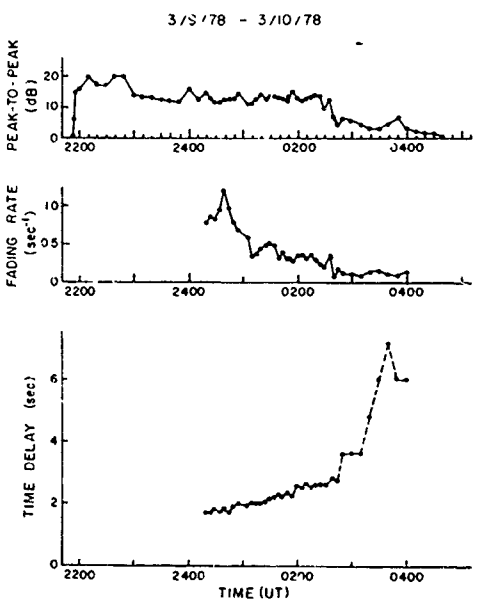


Fig. 7. Peak-to-peak scintillation, fading rate and time delay of scintillating signals during the evening of March 9, 1978.

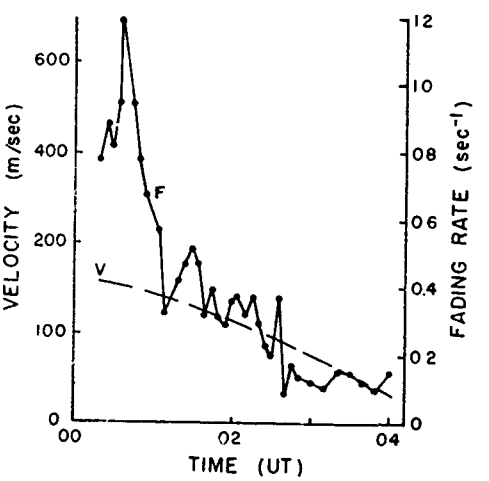


Fig. 8. A comparison of the fading rate (F) and the eastward drift velocity (V).

VHF SCINTILLATION NEAR THE EQUATORIAL ANOMALY CREST IN THE INDIAN SECTOR

A. DasGupta* and A. Maitra
Institute of Radio Physics and Electronics
University of Calcutta
Calcutta 700009, India

ABSTRACT

Some results of observations on the amplitude scintillation of a 136 MHz transionospheric signal recorded at Calcutta, which is situated virtually below the northern crest of the equatorial anomaly in the Indian sector, during the period April 1977 through February 1980 are presented. For the low solar activity period scintillation has been found to be essentially a local summer phenomenon at this location. With the increase in solar activity the equinoctial occurrence increases remarkably and becomes the dominant phenomenon while the local summer occurrence shows little change. The results are discussed in the context of the present theories of equatorial irregularities, as related to the field tube confinement of the equatorial irregularities, and the variation of both the altitude of the F region and the extent of the upwelling motion at the magnetic equator with the solar activity.

INTRODUCTION

In recent years there has been an increasing interest in the understanding of the physics of the nighttime F region irregularities near the magnetic equator (Basu and Kelley, 1979). Even though a number of techniques have been used to explore the irregularities (Basu et al., 1980), the measurement of scintillations of signals from satellites remains the most widely used method for the study of kilometer scale irregularities. In the context of the current experimental and theoretical investigations it is apparent that the equatorial F region plasma within an entire flux tube goes unstable and take part in an upwelling motion (Anderson and Haerendel, 1979; Weber et al., 1980). As a result, the irregularities are also obtained north and south of the magnetic equator.

Most of the long-term observations have, however, been performed at locations like Huancayo and Legon, situated very close to the magnetic equator, and such long term observations from stations off the magnetic equator are lacking. Recently, extremely high level (>20 dB at 1.54 GHz) of amplitude scintillation has been observed (Aarons et al., 1980a) at Ascension Island (7.95°S, 14.4°W), located near the southern crest of the well known equatorial anomaly in F2 ionization in the African zone. Although the above results have focused attention on the problem very few long term scintillation observations have so far been reported from such locations. In this paper we present some results of 136 MHz amplitude scintillation measurements made during the three year period April 1977 through February 1980 at Calcutta (23°N, 88.5°E), which is situated virtually below the northern crest of the equatorial anomaly. We shall examine the seasonal variation of scintillation at this location during the above period and show that both the level of scintillation and the seasonal pattern are controlled by solar activity. This, as we shall discuss, can be related to the variation of the height of F2 layer at the magnetic equator with solar activity and the confinement of equatorial irregularities within the magnetic field tubes.

DATA

A polarimeter of the type described by Fig et al (1977) and received from the Air Force Geophysics Laboratory, Hanscom Air Force Base, Massachusetts, U.S.A. has been in operation at the Haringhata Field Station of the University of Calcutta. This was used to record the amplitude and polarization of the 136 MHz signal from the Japanese geostationary satellite ETS-2 since April 1977. This paper discusses the amplitude scintillations of the signal recorded by the above. The recording time constant of the system is 0.1 sec and the dynamic range is about 22 dB. Scintillation index in dB has been scaled manually every 15 minutes following the method outlined by Whitney et al. (1969). The subionospheric point, namely, the intersection of the propagation path from the station to the ETS-2 satellite with the 400 km ionospheric height, occurred at 21°N lat. and 92°E long. (dip 27°N).

*At present NRC/NAS Senior Resident Research Associate at the Air Force Geophysics Laboratory, Hanscom Air Force Base, MA 01731, U.S.A.

This subionospheric position is to the immediate south of the northern crest of the equatorial anomaly. It should be noted that in this longitude sector the height of the 400 km subionospheric point maps along the earth's magnetic field to an altitude of 800 km above the magnetic equator.

RESULTS AND DISCUSSION

Scintillation around the present location has varied considerably during the period April 1977 to February 1980 when the level of solar activity changed greatly. During 1977, the scintillation index at 136 MHz generally varied between 3-6 dB (DasGupta and Maitra, 1980) but in the later period the index often attained >22 dB and was limited by the dynamic range of the system.

Figure 1 shows the hourly percentage occurrence of nighttime scintillation (SI $\geq$ 3 dB) on a seasonal basis for 1977, 1978 and 1979. It is of interest to note that in 1977, a year of low solar activity, scintillations were observed mainly during the months of May-July. With the increase of solar activity during 1978, scintillations were also observed in the equinoctial periods February-April and August-October and in the months of November-January, in addition to the May-July period. In 1979, the year of recent maximum solar activity, the equinoctial occurrence of scintillations becomes much higher than that during the May-July period and the activity during November-January also registers some increase.

Scintillations usually occur predominately in the time period before local midnight during the equinoxes, while there is no such definite pattern for the other seasons. The temporal character of the equinoctial scintillations has been observed to be patchy, the discrete patches lasing for a few tens of minutes interspersed with absolutely quiet periods. This character is very similar to that observed near the magnetic equator by Aarons et al (1980b). On the other hand, scintillations observed during the May-July period in all the three years 1977-1979 indicate absence of discrete patches and instead show one or two continuous periods of scintillation activity lasing for several hours.

Figure 2 shows the percentage of time the VHF signal scintillates (SI $\geq$ 3 dB) for the different months individually. The monthly mean sunspot numbers are also indicated in the diagram to facilitate the comparison of the scintillation pattern with solar activity. As discussed in the previous paragraph, a semiannual component corresponding to the equinoctial scintillations becomes progressively more prominent with the increase in the solar activity while the annual component identified with scintillation occurrence during the May-July period shows a decreasing trend relative to the semiannual component (Table 1). The average occurrence or the d.c component in 1979 is nearly double of that in the previous year.

Observations at all the equatorial stations in African, American, Asian and Pacific sectors indicate high occurrences of scintillations during the equinoxes (Livingston, 1980; Aarons et al., 1980c). In this respect, the emergence of an equinoctial component of scintillations near the crest of the (daytime) equatorial anomaly in F2 ionization from the present data is very significant. As mentioned earlier the subionospehric location of the present observations are mapped along the earth's magnetic field to an altitude of about 800 km above the magnetic equator. Should the F-region irregularities above the magnetic equator extend to altitudes of 800 km and above, the kilometer scale irregularities at the magnetic equator are expected to be mapped down the field lines to the subionospheric location of the present observations. With increasing solar activity, the equatorial F-region vertical drift velocity around sunset increases (Fejer et al., 1979) and the height of the F-region maximum in the post-sunset hours has been found to attain altitudes as high as 500 km. In the context of the current theories of the generation of nighttime equatorial irregularities and their upwelling motion, field tubes at altitudes of 800 km and above are more likely to go unstable in the sunspot maximum years. Thus the present station is able to record the equinoctial component of the equatorial scintillation only during the years of high solar activity. It seems that when the level of solar activity is decreased, the present location becomes detached from the equatorial scintillation belt due to the decreasing altitude of the F region. Scintillations recorded during the May-July period probably arises from mid-latitude irregularities. As noted earlier, this component does not exhibit significant variation with solar activity.

In the global perspective of equatorial scintillation, the present results do not give any definite clue whether the Indian sector could be grouped either with the Asian sector or the African sector since the present observations show a prominent May-July occurrence of scintillation which changes little with solar activity while observations near the magnetic equator in India for a very limited period show occurrence minimum during the same period (Moorthy et al., 1979). Long term observations on the occurrence pattern of scintillation near the magnetic equator in the Indian sector could only resolve this question.

ACKNOWLEDGEMENTS

We are grateful to Santimay Basu and Sunanda Basu for helpful discussions. Thanks are also due to R.K. De for help in the observational program.

Table 1

Harmonic Analysis of percentage occurrences of scintillations at Calcutta (SI $\geq$ 3 dB)

Year	Constant Term or Average Value	Annual Component		Semiannual Component	
		Amplitude	Phase	Amplitude	Phase
1978	8.5	5.2	4 Jul.	4.9	26 Apr.
1979	17.2	0.8	8 Nov.	6.8	01 Apr.

REFERENCES

Aarons, J., H.E. Whitney, E. MacKenzie and S. Basu (1980a), Microwave equatorial scintillation intensity during the current solar maximu, Radio Sci., accepted for publication.

Aarons, J., J.P. Mullen, H.E. Whitney and E.M. MacKenzie (1980b), The dynamics of equatorial irregularity patch formation, motion adn decay, J. Geophys. Res., 85, 139.

Aarons, J., J.P. Mullen, J.R. Koster, R.F. daSilva, J.R. Medeiros. R.T. Medeiros, A. Bushby, J. Pantoja, J. Lanat and M.R. Paulson (1980c), Seasonal and geomagnetic control of equatorial scintillations in two longitude sectors, J. Atmos. Terr. Phys., 42, 861.

Anderson, D.N. and G. Haerendel (1979), The motion of depleted plasma regions in the equatorial ionosphere, J. Geophys. Res., 84, 4251.

Basu, Sunanda and M.C. Kelley (1979), A review of recent observations of equatorial scintillations and their relationship to current theories of F-region irregularity generation, Radio Sci., 14, 471.

Basu, S., J.P. McClure, Sunanda Basu, W.B. Hanson and J. Aarons (1980), Coordinated study of equatorial scintillation and in situ and radar observations of nighttime F region irregularities, J. Geophys. Res., 85, 5119.

DasGupta, A. and A. Maitra (1980), Scintillation of vhf satellite signals near the edge of the equatorial irregularity belt, Advances in Space Exploration, 8, (Low latitude Aeronomical Processes, ed. A.P. Mitra, Pergamon Press, Oxford and New York), p. 209.

Eis, K.E., J.A. Klobuchar and C. Malik (1977), On the installation, operation, data reduction and maintenance of VHF electronic polarimeters for total electron content measurements, Rept. AFGL-TR-77-0130.

Fejer, B.G., D.T. Farley, R.F. Woodman and C. Calderon (1979), Dependence of equatorial F region vertical drifts on season and solar cycle. J. Geophys. Res., 84, 5792.

Livingston, R.C. (1980), Comparison of multifrequency equatorial scintillation: American and Pacific sectors, Radio Sci., 15, 801.

Moorthy, K.K., C.R. Reddi and B.V. Krishna Murthy (1979), Nighttime ionospheric scintillations at the magnetic equator, J. Atmos. Terr. Phys., 41, 123.

Weber, E.J., J. Buchau and J.G. Moore (1980), Airborne studies of equatorial F layer ionospheric irregularities, J. Geophys. Res., 85, 4631.

Whitney, H.E., J. Aarons and C. Malik (1969), A proposed index for measuring ionopsheric scintillation, Planet. Space Sci., 17, 1069.

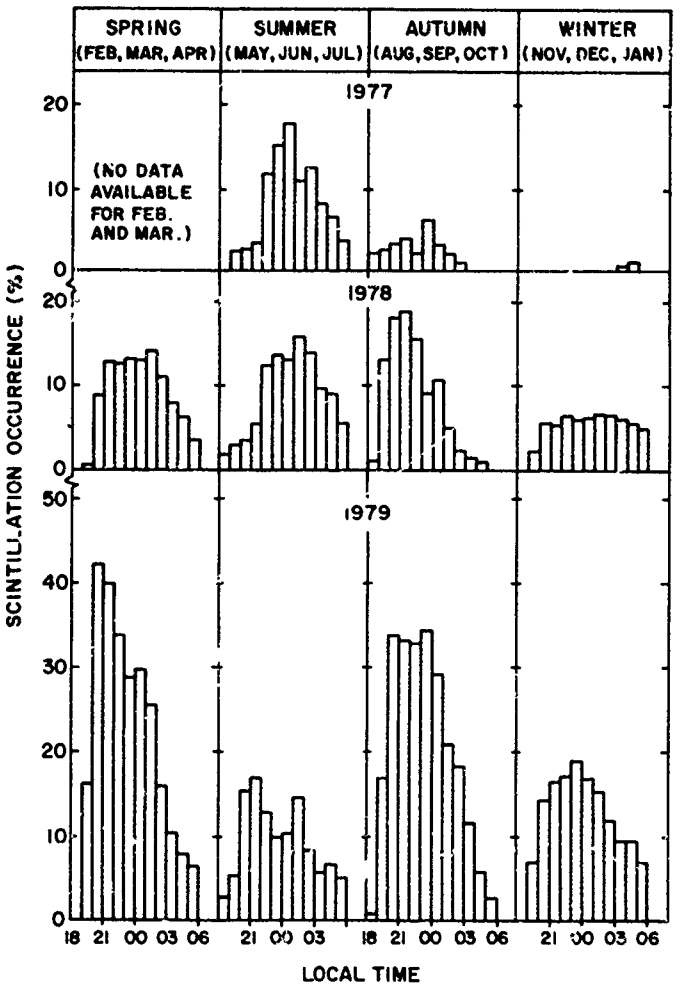


Fig. 1. Percentage occurrence of nighttime scintillations (SI $\geq$ 3 dB) at 136 MHz for different seasons.

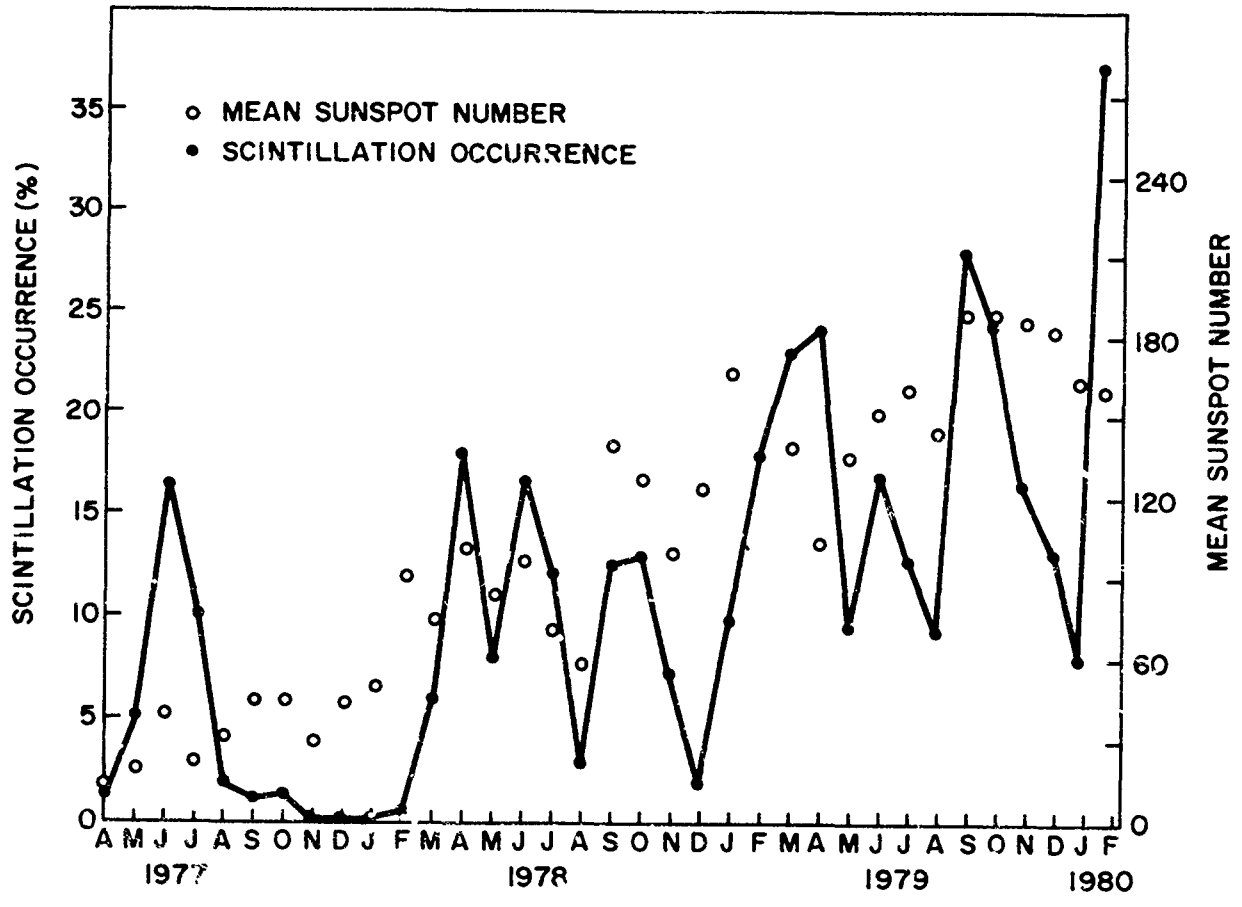


Fig. 2. Monthly variation of scintillation activity. The open circles show the monthly mean sunspot numbers.

CUMULATIVE DISTRIBUTION OF SCINTILLATION INDEX

A. I. Trivedi, M. R. Deshpande and hari Om Vats*
Physical Research Laboratory Ahmedabad, India

ABSTRACT

Amplitudes of radio beacons (40, 140, and 360 MHz) from ATS-6 (Phase II) recorded at Ootacamund, India have been used to determine the scintillation index every quarter hour. Cumulative distributions of scintillation indices have been evaluated with the aim of evolving an economical method of predicting distributions likely to be experienced over long periods. The quarter-hours that have a mean scintillation index values that fall within ± 2.5% of a specified scintillation index level have been grouped together. The centre scintillation index is assigned to the resulting distribution. Such distributions have been obtained over a representative range of centre scintillation index (S.I.) values spaced at 2.5% intervals of S.I. This procedure results in overlapping of half ranges. This improves accuracy and sample size, and it also overcomes to some extent the arbitrary nature of the classification into ranges. Based on these results, a procedure has been suggested for estimating the cumulative distribution of scintillation index, with possible application to estimating link reliability.

PROCEDURE

Amplitudes of radio beacons (40, 140 and 360 MHz) from ATS-6 (Phase II) recorded at Ootacamund have been used to determine the scintillation index as defined by Whitney et al. (1969) every quarter-hour. This scintillation index will be referred to as S_0 in this article. The distribution of these values has been evaluated from this resulting set of S_0 values. The annual mean value of S_0 for the corresponding quarter-hour has been evaluated. Analysis of the data corresponding to various seasons has also been performed. From these primary analyses, the quarter-hours having mean S_0 values falling within ± 2.5% of specified mean S_0 level have been grouped together, and the cumulative distribution of the S_0 values is evaluated and the centre S_0 value is assigned to the resulting distribution. Such distributions have been obtained over a representative range of centre S_0 values spaced at 2.5% intervals of S_0. This procedure results in overlapping of half ranges, but is considered desirable from the accuracy and sample size point of view and also because it overcomes to some extent the arbitrary nature of the classification into various ranges.

RESULTS

Analysis of the complete one-year data of ATS-6 (Phase II) recorded at Ootacamund was carried out with the above mentioned procedure and results for the frequencies of 40, 140 and 360 MHz are shown in Figs. 1, 2 and 3, respectively. The ordinate shows the level of scintillation activity, measured in terms of the scintillation index S_0. The scale on abscissa represents the percentage of times the S_0 value given on the ordinate is exceeded. Individual curves represent the overall distribution of S_0 values, when the individual distribution having a mean S_0 within ± 2.5% of the value shown as the parameter, are added and averaged, each component distribution being that of S_0 values seen over the year at a specific time of the day.

Examination of these curves indicates that the distribution varies in a systematic manner with the mean annual S_0. This variation is expected, since a higher mean signifies greater percentage occurrence of higher S_0 values, extrapolation to ranges which are not covered is possible. Such extrapolated cdf's are shown by the broken curves which have been derived using linear extrapolation based on the two previous ranges. Inference regarding frequency dependence of the cdf may be drawn by comparing the curves corresponding to the same mean value of scintillation index at different frequencies. Such a comparison between say the 35% curve at 40 MHz and extrapolated 35% curve at 140 MHz or the 10% curve at 140 MHz and the 10% curve at 360 MHz shows that the cdf has a noticeable frequency dependence.

---

*Presently at Department of Electrical Engineering and Computer Sciences, University of California, San Diego, supported by NSF Grant ATM 7802625.

This is explained by the fact that at widely separated frequencies, the ionospheric effects have to be very dissimilar in order to yield the same magnitude of the mean S_0 at the two frequencies. At 40 MHz 35% mean S_0 signifies a predominance of weak irregularity condition, while a 35% mean S_0 at 140 MHz would imply dominance of moderate to strong irregularity conditions.

Seasonal variations in the cdf for a specified mean seasonal S_0 level are shown in Fig. 4. Also shown is the annual averaged cdf for the same value of average S_0. While seasonal effects are evident, it is seen that the annual cdf lies within the extreme curves for summer and winter seasons. Furthermore, in the high S_0 range of 50% to 100% the annual cdf lies very close to the worst case of summer distribution. This range is of a specific interest for link performance evaluation, and hence the annual cdf curve may be considered as giving the approximate worst case S_0 distribution within the range. Thus for the intended application this may be regarded as an advantage.

Diurnal dependence of the cdf is given in Fig. 5. Here again E- and F-region effects give rise to differences in the cdf's which may partly be the effect of a reduced sample size. The irregularities in the E-region are mainly responsible for the day-time fading whereas those in the F-region are basically responsible for night-time fading. In deriving the curves of Figs. 1-3, however, these diurnal effects to a great extent have been implicitly accounted for since higher mean S_0 values occur mainly during night-time, and hence night-time data has a dominant influence in the cdf curves for higher parameter values, and the day-time data in the curves for lower parameter values. In the intervening sets of curves diurnal effects tend to be averaged out.

It has not been possible to test the sunspot number of dependence of the cdf's from the data available. However, Fremouw et al. (1977) have suggested a linear variation of the mean scintillation index with R_Z. This suggests that the effect of sunspot activity is accounted for in the effect on the mean level of S_0.

DISCUSSION

The above mentioned observations suggest that if the mean diurnal variation of S_0 at a given frequency can be predicted (e.g., Fremouw et al. 1977 Trivedi et al. 1979) the curves of Figs. 1-3 can be used to derive an approximate prediction for the cdf of scintillation either directly or by suitable extrapolation of mean S_0 curves and interpolation in frequency provided that the frequency does not differ greatly from the frequencies for which the curves have been derived. Similar analyses at intermediate frequencies when available may of course alleviate this drawback.

The procedure suggested from the mean diurnal variation of S_0, the percentage of 24 hours during which the mean S_0 falls within various 2.5% ranges can be calculated. The cdf curve corresponding to the upper limit of a given range (i.e., worst case consideration) can then be multiplied by the percentage to yield the partial percentage of scintillation at various S_0 levels. This percentage is that contributed by the group of S_0 values in that range. Similar partial percentages may be derived corresponding to each of the predicted ranges of S_0 using appropriate percentage diurnal occurrence. A summation of these partial percentages corresponding to a given S_0 level yields the total percentage of time that level of S_0 is likely to be exceeded. Such summations at various levels can be used to generate the overall expected cdf.

The errors introduced by the various averaging processes, arbitrary definition of ranges, the use of long-term mean S_0 values, etc. in cdf prediction have been estimated by applying the above mentioned procedure to the observed mean diurnal variation of S_0. The cdf's obtained by this process are compared with the true observed cdf's in Fig. 6. It may be seen that the agreement is good, and that the estimated cdf in all cases gives a fractionally higher estimate of the percentage occurrence of scintillation at a given level.

While the above test is not on purely independent data it emphasizes the validity of the procedure and its success on application to independent data depends upon the constancy of the individual curves. The present results lead one to expect such a constancy, but clearly further verification is needed.

REFERENCES

W. J., C. L. Rino, A. R. Hessing and V. E. Hatfield, A Transionospheric Communication Channel ., SRI Quarterly Tech. Rep. /, Contract F30602-75-C-0236, July 1977.

Trivedi A. I., Hari Om Vats, R. G. Rastogi and M. R. Deshpande, Low Latitude Aeronomical Processes (Advances in Space Exploration). Cospar Symposium Series, Vol. 8, Ed. A. P. Mitra, Pergamon Press, Oxford and New York, pp. 213, 1979.

Whitney H. E., J. Aarons and C. Malik, Planet. Space Sci. 17, 1069, 1969.

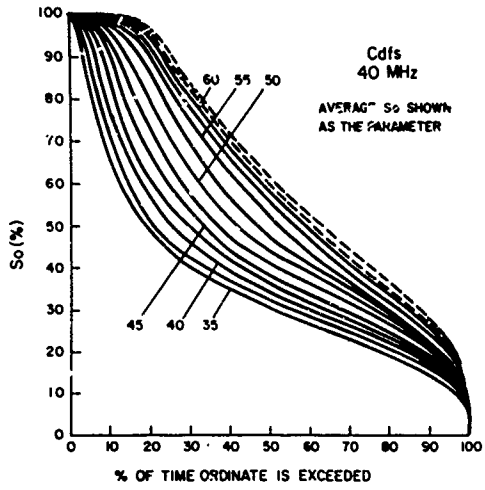


Fig. 1. Percentage of time a given S_o value is exceeded at 40 MHz when average S_o has a value within ±2.5 of the value shown as parameter on these curves.

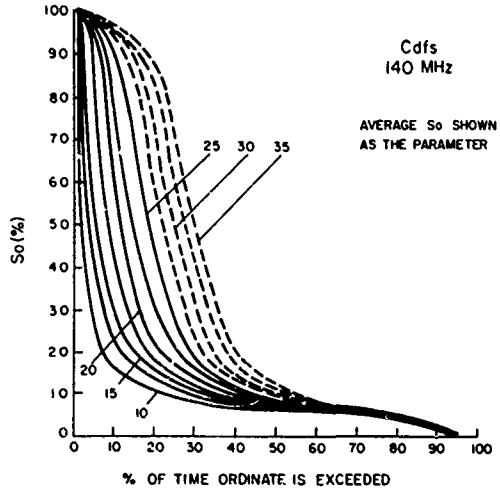


Fig. 2. Percentage of time a given S_o value is exceeded at 140 MHz when average S_o has a value within ±2.5 of the value shown as parameter on these curves.

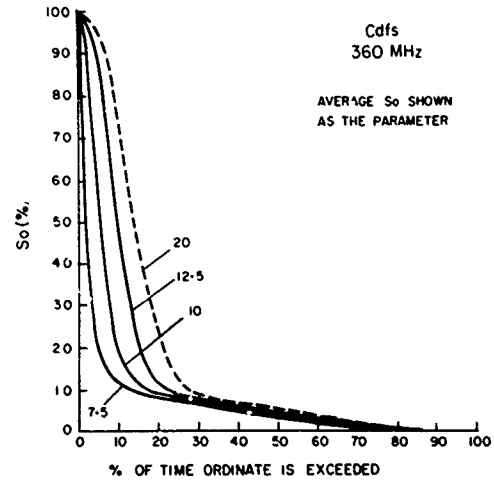


Fig. 3. Percentage of time a given S_o value is exceeded at 360 MHz when average S_o has a value within ±2.5 of the value shown as parameter on these curves.

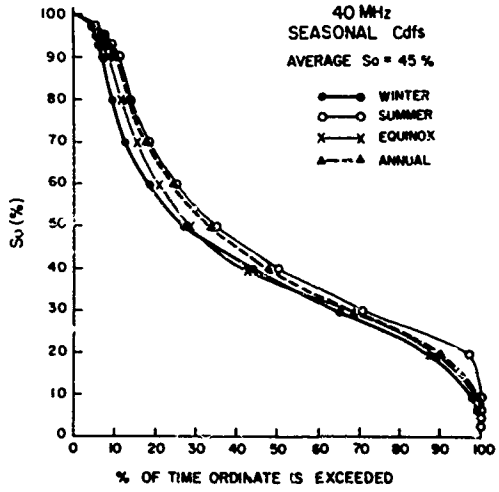


Fig. 4. Distribution of S_o observed during various seasons at 40 MHz. Each distribution has a mean S_o value of 45%.

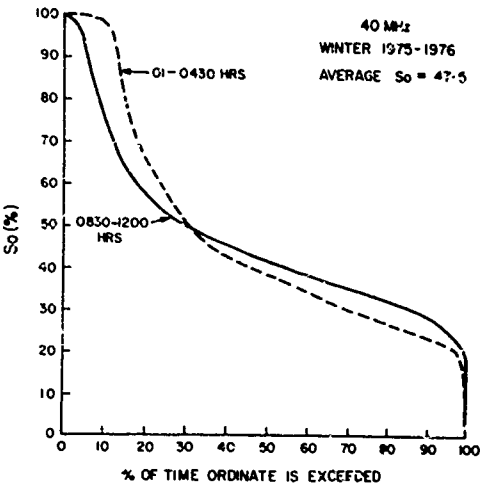


Fig. 5. Comparison of E and F layer scintillation distributions seen at 40 MHz during winter months. Both distributions have identical mean value of S_o as 47.5%.

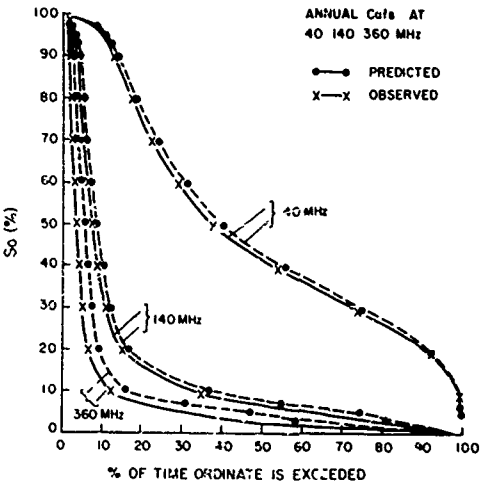


Fig. 6. A comparison of the observed cdf's at 40, 140 and 360 MHz with the predicted cdf's derived from Fig. 1, 2 and 3.

Predicting the Equatorial Spread F/Scintillation Environment by Theoretical and Numerical Simulation Techniques

S. T. Zalesak[+], S. L. Ossakow[+], and P. K. Chaturvedi[*]

[+]Geophysical & Plasma Dynamics Branch, Plasma Physics Division
Naval Research Laboratory, Washington, D.C. 20375

[*]Berkeley Research Associates, Springfield, VA 22150

ABSTRACT

The understanding of equatorial spread F (ESF) phenomena has increased significantly during the past several years. Associated ESF phenomena involve ionospheric irregularities spanning some 5-6 orders of magnitude in scale size. The largest scale sizes (~ 100 meters to many kilometers), which cause scintillation phenomena, are caused by a plasma fluid type Rayleigh-Taylor (R-T) instability mechanism on the bottomside of the nighttime equatorial F region. Plasma density bubbles (depletions) are formed on the bottomside by these ionospheric irregularities and then rise nonlinearly to the topside, by $\underline{E} \times \underline{B}$ polarization motion, causing irregularities above the F peak. Our previous work has focused on the pure gravitationally-driven R-T instability in equatorial F region geometry, i.e., a two dimensional isolated equatorial F region plane with both altitude and east-west extent. Here we present the results of recent analysis and numerical simulations of the effect of a background E region of finite Pedersen conductivity coupled with an eastward neutral wind at the equator. The results show the effect of an eastward neutral wind in causing a westward tilt in ESF large scale bubbles. In addition, the results offer an explanation for the "fishtails," C's, and westward tilts observed on equatorial radar backscatter maps of ESF plumes, i.e., small scale size irregularities. The numerical simulation results also indicate the presence of secondary scintillation causing instability mechanisms acting on the westward walls of bubbles at low altitudes and on the eastward walls at higher altitudes.

INTRODUCTION

In our previous studies of evolving equatorial spread F (ESF) bubbles and plumes in the equatorial ionosphere [Scannapieco and Ossakow, 1976; Ossakow et al., 1979; Zalesak and Ossakow, 1980] we focused our attention on showing that the motion and structure of the experimentally measured ESF environment (bottomside and topside spread F, bubble formation, and evolution) could be explained in terms of the nonlinear evolution of the gravitationally driven collisional Rayleigh-Taylor (R-T) instability. Through the use of numerical simulation techniques, we were able to demonstrate that after its initial linear growth phase, the R-T instability on the bottomside of the nighttime equatorial F region evolves nonlinearly into bubbles or plumes of low density plasma (depletions) which rise by $\underline{E} \times \underline{B}$ polarization motion to well beyond the F peak. Thus it was possible to understand the previously inexplicable existence of both bottomside and topside ESF. These studies showed both qualitative and quantitative agreement with observations [Kelley et al., 1976; Woodman and LaHoz, 1976; McClure et al., 1977] in terms of time scales, depletion levels, and plume morphology.

In this present study we would like to address the influence of an eastward neutral wind at the equator, combined with the presence of a background E region of finite Pedersen conductivity, coupled electrically to the equatorial F region plasma along magnetic field lines. We shall find that, under some simple assumptions, this situation results in a shear in the unperturbed equatorial F region plasma motion, even when no shear exists in the F region neutral wind. This shear is such as to bend passive vertically aligned structures into "C" shaped structures. A numerical simulation is presented in which it is verified that a rising ESF bubble is indeed caught up in this shear, as it rises, causing at late times the "C"'s, "fishtails" and westward tilts seen on backscatter radar plumes [Woodman and LaHoz, 1976; Tsunoda, 1981] and the westward motion of large scale bubbles [McClure et al., 1977].

THEORY AND MODEL

In Fig. 1, we show the geometry of the physical phenomenon we are attempting to model. The equatorial F region plasma responds to the effects of the earth's magnetic field, gravity, collisions with the neutral atmosphere, and electric fields. Since the conductivity along magnetic field lines is extremely high, these electric fields can depend on the dynamics of plasma far from the equatorial region, but connected to the equatorial region by magnetic field lines. We find that the physical quantity dominating the evolution of the collisional Rayleigh-Taylor instability is the magnetic field line integrated Pedersen conductivity, and that the primary contribution to that quantity comes from plasma in the local region near the "computational plane" shown in Fig. 1. This fact has been the basis for our previous theoretical and numerical studies of equatorial spread F [Scannapieco and Ossakow, 1976; Ossakow et al., 1979; Zalesak and Ossakow, 1980], and has enabled us to study the phenomena of interest using just a single two-dimensional computational plane.

We do not propose here to analyze the problem in the complete three-dimensional geometry, but rather, as a first step, to modify our two dimensional model to take into account the presence of other plasma, and hence Pedersen conductivities and forces, in regions far from the equatorial plane, but connected to the equatorial F region plasma along magnetic field lines. For instance this could be the northern and southern hemisphere E region plasma shown in Fig. 1. This modification is shown in Fig. 2, where we show three distinct layers of plasma connected by magnetic field lines. The center layer is the same compututational plane as we have used in our previous work [Scannapieco and Ossakow, 1976; Ossakow et al., 1979; Zalesak and Ossakow, 1980], and represents the equatorial nighttime F region plasma. The upper and lower layers represent the remaining northern and southern hemisphere plasma respectively, including the E region plasma. The problem is still essentially two dimensional in that we do not allow transport of ions between layers (nor do we allow for any physical variable to depend explicitly on z, the direction along the magnetic field). We do, however, allow electron currents to flow along magnetic field lines between layers to preserve electrical neutrality. Also, within the context of this model, we will finally take the E region layers to act as a passive load, i.e., we do not allow for any motion in layers 1 and 3 and those layers are assumed to remain uniform. Thus, as a first cut we are taking our previous equatorial plane simulations [Scannapieco and Ossakow, 1976; Ossakow et al., 1979; Zalesak and Ossakow, 1980] and adding a passive E region load to the circuit to allow for short circuiting effects. Under the assumptions that (i) the electric fields of interest are electrostatic and, hence, derivable from a scalar potential; and (ii) that the conductivity along magnetic field lines is extremely large and, hence, the potential is constant along a field line, we are left with a problem similar to the multilevel barium cloud striation problem [Lloyd and Haerendel, 1973; Scannapieco et al., 1976; Doles et al., 1976]. We will now briefly review the multilevel equations appropriate to our ESF problem.

We will not derive the equations here; rather we shall state the governing equations and list the assumptions made in deriving them. The equations to be solved are

$$\frac{\partial n}{\partial t} + \nabla_\perp \cdot (n \underline{v}_i) = 0 \tag{1}$$

$$\nabla_\perp \cdot [(\Sigma_1 + \Sigma_2 + \Sigma_3)\nabla_\perp \phi] = \frac{\partial}{\partial x}\left[\Sigma_2 \frac{Bg}{\nu_{in} c}\right] - \frac{\partial}{\partial y}\left[\Sigma_2 \frac{B U_n}{c}\right] \tag{2}$$

$$\underline{v}_i = \frac{\nu_{in}}{\Omega_i} \frac{c}{eB} (e \underline{E} + m_i g \hat{y} + \nu_{in} m_i U_n \hat{x}) + \frac{c}{eB} (e \underline{E} + m_i g \hat{y} + \nu_{in} m_i U_n \hat{x}) \times \hat{z} \tag{3}$$

$$\underline{E} = - \nabla_\perp \phi(x,y) \tag{4}$$

$$\nabla_\perp = \hat{x} \frac{\partial}{\partial x} + \hat{y} \frac{\partial}{\partial y}, \quad \hat{z} = \underline{B}/|B| \tag{5}$$

where n, $\underline{v}_i$, e, m_i, and ν_{in} are the ion number density, velocity, charge, mass, and collision frequency with neutral species respectively, U_n is the westward neutral wind velocity, and the $\hat{x}$, $\hat{y}$, and $\hat{z}$ directions refer to Fig. 2. Here Σ_1, Σ_2 and Σ_3 refer to the magnetic field line integrated Pedersen conductivities in layers 1, 2, and 3 respectively. All other quantities in (1) through (3) refer to layer 2 (see Fig. 2). The earth's magnetic field B is taken to be 0.5 gauss, g is - 980 cm/sec^2, and c is the speed of light.

Equations (1) through (3) were derived from the plasma fluid ion and electron continuity and momentum equations in each layer, subject to the following assumptions: 1) Quasineutrality, i.e., $n_i \approx n_e \equiv n$ everywhere; 2) the electric field is electrostatic in nature and hence derivable from a potential (Eq. 4); 3) the conductivity along magnetic field lines is sufficiently high that the field lines are in fact equipotentials, i.e., $\phi = \phi(x,y)$; 4) the slight convergence of the magnetic field lines with latitude is neglected; 5) the intertial terms (see Ossakow et al., 1979) in the

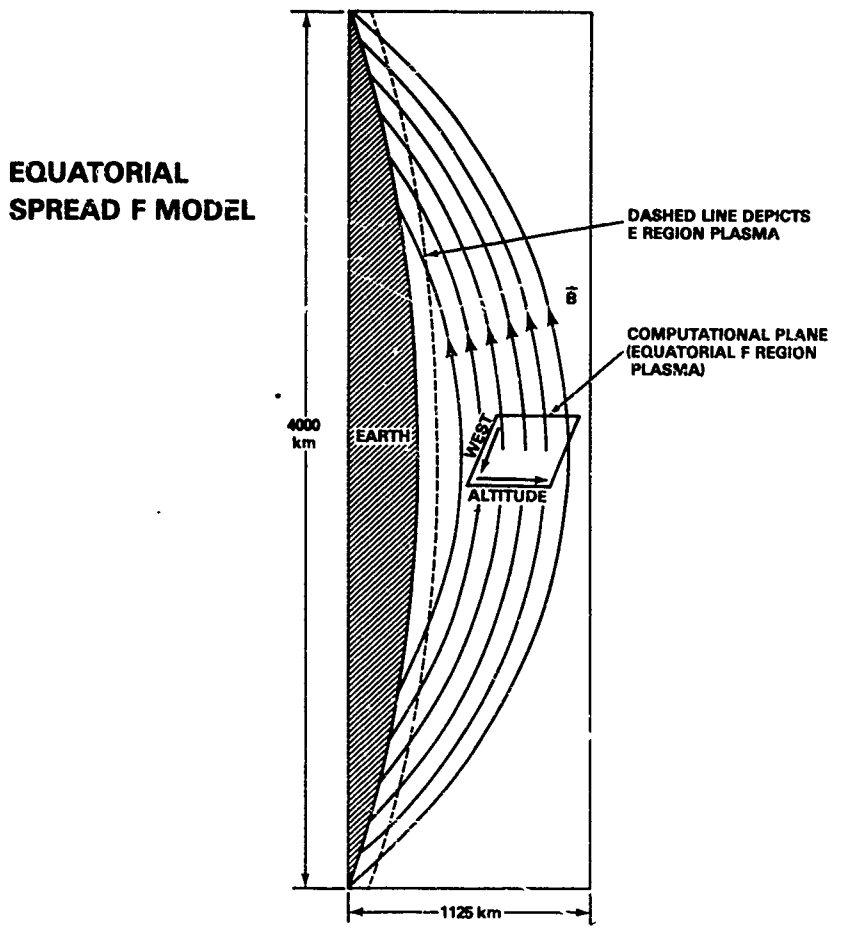


Figure 1. Diagram of the equatorial ionosphere and of the neighboring regions which have physical relevance to equatorial spread F (ESF) processes, including the E region plasma at higher and lower latitudes. These regions are electrically coupled to the equatorial F region ionosphere by the high conductivity along magnetic field lines. Plasma is actually distributed all along these field lines, but in this study we shall make the assumption that this system can be modeled accurately by three planes of plasma connected by straight field lines, as shown in Figure 2. One of these three layers (layer 2 in Figure 2) is shown here as the "computational plane".

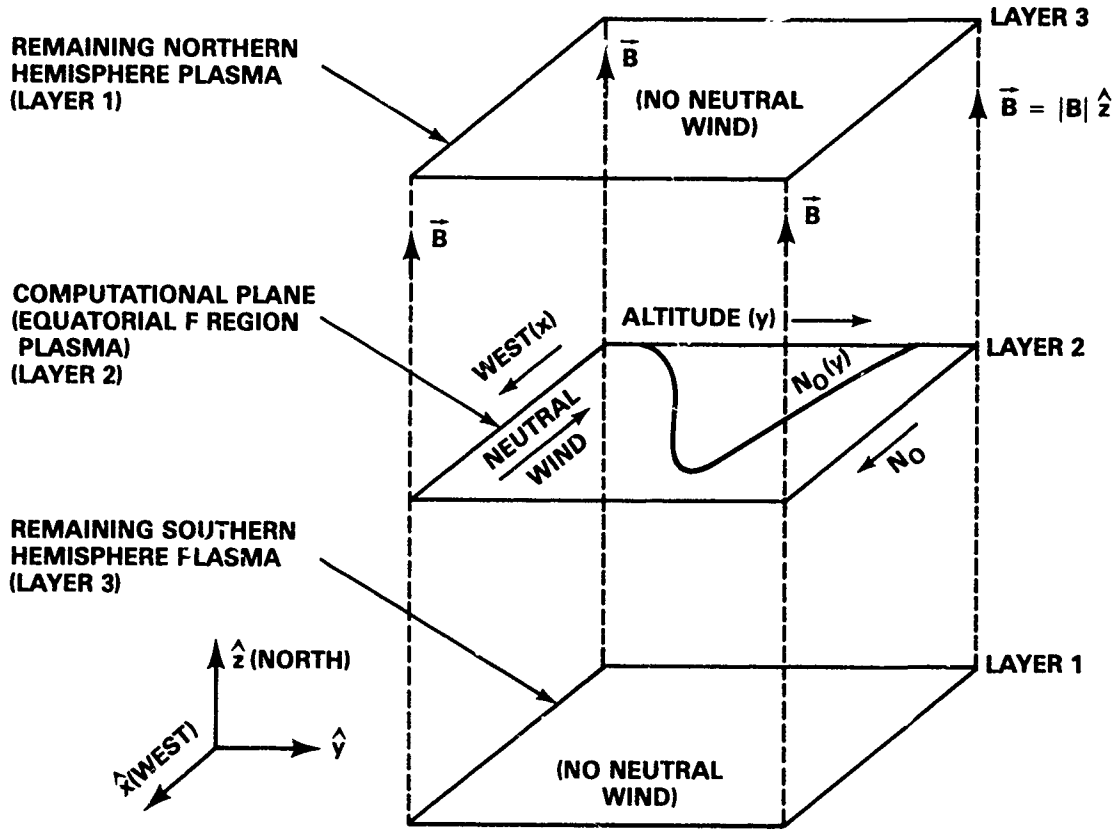


Figure 2. The "three layer" model of the physical system depicted in Figure 1. All plasma in the vicinity of the equatorial plane has been compressed into layer 2, while the remaining northern and southern hemisphere plasma has been compressed into layers 1 and 3 respectively. Further, the magnetic field lines have been straightened so we can deal in cartesian coordinates x, y, and z as shown in the figure. The plasma in layers 1 and 3 is assumed to be uniform and free of any external driving force such as a neutral wind. The equatorial layer 2 is assigned a realistic initial distribution of electron density $N_o(y)$, and ion-neutral collision frequency, along with a neutral wind which may vary with altitude, but which is taken to be uniform and eastward, and equal to 150 m/sec in this study. In addition, gravity points in the negative y direction.

ion and electron momentum equations can be neglected; 6) all collisional effects except those of ions on neutrals are neglected; 7) terms of order ν_{in}^2/Ω_i^2 are neglected; 8) the plasma compressibility associated with Pedersen mobility in the E region (layers 1 and 3) is small enough so that we may treat Σ_1 and Σ_2 as constants during the course of our simulations; and 9) there are no neutral winds or external electric fields in layers 1 and 3.

Equations (1) and (2) constitute the system of equations we must solve. In general it will be necessary to resort to numerical means for this task, but for the case of an unperturbed laminar ionosphere it is both possible and useful to find a simple analytic solution to the plasma flow field, which is an illuminating example.

Suppose Σ_1, Σ_2, Σ_3, are functions only of y (altitude in the equatorial plane). Then for any set of boundary conditions on ϕ which does not itself impose an x-dependence on ϕ, we find that $\phi = \phi(y)$. Then (2) becomes

$$\frac{\partial}{\partial y}\left((\Sigma_1 + \Sigma_2 + \Sigma_3)\frac{\partial \phi}{\partial y}\right) = -\frac{\partial}{\partial y}\ \Sigma_2 \Omega_i \frac{m_i}{e} U_n \tag{6}$$

the general solution of which is

$$(\Sigma_1 + \Sigma_2 + \Sigma_3)\frac{\partial \phi}{\partial y} = -\Sigma_2 \Omega_i \frac{m_i}{e} U_n + J_{oy} \tag{7}$$

where J_{oy} is a constant, and $\Omega_i = eB/(m_i c)$.

Assuming that $\Sigma_2 \to 0$ as $y \to \pm \infty$ and demanding that $\partial\phi/\partial y$ (or equivalently the total current) vanish at $y = \pm \infty$ we get $J_{oy} = 0$. Recalling that $\partial\phi/\partial y = -E_y$ we obtain

$$E_y = \frac{\Sigma_2}{\Sigma_1 + \Sigma_2 + \Sigma_3}\ \Omega_i \frac{m_i}{e} U_n \tag{8}$$

The $\underline{E} \times \underline{B}$ plasma motion produced by this electric field is given by

$$v_x = \frac{cE_y}{B} = \frac{\Sigma_2}{\Sigma_1 + \Sigma_2 + \Sigma_3}\frac{c}{B}\Omega_i \frac{m_i}{e} U_n \tag{9}$$

$$v_x = \frac{\Sigma_2}{\Sigma_1 + \Sigma_2 + \Sigma_3}\ U_n = f\, U_n$$

where

$$f = \Sigma_2/(\Sigma_1 + \Sigma_2 + \Sigma_3) \tag{10}$$

Note that the plasma drifts at a fraction f of the neutral wind velocity, and that the fraction is simply the ratio of the "local" equatorial plane Pedersen conductivity to the total field line conductivity on a given field line (Note: what we have in mind here and in the numerical simulations is that our magnetic field line integration for the equatorial F region is over a very limited extent in z, i.e., almost a delta function plane (a very thin layer for the F region), and that regions 1 and 3 constitute the rest of the field line connected ionosphere as a load on the circuit). This simple equation has some remarkable consequences in terms of the motion of structures (spread-F plumes, for example) imbedded in the equatorial ionosphere. Suppose that Σ_2 is a function of altitude with a peak Σ_2^{max} at altitude h_{max}. Suppose further that Σ_1 and Σ_3 are constants such that $\Sigma_1 + \Sigma_3 = 0.1\ \Sigma_2^{max}$, and that we impose a uniform eastward neutral wind of 100 m/sec on level 2 (the equatorial plane). We now create a model ionosphere (see Table 1) and tabulate the eastward plasma velocity as a function of altitude:

Table 1

Altitude (km)	Σ_2/Σ_2^{max}	Eastward plasma velocity (m/sec)
600	0.1	50
500	0.5	83
400 (h_{max})	1.0	91
300	0.1	50
200	0.01	9

Note that even though there is no vertical shear in the neutral wind velocity, the plasma flow field contains a large shear with opposing signs on either side of h_{max}. The effect of this shear is to bend any passive vertical structure imbedded in this flow field into a "C" shape.

The above result is quite satisfying in that it offers a qualitative explanation of the "C"'s, "fishtails", and other tilted structures seen by Woodman and LaHoz [1976] and Tsunoda [1981] in their observations of coherent radar backscatter from the meter-scale size irregularities associated with ESF plumes. However, the above analysis is valid only for passive structures imbedded in a laminar unperturbed ionosphere, conditions which are simply not met in the ESF environment. Numerical simulations are necessary to prove the case unequivocally.

NUMERICAL SIMULATION RESULTS

Equation (1) is solved numerically using the fully multidimensional flux-corrected transport (FCT) techniques of Zalesak [1979]. Briefly, FCT is a technique originally developed by Boris and Book [1973] for solving equations of the form (1) where steep gradients in n are expected to form. The fluxes used in the algorithm are nonlinear weighted averages of fluxes computed by high and low order finite differences. The high order fluxes are weighted as heavily as possible subject to the constraint that nonphysical oscillations are not introduced. Equation (2) is solved using the direct elliptic solver of Madala [1978].

The numerical calculation to be presented was performed on a two-dimensional cartesian mesh using 40 points in the x (east-west) direction and 140 points in the y (vertical) direction. The (uniform) grid spacing was 3 km in the y direction, and 5 km in the x direction. The bottom of the grid corresponds to 253 km altitude and the top of the grid to 676 km altitude. Periodic boundary conditions were imposed on both n and ϕ in the x direction. In the y direction transmissive boundary conditions were imposed on n ($\partial n/\partial y = 0$) and Neumann boundary conditions were imposed on ϕ. Specifically, the normal derivative of ϕ was chosen such that the normal component of the total current (the sum over all three layers) was zero for the unperturbed state. This implies

$$(\Sigma_1^\circ + \Sigma_2^\circ + \Sigma_3^\circ)\frac{\partial \phi}{\partial y} + \left[\Sigma^\circ \Omega_i \frac{m_i}{e} U_n\right]_2 = 0 \qquad (11)$$

at the upper and lower boundary, where Σ° is the Pedersen conductivity of the initial unperturbed state.

The plots which will be presented are contours of constant n(x,y,t). Superimposed on each contour plot is a dashed line depicting $n_o(y)$, the initial unperturbed electron density profile in layer 2, for reference purposes. Our $n_o(y)$ profile is such that the F_2 peak is located at 434 km altitude, and the minimum electron density scale length $L = n_o(\partial n_o/\partial y)^{-1}$ is 10 km. The ion-neutral collision frequency $\nu_{in}(y)$ used in the calculation can be found in Ossakow et al. [1979]. The initial perturbation used to start each calculation was a mode 1 sine wave in the x direction:

$$\frac{n(x,y,0)}{n_o(y)} = 1 - e^{-3}\cos(\pi x/100) \qquad (12)$$

A uniform eastward neutral wind of 150 m/sec was imposed over the entire equatorial plane ($U_n = -150$ m/sec). In addition, it was assumed that $(\Sigma_1 + \Sigma_3) = 0.12\,\Sigma_2^{max}$. In Fig. 3 we show the isodensity contours of n(x,y) for our initial conditions (laminar ionosphere $n_o(y)$ plus perturbation (12)). The contours are labeled for later reference purposes. Figure 4 shows isodensity contours of n(x,y) at four different times during the simulation. The 40 x 140 mesh has been extended to 80 x 140 for plotting purposes (recall we have periodic boundary conditions in the east-west direction), and we have placed ourselves in a reference frame moving eastward at 68 m/sec (to minimize computer time). We see the usual plume formation and subsequent rise, but we also see the effects of the plasma shear associated with the laminar ionosphere: the plumes are bent into "C" shaped structures. That is, we see a westward tilt of the plume with altitude above ~ 360 km, and an eastward tilt with altitude below ~ 360 km. The shapes depicted in Fig. 4 are similar to those seen on radar backscatter maps of small scale (≤ 3m) irregularities [Woodman and LaHoz, 1976; Tsunoda, 1981]. Presumably the radar backscatter maps are signatures of the large scale size bubbles depicted in Fig. 4, i.e., the steep plasma density gradients associated with the bubbles in Fig. 4 drive the radar backscatter observed irregularities. The westward and upward motion of the bubbles depicted in Fig. 4 are in agreement with the satellite in situ measurements of McClure et al. [1977]. Small scale rippling on the high altitude eastward wall of the bubble is also exhibited in Fig. 4.

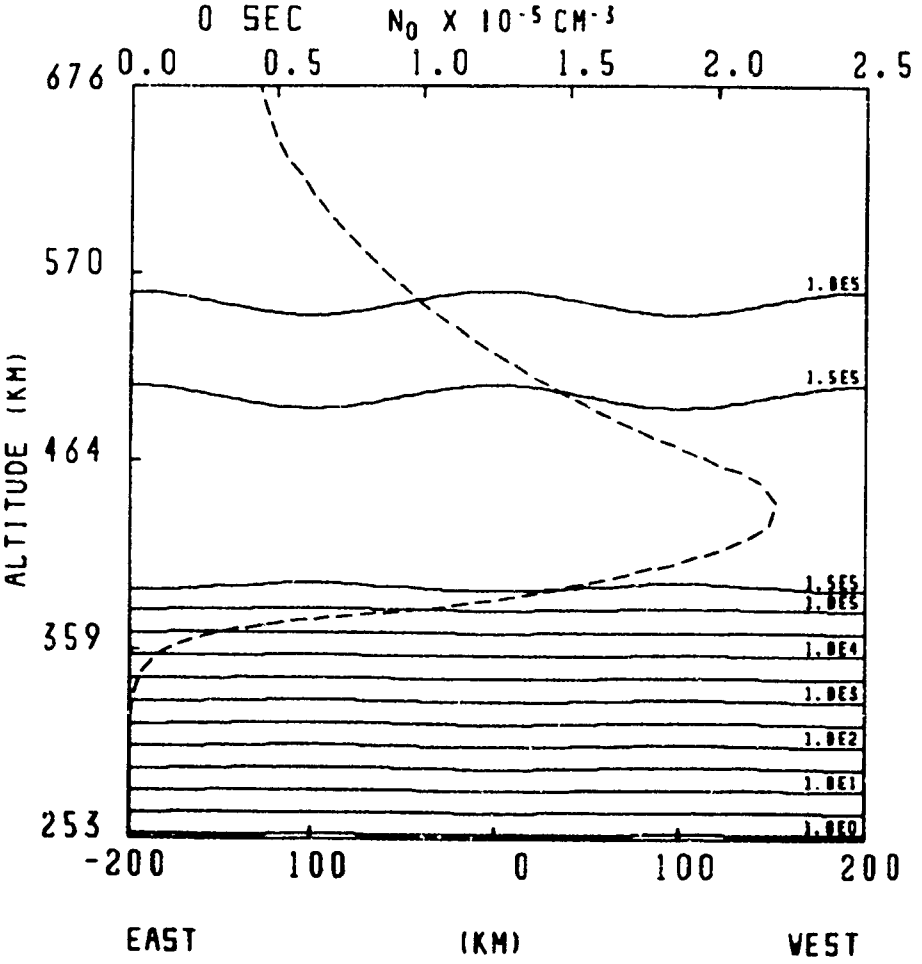


Figure 3. Iso-electron density contours for the initial perturbed state in layer 2. This represents the initial conditions for our numerical simulation. The contours are labeled in units of electrons/cm^3 in E format notation (1.0E1 = 1 x 10^1, etc.). The unperturbed ionosphere was initially laminar (independent of x, the east-west direction) and is exhibited by the dashed curve showing $N_0(y)$, at any point in the east-west (x) direction. This curve is labeled at the top of the figure. The perturbation has a maximum amplitude of e^{-3} in relative electron density, is a pure mode 1 sine wave in x, and is independent of altitude y, as described in the text. The observer is looking southward so that $\underline{B}$ is out of the figure.

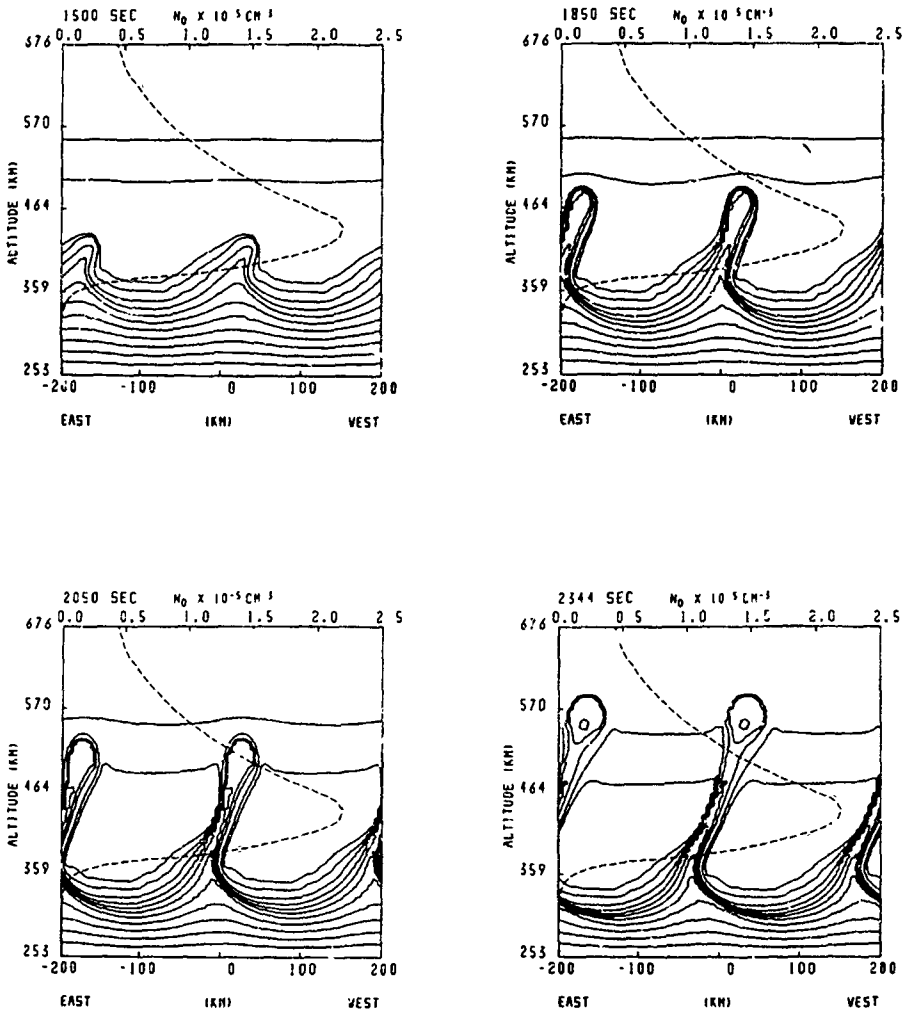


Figure 4. Sequence of four iso-electron density contours for our numerical simulation, using the initial conditions depicted in Figure 3. Times shown are 1500, 1850, 2050, and 2344 seconds. Note that the plasma shear associated with the initial unperturbed laminar ionosphere (see text) superimposes itself on the normally vertical plume rise, resulting in westward tilts of the plume with altitude above ~ 360 km, and eastward tilts with altitude below 360 km, in agreement with experimental observations. Also note the plasma gradient steepening and apparent onset of secondary instabilities on the east wall of the plume at the higher altitudes, and on the west wall at lower altitudes.

CONCLUSIONS AND DISCUSSION

It would appear from the analysis and results we have presented here that we have a satisfactory explanation for the westward tilts, C-shapes and "fishtails" seen on backscatter radar maps of ESF irregularities [Woodman and LaHoz, 1976; Tsunoda, 1981] and the westward and upward movement of large scale plasma depletions (bubbles) as observed by satellite in situ measurements [McClure et al., 1977]. Our plans for the future include improving the model to allow for a more realistic treatment of the E regions: i.e., solving the ion continuity equation numerically and allowing driving terms such as tidal E region neutral winds. There is also to be studied the occurrence of secondary instabilities on the westward and eastward walls of the bubbles as a result of the neutral wind and of the deformation of the plume by the plasma wind shear. For example, in Fig. 4 we see that at the latest time shown (2344 sec) we have a steepening of plasma gradients and the apparent onset of instability on the east wall of the bubble at the higher altitudes and on the west wall of the bubble at lower altitudes. Presumably if we were able to carry the calculations later in time with increased resolution (recall that in the present simulation the minimum grid size is 3 km) in these regions, we would see the further structuring of these steepened regions which would result in scintillation causing irregularities.

It should also be pointed out that our previous numerical simulation results [Ossakow, et al., 1979] on rising ESF bubbles (without a neutral wind and E region) have been utilized in model computations of radio wave scintillation caused by equatorial ionospheric bubbles [Wernik et al., 1980]. It is anticipated that the results presented in this paper will also be of value in such scintillation modeling.

ACKNOWLEDGEMENT

This work was supported by the Defense Nuclear Agency and the Office of Naval Research.

REFERENCES

Boris, J. P., and D. L. Book, Flux-corrected transport, I, Shasta, a transport algorithm that works, J. Comp. Phys., 11, 38, 1973.

Doles, J. H., III, N. J. Zabusky, and F. W. Perkins, Deformation and striation of plasma clouds in the ionosphere, 3, Numerical simulations of a multilevel model with recombination chemistry, J. Geophys. Res., 81, 5987, 1976.

Kelley, M. C., G. Haerendel, H. Kappler, A. Valenzuela, B. B. Balsley, D. A. Carter, W. L. Ecklund, C. W. Carlson, B. Hausler, and R. Torbert, Evidence for a Rayleigh-Taylor type instability and upwelling of depleted density regions during equatorial spread F, Geophys. Res. Lett., 3, 448, 1976.

Lloyd, K. H., and G. Haerendel, Numerical modeling of the drift and deformation of ionospheric plasma clouds and of their interaction with other layers of the ionosphere, J. Geophys. Res., 78, 7389, 1973.

Madala, R. V., An efficient direct solver for separable and non-separable elliptic equations, Monthly Weather Rev., 106, 1735, 1978.

McClure, J. P., W. B. Hanson, and J. H. Hoffman, Plasma bubbles and irregularities in the equatorial ionosphere, J. Geophys. Res., 82, 2650, 1977.

Ossakow, S. L., S. T. Zalesak, B. E. McDonald, and P. K. Chaturvedi, Nonlinear equatorial spread F: Dependence on altitude of the F peak and bottomside background electron density gradient scale length, J. Geophys. Res., 84, 17, 1979.

Scannapieco, A. J., and S. L. Ossakow, Nonlinear equatorial spread-F, Geophys. Res. Lett., 3, 451, 1976.

Scannapieco, A. J., S. L. Ossakow, S. R. Goldman, and J. M. Pierre, Plasma cloud late time striation spectra, J. Geophys. Res., 81, 6037, 1976.

Tsunoda, R. T., Time evolution and dynamics of equatorial backscatter plumes, 1., Growth phase, J. Geophys. Res., 86, 139, 1981.

Wernik, A. W., C. H. Liu, and K. C. Yeh, Model computations of radio wave scintillation caused by equatorial ionospheric bubbles, Radio Sci., 15, 559, 1980.

Woodman, R. F., and C. LaHoz, Radar observations of F region equatorial irregularities, J. Geophys. Res., 81, 5447, 1976.

Zalesak, S. T., Fully multidimensional flux-corrected transport algorithms for fluids, J. Comput. Phys., 31, 335, 1979.

Zalesak, S. T., and S. L. Ossakow, Nonlinear equatorial spread F: Spatially large bubbles resulting from large horizontal scale initial perturbations, J. Geophys. Res., 85, 2131, 1980.

THE STP/S3-4 SATELLITE EXPERIMENT: EQUATORIAL F-REGION IRREGULARITIES

M. Singh[+], E. P. Szuszczewicz and J. C. Holmes
E.O. Hulburt Center for Space Research
Naval Research Laboratory
Washington, DC 20375

[+] Sachs/Freeman Associates, Inc.
Bowie, MD 20715
On leave from Punjabi University, Patiala, India

ABSTRACT

The S3-4 "in situ" satellite observations near the nighttime equator (2230 LT) have shown irregularity structures with depletions in plasma density up to 3 orders of magnitude and ranges in horizontal extent from less than 1 km to tens of kms. The holes have sharper gradients on the eastern boundary than across the western counterpart, and the power spectral indices in the irregularities lie between 1.9-2.5 in the intermediate wavelength domain. The east-west asymmetry of the depletions is also shown in the irregularity intensity and spectral strengths. The variations in density are more intense on the western wall of the holes suggesting a scintillation enhancement on the western boundary. In addition, the results of power spectral analyses support the role of the Rayleigh-Taylor instability in the generation of intermediate scale size irregularities during the occurrence of equatorial spread-F.

INTRODUCTION

Equatorial F-region irregularities have drawn much interest in recent years because of their undesirable effects on transionospheric communications and their cause-effect relationship with fundamental plasma instability processes. In efforts to understand the causative mechanism(s), considerable advances have been made in areas of detailed ground-based radar observations and "in situ" measurements. The radar observations (e.g., Woodman and La Hoz, 1976; Tsunoda et al. 1979) have found that meter-size irregularities primarily populate the bottomside F-region in the early evening, at later times tend to rise up and break away from their lower altitude source regime and develop structures extending up to the 700-1000 km region. These structures have come to be called "plumes".

On the macroscale (100's of meters to 10's of km) "in situ" nighttime equatorial measurements have revealed large biteouts in plasma density ranging up to three orders of magnitude (Hanson and Sanatani, 1973) and considered just as characteristic of spread-F as the much less intense (meter size) irregularities observed by radar. Later works on biteouts (Brinton et al. 1975; McClure et al. 1977) showed that ion composition inside and outside the holes can be vastly different. The molecular ions can be more abundant inside the holes than outside the holes, and the holes can vary from a few km to tens of km in the horizontal extent.

In examining potential relationships between radar plumes and ionospheric depletions, Szuszczewicz (1978) suggested that equatorial holes and spread-F were the same phenomena with small scale irregularities imbedded within the large scale depletions. He argued that a chemical volume of ion density on the bottomside (containing the signature of bottomside species) could move upward through a stationary neutral atmosphere and appear as biteouts at higher altitudes with much smaller structures (down to the meter range) populating the density gradients which bounded the macroscale depletion. This model was in concert with the numerical results of Scannapieco and Ossakow (1976) and the drift measurements of McClure et al. (1977).

To further study and definitively unfold the detailed relationships between large scale depletions, meter size irregularities and chemical transport processes, a coordinated investigation was conducted which involved simultaneous observations by radar and "in situ" rocket-borne diagnostics (Szuszczewicz et al., 1980). The combined observations have shown that:

(a) During conditions of well-developed equatorial spread-F the most intense "in situ" irregularities occurred on the bottomside F-layer gradient.

(b) Within a large scale topside F-layer depletion radar backscatter and "in situ" irregularity strengths maximized near the depletion's upper wall.

(c) Ion composition within a topside depletion provided signatures of its bottomside source domain and estimates of average maximum vertical drift velocity. For long-lived depletions, it was found that molecular-ion signatures (NO^+ and O_2^+) can be lost while bottomside levels of N^+ can be maintained when $[O^+] \approx N_e >> [NO^+] + [O_2^+]$; and finally,

(d) Large scale fluctuations of O^+ accompanied by a near-constant level of NO^+ and O_2^+ on the bottomside F-layer gradient suggests that neutral atmospheric turbulence was not a major source for bottomside ionospheric plasma irregularities and the associated triggering of equatorial spread-F.

To complement the vertical profile information provided by the rocket observations, we present and analyze a sample of "in situ" measurements conducted on the STP/S3-4 satellite carrying a pulsed plasma probe experiment. The probe experiment employed a self consistent test for measurement integrity, while determining electron density and temperature as well as density and mean ion mass fluctuations at 5-20 meter resolution. The S3-4 experiment has been discussed by Szuszczewicz et al. (1981). In this paper, we discuss some of the recent results with emphasis on general horizontal morphology and relationships to basic instability processes and associated scintillation effects.

EXPERIMENTAL RESULTS

Plasma Depletions

The data reported here was made available by NRL-747 paired-pulsed-plasma-probes (P^4) experiment (Szuszczewicz et al. 1981) on STP satellite S3-4 in a sun-synchronous orbit at lower F-region altitudes. The satellite crossed the nighttime equator at about 2230 LT when the the frequency of occurrence of spread-F was high. One probe was biased to respond to variations in plasma electron saturation currents ($I_e \propto N_e$) and the other probe responded to the ion saturation currents ($I_i \propto N_e / \sqrt{M_i}$). Subject to the selection of a number of commandable modes of operation, either probe could be repetitively pulsed from its fixed-bias level using a special electronic procedure (Holmes and Szuszczewicz, 1975) to generate conventional Langmuir characteristics for full determinations of electron density N_e, temperature T_e and plasma potential V_∞. The different modes of operation and experimental details are discussed by Szuszczewicz et al. (1981).

Figures 1 and 2 present samples of nighttime equatorial irregularity structures as measured by the currents collected by the electron (I_e) and ion (I_i) probes. The Figures show that the holes extend from a few kms to tens of kms, with depletion levels ranging from a factor of 3 (hole A) to a factor of 500 (hole F). A cursory analysis of Figures 1 and 2 reveals three particularly interesting observations:

(a) The density gradients on the opposing sides of each hole are different. Noting that the inclination of the satellite orbit is 96.4° (traveling in the east-to-west direction with time increasing left-to-right) it can be seen that density gradients in holes C-E (Fig. 1) and F (Fig. 2) are sharp on the eastern boundary and soft on its western counterpart. In the case of hole A, B and H (Fig. 1), the density gradients are not quite as sharp in comparison with holes C-F but the average density gradients are still softer on the western boundary. For the holes L (Fig. 1) and N' and L' (Fig. 2), the difference in density gradients on the opposing sides does not constitute any specific behavior while in the hole C' (Fig. 2) the density gradient is sharp on the western boundary. These observations can be summarized by noting that the density gradients are different across the boundaries of each depletion, with a preference for the density gradients to be sharper on the eastern side.

(b) Another interesting feature of the holes is their similarities in structural morphology. Holes A, B and H are similar in their horizontal profiles. A similarity feature can also be identified in the depletion L (Figure 1) and L' (Figure 2). Both of these depletions have a center point (M and M' respectively) around which the structures look similar. At the central point (M and M'), the plasma density approaches the background (undisturbed) level suggesting that one large hole is breaking up or alternatively two smaller ones are merging.

(c) The top panel of Figures 1 and 2 displays relative density fluctuations, $\delta I/I$ determined by variations about linear detrends over sliding 2.1 km intervals throughout the depletions. The fluctuations ($\delta I/I$) easily identifies the degree of disturbance. The percent variation in the holes A, B and H (Figure 1) is less than 20% in most parts of the holes, while in other holes of Figures 1 and 2 the variation can be more than 50% in some parts of the holes. In addition, the variations can be different across the opposing boundaries of the holes. We discuss this more fully below.

East-West Asymmetry

Further illustrations of F-region irregularities appear in data revs 2122 and 2123 shown in Figure 3. Rev 2123 shows four depletions (numbered 1 through 4) with the depletion level extending to more than two orders of magnitude. The rev 2122 shows three depletions (numbered 5 through 7) with depletions in density up to one order of magnitude.

The discussion is facilitated by identifying certain features in rev 2123. First, there are clearly defined regions of undisturbed background ionosphere, marked alphabetically A through D; the smoothness of the relative density and the corresponding 0% fluctuations attest to their undisturbed nature. Focussing on depletions 3 and 4, we see that the irregularity intensities are 2 to 3 times larger on the western boundary than on eastern counterparts. This same relationship is true in depletions 1 and 2, but only after a qualification that suggests that 1 and 2 are halves of a larger depletion bounded by A and B. This is supported in part by the non-existence of a quiescient ionosphere between the two. When viewed from this perspective the western boundary is approximately twice as intense in irregularity intensity as the eastern boundary. In rev 2122, the depletions 6 and 7 show the irregular intensity as more intense on the western boundary by a factor of 3-4. (On the basis of similar considerations applied to Figures 1 and 2, we note that the irregularity intensity is greater on the western boundaries of depletions D, E and H (Figure 1) and F (Figure 2), while the eastern boundary in holes A-C (Figure 1) is more intense. Based on irregularity intensity ($\delta I/I$) observations, we conclude that there is a high probability of occurrence of more intense fluctuations across the western boundary of bottomside F-region depletions.

The asymmetry in irregularity strengths and relationships to plasma instability mechanisms can be explored further through power spectral density analyses. We present in Figures 4 and 5 just such results for each of the boundaries in Figure 3 (1E and 1W refer respectively to the eastern and western boundary of depletion number 1). Though the experiment provides the density fluctuation power spectra with a maximum Nyquist frequency of 400 Hz (19 m resolution along the orbit and 3-5 m resolution perpendicular to magnetic field) the data in this figure has been decimated by a factor of 3 in order to include greater spatial domains across the depletion walls. Power spectral analyses are presented across the boundaries of each of the seven depletions with spectral indices (n, in the equation $P = P_o f^{-n}$) ranging from 1.9 to 2.5. More important however, is the spectral strengths (P_o), found to be 1.4 to 11.3 (entered as $P_1 \equiv P_o$ (west)/P_o (east) in Figures 4 and 5) times more intense on the western boundaries, with these intensities extending down to a 15 meter wavelength perpendicular to the geomagnetic field. The spectral indices support the work of Keskinen et al. (1980) which predicts the same approximate range of values for horizontal irregularity structures perpendicular to B. The east-west asymmetry in the depletions is apparent from the linear detrend of data and the spectral strengths.

We now show that the asymmetry bears on scintillation observations. To do this we note that scintillations depend upon $\Delta N (\Delta N \propto \Delta I)$ rather than $\Delta N/N$ (as calculated by $\frac{(\Delta I)}{I}$ and the power spectral analysis $P_o \propto (\frac{\Delta I}{I})^2$). For the purpose of scintillation-effect calculations, we have determined $|\Delta I|^{r.m.s.}$ in the wavelength

domain 80 m to 8 km by assuming that $P^{r.m.s.}$ in this domain is equal to $(|\Delta I|^{r.m.s.}/\bar{I})^2$. Defining P_2 as

$$P_2 = \frac{(\Delta I)_{r.m.s.}\text{(in the wavelength range 80m-8km on the west wall)}}{(\Delta I)_{r.m.s.}\text{(in the wavelength range 80m-8km on the east wall)}}$$

we find the values of P_2 ranging from 1.1 to 7.7. These results suggest enhanced scintillation effects on the western wall of the depletions, in agreement with the observations of Livingston et al. (1980). Furthermore, the radar measurements (Tsunoda, 1979) also show that the bottomside backscatter strength is often asymmetric in east-west plane with stronger backscatter from the western wall of a plume. The combined observations support a model of E-W asymmetry which allows for a neutral-wind driven instability growth rate enhancement on the western side of a rising bottomside F-region depletion. The maximum growth rate occurs on the western wall of the rising depletion, where the electron density gradient is most closely aligned with the plasma drift velocity vector in the reference frame of the neutral wind (Tsunoda, 1979; Zalesak et al. 1980).

COMMENTS AND CONCLUSIONS

The results of the high resolution S3-4 satellite experiment for equatorial F-region irregularities show large scale plasma depletions (1 km to 10's of km wide) with smaller scale irregularities superimposed (smallest detectable scale size $\gtrsim$ 10 meters perpendicular to $\bar{B}$). The depletions (which may in fact represent the bottomside upwelling process that has come to be identified with the lower F-region manifestations of spread-F) show east-west asymmetry with the irregularity intensity and spectral strengths generally more intense on the western boundary. Associated calculations over density fluctuations in the range 80 meters to 8 km suggest that scintillation effects would be similarly more intense on the western boundary. We find these observations consistent with radar (Tsunoda, 1979) and scintillation measurements (Livingston et al. 1980) as well as the recent computational work of Zalesak et al. (1980). Furthermore, we find the horizontal power spectral indices to lie between 1.9-2.5, supporting the role of Rayleigh-Taylor instability (Keskinen et al. 1980) and the recent theoretical and experimental comparison in the vertical plane (Keskinen et al. 1981).

REFERENCES

Brinton, H.C. H.G. Mayr, and G.P. Newton, "Ion composition in the nighttime equatorial F-region. Implications for chemistry and dynamics", (abstract) EOS Trans. AGU 56,1038 (1975).

Hanson, W.B. and S. Sanatani, "Large N_i gradients below equatorial F-peak", J. Geophys. Res. 78, 1167 (1973).

Holmes, J.C. and E.P. Szuszczewicz, "A versatile plasma probe", Rev. Sci. Instr. 46, 592 (1975).

Keskinen, M.J., S.L. Ossakow and P.K. Chaturvedi, "Preliminary report on numerical simultations of intermediate wavelength collisional Rayleigh-Taylor instability in equatorial spread-F", J. Geophys. Res. 85, 1775 (1980).

Keskinen, M.J., E.P. Szuszczewicz, S.L. Ossakow and J.C. Holmes, "Nonlinear theory and experimental observations of the local collisional Rayleigh-Taylor instability", J. Geophys. Res. (1981, in press).

Livingston, R.C., E. Weber and T. Bauchau "Multi-scale observations of irregularities in an isolated and decaying plasma bubble", Paper presented at International Symposium on Equatorial Aeronomy, Aguidilla, P.R., July 1980.

McClure, J.P., W.B. Hanson and J.N. Hoffman, "Plasma bubbles and irregularities in the equatorial ionosphere", J.Geophys. Res. 82, 2650 (1977).

Szuszczewicz, E.P., "Ionospheric holes and equatorial spread-F" Chemistry and transport", J. Geophys. Res. 83, 2665 (1978).

Szuszczewicz, E.P., R.T. Tsunoda, R. Narcisi, and J.C. Holmes, "Coincident radar and rocket observations of equatorial spread-F", Geophys. Res. Lett. 7, 537 (1980).

Szuszczewicz, E.P., J.C. Holmes and M. Singh, "The STP/S3-4 ionospheric irregularity satellite experiment", J. Geophys. Res. (1981, submitted).

Tsunoda, R.T., "The growth and decay of equatorial backscatter plumes", DNA report, SRI International, Nov. 1979.

Tsunoda, R.T., M.J. Baron, J. Owen and D.M. Towle, "Altair: An incoherent scatter radar for equatorial spread-F studies", Radio Sci. 14, 1979.

Woodman, R.F. and C. La Hoz, "Radar observations of F-region equatorial irregularities", J. Geophys. Res. 81, 5447 (1976).

Zalesak, S.T., S.L. Ossakow and P.K. Chaturvedi, "An explanation of westward tilts, fishtails and C's in the equatorial spread-F ionosphere", EOS Trans. AGU 61, 1059 (1980).

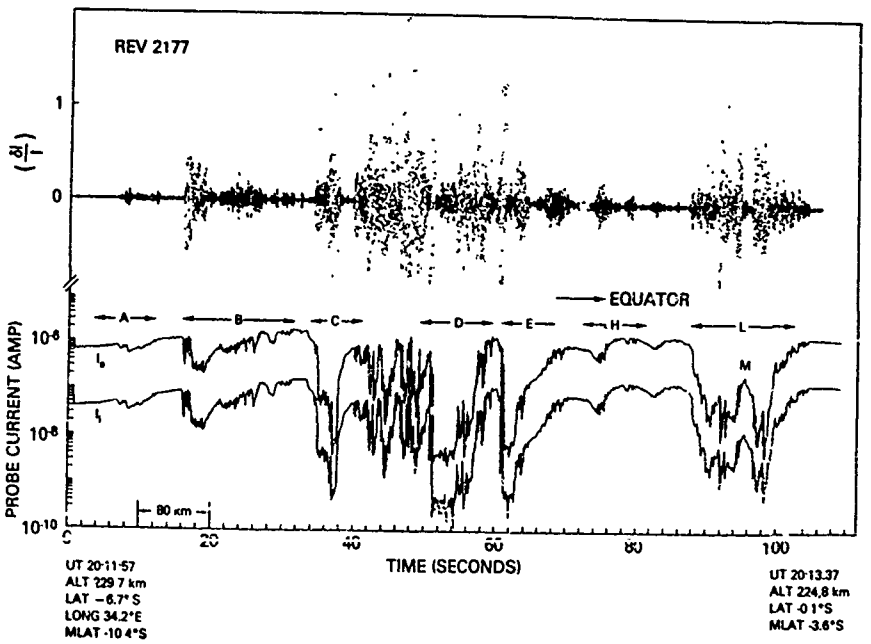


Figure 1. Horizontal profile of ionospheric F-region plasma density as indicated by electron (I_e) and ion (I_i) saturation currents measured on rev #2177. $\delta I/I$ is the relative irregularity intensity as calculated by variations about linear detrends over sliding 2.1 km intervals.

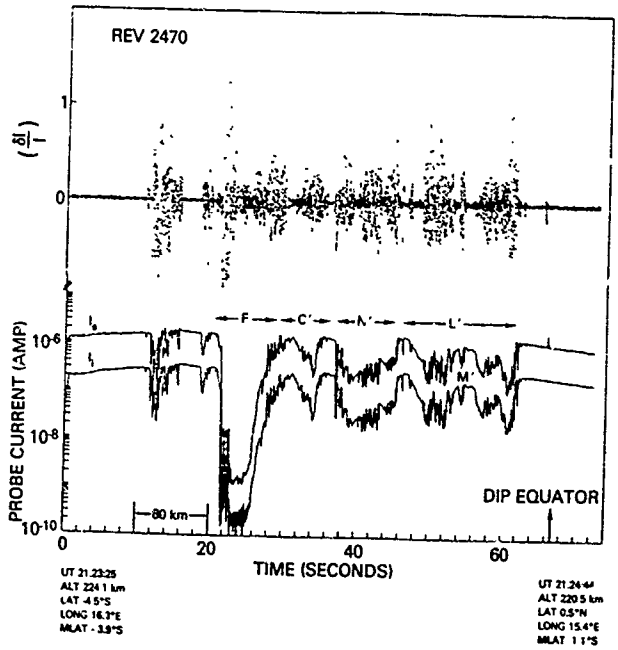


Figure 2. Horizontal profile of ionospheric F-region plasma density as indicated by electron (I_e) and ion (I_i) saturation currents measured on rev #2470. $\delta I/I$ is the relative irregularity intensity as calculated by variations about linear detrends over sliding 2.1 km intervals.

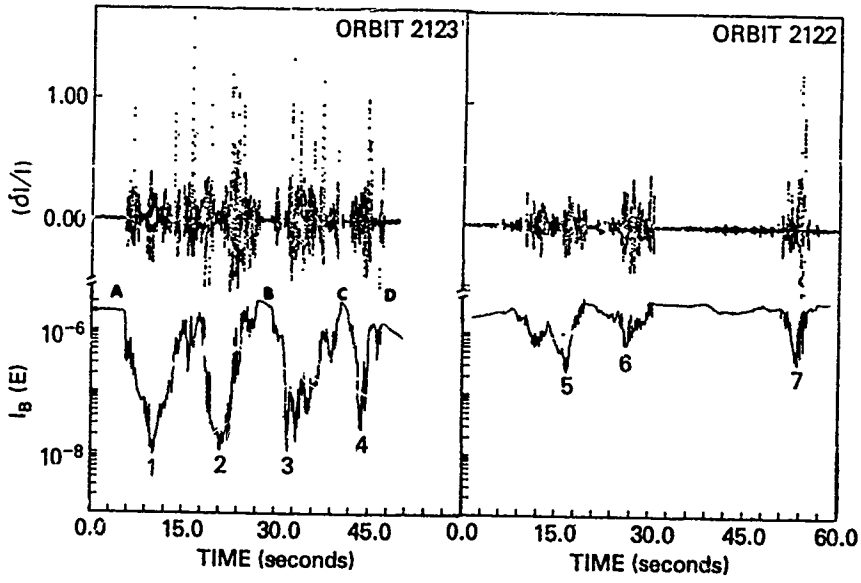


Figure 3. The electron density variation (as indicated by $I_B(E)$) for rev #2122 and 2123 along with irregularity intensity ($\delta I/I$) determined over contiguous 2.1 km intervals throughout the depletion.

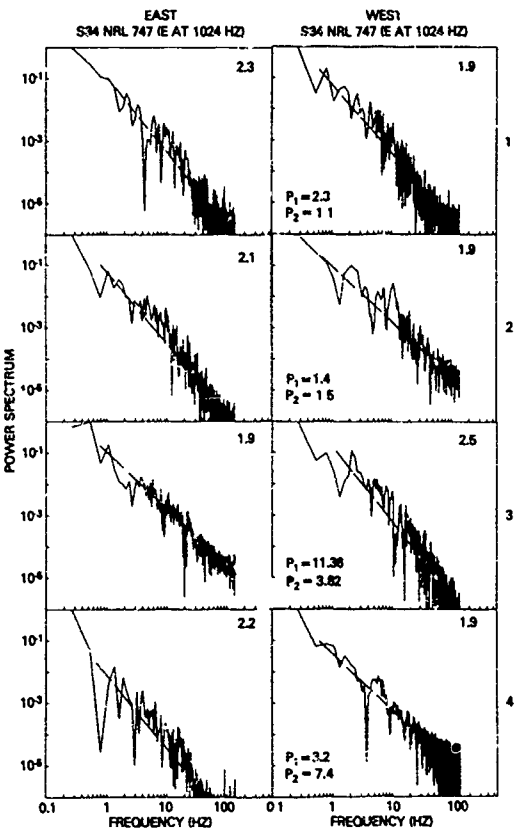


Figure 4. Power spectral analyses across the boundaries of the four depletions (1 through 4) of rev #2123 in Figure 3. The number in the upper right portion of each panel is the spectral index (n in the equation $P = P_o k^{-n}$). The value P_1 in each right panel is the ratio of spectral strengths (P_o) of western to eastern boundary $P_1 = P_o(West)/P_o(East)$. The value P_2 is the ratio of r.m.s. value of spectral strength in the wavelength domain (800 m-8 Km) of western to eastern boundary ($P_2 = P_{r.m.s.}(West)/P_{r.m.s.}(East)$).

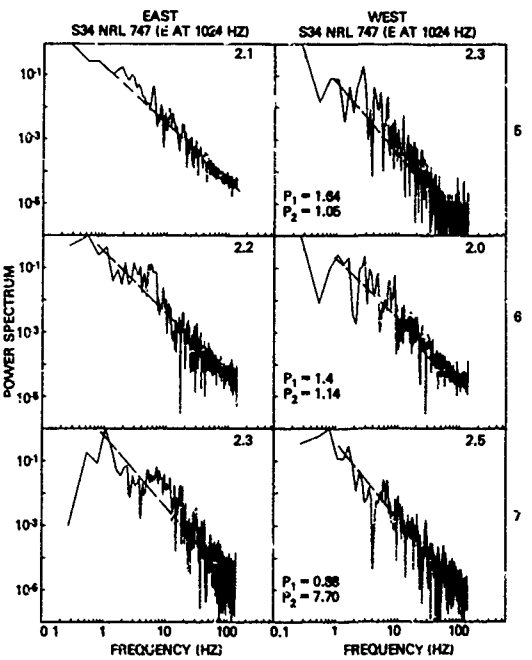


Figure 5 Same as Figure 4 for depletions 5 through 7 of rev #2122 in Figure 3.

THE GLOBAL DISTRIBUTION OF EQUATORIAL SPREAD-F IN THE TOPSIDE IONOSPHERE DURING EQUINOCTIAL PERIODS

E.R. Young
Regis College Research Center
Weston, Massachusetts 02193

W.J. Burke, F.J. Rich, R.C. Sagalyn, and M. Smiddy
Air Force Geophysics Laboratory, Space Physics Division
Hanscom Air Force Base, Massachusetts 01731

ABSTRACT

On the basis of 331 equatorial passes of DMSP we have determined that two distinct types of topside ion depletions occur in the evening sector during equinoctial periods. Firstly, during the early evening (18.9 LT) a smooth type of depletion occurs with over 80% probability of observation. This feature is aligned parallel to the magnetic equator and is probably related to the equatorial fountain effect and general F-region dynamics. Secondly, we have seen irregular depletions later in the evening (19.6h-21.9h) which are the topside signature of spread-F. Their total probability of occurrence increases from ~ 20% at 19.6h to ~ 60% at 21.9h. The association of these later depletions with a threshold of topside density and the expansion of their longitudinal distribution into regions of higher magnetic field as the night progresses indicates that spread-F occurs during conditions of a rising ionosphere and/or eastward electric field.

INTRODUCTION

In this study we present an analysis of irregular ion density "biteouts" observed in density profiles from evening equatorial passes of DMSP-F2 and F4 at 840 km. We presume, and our data supports, their identification as the topside signature of field aligned depletion regions, or "bubbles", associated with spread-F, which have percolated through the F-region peak from the bottomside. McClure et al., (1977) has observed an average vertical drift speed of 150 m/s in bubbles near the F-peak. Such irregularities have been studied using a variety of in situ and remote techniques.

Direct measurements of plasma density biteouts have been made in the bottomside (Kelley et al., 1976), near the F-region peak (McClure et al., 1977; Basu et al., 1976) and in the topside (Dyson and Benson, 1978; Burke et al, 1979 a,b, 1980), while remote measurements have been made using radar backscatter (Woodman and LaHoz, 1976) and scintillations of radio signals. It should be mentioned although such bubbles are capable of producing observed equatorial radio scintillations, it is not clear that they are the only source (Basu et al., 1976), McClure et al., (1977).

A recent review by Fejer and Kelley (1980) suggests that there is a growing consensus that spread-F begins in the bottomside of the F-layer as a collisional, Rayleigh-Taylor instability. For spread-F bubbles to percolate higher than the F-peak the Rayleigh-Taylor mechanism must evolve into the non-linear regime (Scannapeico and Osakow, 1976; Ossakow et al., 1979; Ott, 1978). Recent numerical simulations by Anderson and Haerendel (1979) and Ossakow et al., (1979) have stressed the roles of post sunset eastward electric fields, the height of the initial disturbance and the height of the F peak in determining whether bottomside depletions can rise into the topside layer.

The longitudinal variability of equatorial F-region irregularities at altitudes $\geq$ 400 km were studied by Basu et al., (1976) using measurements from the OGO 6 satellite. Their study showed that the frequency of occurrence for irregularities capable of producing $\geq$ 4.5 db scintillations was greater than 60% within the 95°W to 20°E longitude sector. A secondary maximum of 40% occurrence was found over the central Pacific. That is, large-amplitude irregularities are more apt to be found in the American-Atlantic and central Pacific sectors than in the Asian sector. A similar longitudinal distribution of spread-F plasma bubbles was observed by means of ISIS 1 instrumentation in the topside ionosphere (Burke et al., 1979a). Simultaneous measurements by the topside sounder on ISIS show that the plasma depletions extend along entire tubes of magnetic flux (Dyson and Benson, 1978). Both the OGO 6 and the ISIS 1 results were obtained in the evening sector, under magnetically quiet conditions during the November, 1969 to January 1970 period. Burke et al., (1980) using measurements from the

DMSP/F2 satellite found 10 examples of spread-F at an altitude of 840 km near 0700 LT during the magnetic storm of September, 197. Most of the examples were found in the Central Pacific and American longitudinal sectors.

The global distribution of quiet-time F-layer irregularities has a seasonal dependence that has been reviewed by Aarons (1977). At Huancayo, Peru the percentage of occurrence for 137 MHz scintillation indices of > 60 has a broad maximum during the summer (southern hemisphere) months. The percentage occurrence during winter months is nearly zero. A similar seasonal distribution was found at Kwajalein in the central Pacific. In the Indian sector bottomside spread-F is largely independent of season (R.G. Rastogi, priv. comm., 1979). Finally, we note that the magnitude of the near sunset enhancement of the zonal electric field and the local time of its reversal to its night-time westward direction as measured at Jicamarca, Peru, shows both seasonal and solar cyclic variations (Fejer et al., 1979). Near solar minimum the enhancement only appears near the equinoxes. It is present during all seasons at solar maximum with the greatest (least) intensity and with the latest (earliest) local time of the reversal at the time of the summer (winter) solstice. These measurements, in conjunction with scintillation frequencies at Huancayo (Aarons, 1977) indicate that zonal electric fields p.ay a central role in the generation of intense spread-F.

This report extends the above mentioned global studies by presenting a large set of equatorial, total ion density measurements from which we determine the longitudinal distribution of spread-F irregularities reaching the topside ionosphere. The measurements were taken by means of instrumentation on two DMSP (Defense Meteorological Satellite Program) satellites in the evening local time sector. Effects due to magnetic activity and seasonal differences at northern and southern ends of field lines were minimized by limiting ourselves to data taken during magnetically quiet periods within $\pm$ 10 days of the five equinoxes between September 1977 and September 1979.

The DMSP instrumentation is described briefly in the section immediately following. In the observation section, several typical equatorial passes are used to illustrate the consistency of density profiles within longitudinal sectors. Such consistency implies a high degree of magnetic control. Analysis of the distribution of spread-F leads to the conclusions that the longitudinal average probability of observing spread-F increases from $\simeq$ 20% at 19:30 LT to $\geq$ 50% after 21:00 LT, and that the longitudinal maximum shifts eastward from the Atlantic - African sector in the early evening to Asian sector at later times. Additionally we find that the occurrence of spread-F is associated with the crossing of a threshold of low latitude topside density the level of which depends in a complicated way upon season, local time and/or magnetic activity. In the discussion section we relate the present measurements to previously reported observations and theoretical results.

INSTRUMENTATION

DMSP satellites are three-axis stabilized vehicles in sun-synchronous, circular, polar orbit at an altitude of 840 km. The DMSP/F2 (F4) satellite was launched in August 1977 (June, 1979) with an ascending node at 0700 (1000) and descending node at 1900 (2200) LT. Because the F2 orbit is not inclined exactly at 98.7° it is subject to a slow precessional motion toward later local times. The precessional drift is ~ 1 hr local time per year. Thus, the data set covers approximately a three hour swath in the post-sunset local time sector.

The ion detectors are planar RPA's (retarding potential analyzers) mounted on 0.75 m booms. Each one has a gridded circular aperture of 2.54 cm diameter which is oriented normal to the motion of the satellite. The sampling rate is 7 per second. A complete description of the instrumentation and data reduction is given by Smiddy et al., (1978).

OBSERVATIONS

Our data set consists of ion density measurements on 331 equatorial crossings by DMSP in the evening sector. These passes are chosen to provide dense longitudinal coverage of six sequential quiet days in each of five successive equinoctial periods, as shown in Table 1. Of these, 172 show depletions at or near the magnetic equator.

Phenomenologically, the 1977 data differs from all the rest in that the depletions which occur are smooth, and as we shall see, can not be identified with spread-F. General features of the 1977 and other data subsets are illustrated in figures 1 and 2, respectively, where we compare density measurements taken in narrow longitudinal bins, within $\pm$ 30° Lat of the magnetic equator, on four successive days. We note that the ion density signature shows a high degree of repeatability within a given longitudinal bin, during a given equinox. Great differences are noted, on the other hand, when comparing the data of 1977 to that of other periods, or when comparing the results form different longitudinal bins, even on the same day. A high degree of magnetic control has also been found in quiet time morningside measurements by DMSP (Burke et al., 1979b).

FIG	PERIOD	LOCAL TIME	TOTAL PASSES	DEPL. PASSES	% DEPL.
3	1977, Sept. 4-9	18.9	68	56	82%
4a	1978, March 23-27	19.6	62	12	19%
4b	1978, Sept. 13-18	20.0	60	20	33%
4c	1979, March 11-15	20.6	39	25	64%
4d	1979, Sept. 12-17	21.5	56	32	57%
4e	1979, Sept. 12-17	21.9	46	27	59%

TABLE 1: TOTAL DEPLETION OCCURRENCE

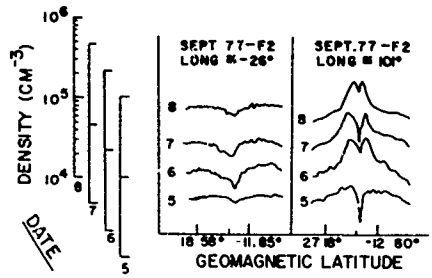


Fig. 1. Sept. 1977 ion density data from DMSP-F2 at 18.9 LT in two longitude bins on four successive days.

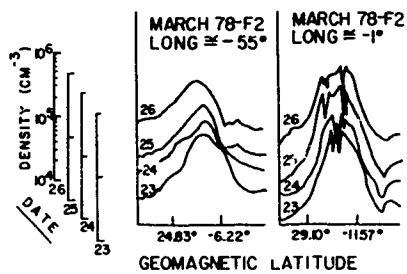


Fig. 2. March 1978 ion density data from DMSP-F2 at 20.00H LT in two longitude bins on four successive days.

The two sets of examples given in Figure 1 illustrate the general features of the September 1977 data subset. At longitudes near -26°, the ion density has a fairly uniform level of $\sim 0.7 \times 10^5$ between + 25° and -25° magnetic latitude except for a relatively smooth depletion near the magnetic equator. Near 101° longitude, on the other hand, the density rises slowly from $\sim 0.5 \times 10^5$ at $\pm$ 30° to $\sim 0.9 \times 10^5$ at $\pm$ 15° magnetic latitude. It then rises abruptly to $\sim 2.0 \times 10^5$ near the magnetic equator, and finally exhibits a smooth but narrow depletion at the equator. The smooth depletions were found in 82% of September 1977 passes, but in none of the 263 passes from the other equinoctial periods. In Figure 3, the symbols "X" and " ● " denote the presence or absence of the equatorial depletion. Each symbols position on the graph indicates the longitude of equatorial crossing and peak low-latitude density on that pass. The depletions occur at all longitudes but have a slight minimum in the American-Pacific sector - 64% probability compared to ≅ 100% probability in the African-Asian sector. Their occurrence is unrelated to peak density.

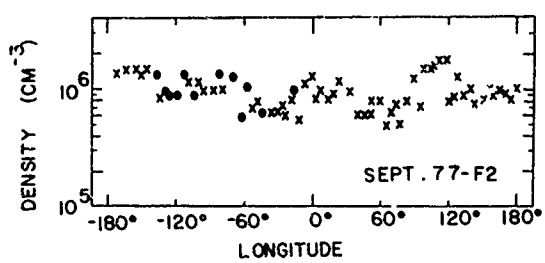


Fig. 3. Equatorial peak density vs. longitude for DMSP-F2 during Sept. 1977.

Anticipating our discussion below several comments may be made: (1) the equatorial depletions shown in Figure 1 are not examples of spread-F (2) The depletions are similar in morphology to those found during quiet times at the equator near 0700 LT (Burke et al., 1979b). They are most likely indicative of general near-sunset topside dynamics (3) No example of spread-F appears in the September 1977 data.

Measurements plotted in Figure 2, taken from March 1978, illustrate the two characteristics latitudinal profiles of density found in the four other equinoctial periods. Both types of profile are characterized by density maxima near the magnetic equator. The examples from the -55° longitude sector show smooth variations in density while those from the -1° sector show sharp irregular biteouts. It is seen that more than one biteout can be encountered during a single pass. These biteouts are spread-F bubbles that have reached an altitude of 840 km in the topside ionosphere. Finally, we note that the maximum density found in the longitudinal sector with spread-F biteouts exceeds that in the sector where they are absent.

Figure 4 is a plot of the maximum densities measured by DMSP during the last four equinoxes as functions of longitude. Data plotted in the top four panels come from F2 while those in the bottom panel are from F4. Recall that F2 precesses from 19.6 LT in March, 1978 to 21.5 LT in September 1979. F4 was at 21.9 LT. The symbols "+" and " ● " indicate whether or not at least one biteout was encountered. Several impressions may be gained from a cursory study of the distribution of points. (1) The frequency of occurrence for encountering spread-F increases from 19% at 19.6 LT to >50% in the 21-22 local time sector. (2) In any given equinoctial period spread-F appears most frequently in the high density portion of the distribution of points. (3) At early local times (March and September, 1978) the examples of spread-F are narrowly distributed about 0° longitude. The distribution first broadens then shifts eastward at later local times. Although an impression of an eastward shift could result from the data gap in F4 data between -150° and -45°, we note also that spread-F was not encountered frequently in this longitude sector by F2 during September, 1979.

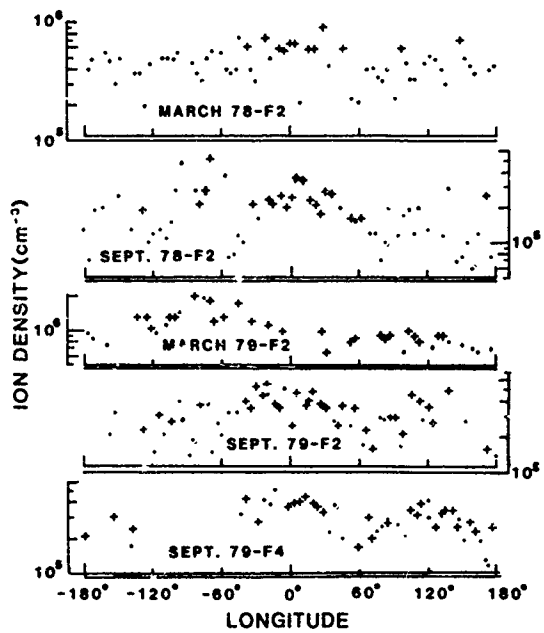


Fig. 4. Equatorial peak density vs. longitude during the late evening. Symbols "+" and " ● " represent cases with and without spread-F, respectively. Data are presented in order of increasing local time (19.6, 20.05, 20.6, 21.5 and 21.9).

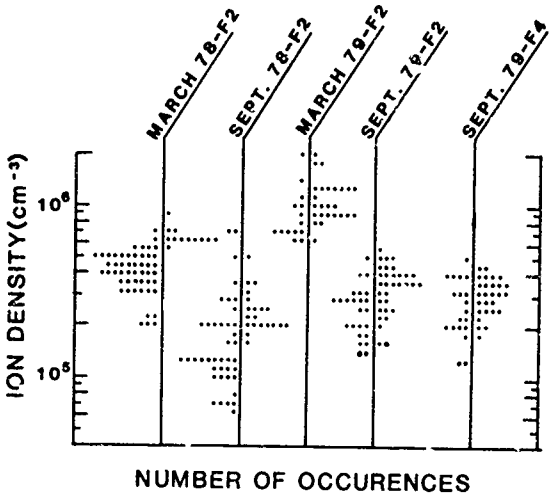


Fig. 5. Density distribution of passes with spread-F (right of dividing line) and without spread-F for each of the five data sets displayed in Fig. 4.

To better appreciate the second impression gained from Figure 4 we leave replotted densities measured in each equinox period irrespective of longitude (Figure 5). Data points plotted to the right (left) of the various separation lines indicate densities when spread-F was (was not) detected. In March and September, 1978 when the orbit was at relatively early local times, there is a sharp turn-on in densities below which spread-F is not observed. At later local times, the distribution of densities with and without spread-F become more comparable. However, both in March and September, 1979 the average values of densities when spread-F was present exceed those when it was absent.

For reasons that are not understood, Figure 5 clearly shows that topside densities in March are significantly greater than those in September. This does not seem to be a simple local time effect. The densities measured in September of 1978 and 1979 at three distinct local times are comparable. Also, the densities in March 1978 (19.6 LT) are less than those in March 1979 (20.6 LT). Figure 1 of Fejer et al., 1979 shows that at the equinox the zonal electric field reverses direction between 19.0 and 19.5 LT. Thus, during equinoctial periods the density at 840 km would be expected to decrease after 19.5 LT when the F-layer as a whole moves downward.

To simplify our perception of the longitudinal distribution of topside spread-F bubbles, we have grouped their occurrence in bins of 60° width in Figure 6. As an aid to interpretation we have also plotted in the bottom two panels, the magnetic field strength at 350 km, the altitude appropriate to spread-F origin, and the magnetic declination at 840 km, the orbital altitude where the angle of crossing field lines may be important. As discussed below, the 98.7° orbital inclination provides some geometrical bias favoring observations of the field-aligned features of spread-F in regions of westward declination.

Following the data sequence from the top to the bottom panel we note that in the 19.6-20.0 LT sector the spread-F bubbles are almost exclusively limited to the American to African longitude sectors. During March 1979, near 20.6 LT spread-F observations are uniformly distributed in longitude. In the post 2100 LT sector there is a definite shift to the east so that spread-F is most frequently seen at Asian longitudes. The significance of these distributions for global models of equatorial spread-F is discussed in the following sections.

DISCUSSION

Smooth depletions near the magnetic equator have been previously observed in the morning on DMSP F2 (Burke et al., 1979b) and Alouette (Goldberg et al., 1964) as well as in the evening on OGO-6 (Hanson and Sanatani, 1973; McClure and Hanson, 1973). The Alouette observations, at altitudes between 440 and 710 km, and some of the OGO-6 observations show depletions of a much broader geographical extent, typically 20° latitude. Others such as that of orbits 2404 (McClure and Hanson, 1973) and 2265 and 2266 (Hanson and Sanatini, 1973) on OGO-6 in the evening show the same extremely sharp gradi-

ents seen in our September 1977 data set and in the morning on DMSP (Burke et al., 1979b). The current September 1977 data set data set and the morning September 1977 quiet time data set of Burke et al., (1979b) display smooth depletions with rates of occurrence of 82% and 100%, respectively. Although this type of depletion has been observed previously, it has not been seen with such high rates of occurrence, probably because it is very restricted in season and local time.

McClure and Hanson, (1973) and Hanson and Sanatani, (1973) note that the minimum of the measured depletion occurs when the satellite is moving tangent to the magnetic apex shell (Van Zandt et al., 1972), i.e., the field line which crosses the equator at the lowest altitude, rather then at the dip equator. Such a distinction does not exist on a constant altitude satellite such as DMSP.

An examination of the width of smooth depletions in our September, 1977 data set has shown that they are narrowest when the satellite is moving perpendicular to the equator, i.e., tangent to the magnetic field. This fact, plus the closeness to the magnetic equator of all the smooth depletions described in this section, indicates that the measured depletions are part of an equatorially aligned structure, rather than a field-aligned structure, as usually presumed for spread-F.

The broader depletions observed by Alouette have been represented by the analytic solutions of Goldberg et al., (1964) for a specified equatorial density as boundary condition, without ExB drift. McClure and Hanson, (1973) and Hanson and Sanatani, (1973) used a similar model of field-aligned transport but point out the experimental evidence for a vertical ExB drift. They explain the sharp equatorial depletions seen by OGO-6 in terms of the bottomside of the F-layer having been lifted above the satellite orbit at the equator.

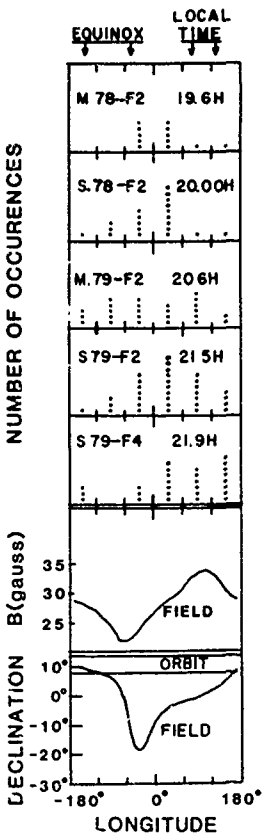


Fig. 6. Longitudinal distributrion of spread-F occurrence in 60° longitudinal bins for the five data sets. The magnetic field strength and declination are given in the bottom two frames.

Vertical scale heights of 1/3 km at the equator have been derived by Hanson and Sanatani, 1973 using a diffusive equilibrium distribution parallel to each field line. This very steep gradient is probably unstable too and therefore, limited by the production of drift waves (Kadomtsev, 1965).

Although the smooth depletions seen in our data set appear morphologically similar to those of Hanson and Sanatani (1973), their greater altitude makes an explanation in terms of a lifted bottomside unlikely. Vertical drift both across and along magnetic field lines (Burke et al., 1979b) may both occur, and this type of feature deserves much further study.

Previous in situ studies of equatorial spread-F in the evening have been less extensive or less seasonally selective than the current study, as tabulated below. In spite of differences, which will be detailed below, all the studies found irregularities near the Atlantic sector, although the maximum could be displaced in either direction.

Despite similarities, there were significant differences. Both the OGO-6 (Basu et al., 1976) and the INTERCOSMOS-2 (Kutiev et al., 1976) study showed a heavier preponderance of irregularities in the American Atlantic sector (0°-90°W) while the ISIS study (Burke et al., 1979a) and the current study showed a shift to the Atlantic-African sector (40°W to 40°E). The current study also shows a strong secondary maximum in total distribution over Asia, i.e., 100-150° E. reflecting the later evening part of the data, which is not shown by the other studies. Also the minimum of our longitudinal distribution occurs over the central Pacific, near 180°, which is the location of the secondary maximum of occurrence observed by Basu et al., (1976).

SATELLITE	AUTHORS	DATA SET
OGO-6	Basu et al., 1976	250 PASSES, NOV-DEC, 1969-1970
INTERCOSMOS-2	Kutiev et al., 1976	44 PASSES, UNSPECIFIED SEASON
ISIS-1	Burke et al., 1979a	300 PASSES, ENTIRE YEAR
DMSP-F2, F4	Present Study	331 PASSES, 1977-1979, EQUINOX

TABLE 2: PREVIOUS IN SITU STUDIES

It should be noted that the orbital inclination of DMSP and the declination of the magnetic field may combine to produce a geometrical bias in our observed global distribution of spread-F. In the evening sector DMSP satellites cross the geographic equator at an angle of 8.7° to the west of south. Thus, orbits cross relatively wide swaths of magnetic longitude in regions, such as the Atlantic, where declination is westward. In regions of eastward declination, such as the Pacific, orbits lie close to magnetic meridians. Since spread-F bubbles are field-aligned, they are more apt to be observed in regions of eastward declination where more flux tubes are being sampled. Satellites such as OGO 6 (inclination = 82°) would have different biases in this regard.

Our study is also unique among the in-situ studies in that it traces the development of the longitudinal distribution as a function of local time, as shown in Figure 6. There have, however, been local time distribution studies of the related phenomenon of radio scintillations (Aarons, 1975) which have shown similar increases of total occurrence frequency, during the evening, between 19.5 and 2200 LT.

The slow migration of the longitudinal distribution eastward with later local times is consistent with a picture in which the bubbles are rising with a velocity determined by the drift speed (Ossakow and Chaterverdi, 1978)

$$V = \frac{\Delta N}{N}\left[\frac{g}{\nu} + \frac{cE_o}{B}\right]$$

where c is the speed of light g is the acceleration of gravity, ν the ion-neutral collision frequency, E_0 the ambient electric field just outside the bubble, B the geomagnetic field and $\Delta N/N$ is the fractional depletion of the bubble relative to the ambient ion density. Therefore the drift is slower in regions of higher magnetic field, where consequently the bubbles appear later. That the bubbles are indeed rising is also confirmed by the development of the altitude distribution as a function of local time, which has been shown to possess a rising upper envelope (Burke et al., 1979a). Perhaps the most direct indication of the rise of F-region plasma bubbles is provided by simultaneous velocity and density measurements aboard AE-C (McClure et al., 1977). A study remains to be done of calculated equatorial height-envelope vs. local time, using the current data.

The association of spread-F with a rising ionosphere has also been investigated using bottomside ionosonde data (Rao, 1966). Seasonal differences were shown to exist between solstice and equinox data but vernal-autumnal differences were not investigated. In any one season, during all local times which Rao, (1966) investigated, however, and at both Kodaikanal and Huancayo, spread-F was associated with higher values of h'F, (the minimum virtual height of the F-region). The equinoctial value of the threshold for this parameter was about 400 km at both stations.

Although the mean value of topside density, and the mean value of h'F seen on nights with spread-F are both higher, at all local times, than the mean values on nights without, the sharpness of the threshold is reduced at later local times, as observed with both kinds of data. This is to be expected as the rising bubble either elongates itself into a vertical sheet or trails a plume of smaller scale irregularities behind it (Kelley et al., 1976) at later times. The elongation or pluming effect is also reflected in an increase of total frequency of occurrence at later local times as shown in Table 1. In short, an irregularity detected at earlier times is more likely to be near the top of a rising bubbles while one detected later could be either near the top of a rising growing bubble, or in the vertical sheet or decaying wake behind an older bubble. The actual level of the observed threshold may depend not only on local time and magnetic conditions, but upon season, solar activity and solar cycle as well (Basu, priv. comm., 1981).

CONCLUSION

We have shown that two types of depletions are present in evening equatorial ion density measurements. The first is a smooth depletion aligned closely along the magnetic equator which reflects the overall dynamics of the F-region. The second is an irregular field-aligned depletion which is the topside signature of spread-F. The occurrence of the latter depletions is associated with an enhancement of peak density on neighboring field lines. This association is consistent with the production of spread-F by a rising ionosphere. The threshold of topside density for spread-F appears to depend on season, solar activity, solar cycle, local time and magnetic conditions in some as yet unexplained way. The longitudinal distribution of the spread-F appears to migrate to regions of higher magnetic field at later local times in a way consistent again with a picture of rising bubbles of depleted plasma. Finally, both the topside density and spread-F occurrence exhibit a longitudinal repeatability which indicates a detailed control by the magnetic field of the earth.

ACKNOWLEDGEMENT

This work was supported, in part, under USAF contract # F19628-80-C-0116 with Regis College.

REFERENCES

Aarons, J., Global morphology of ionospheric scintillations II, AFCRL-TR-75-0135, Air Force Cambridge Res. Lab., Hanscom Air Force Base, MA, 1975.

Aarons J., Equatorial scintillations: A review, IEEE Trans. Antennas Propagat., 25, 729, 1977.

Anderson, D.N., and G. Haerendel, The motion of depletion plasma regions in the equatorial ionosphere, J. Geophys. Res., 84, 4251, 1979.

Basu, Suuanda, S. Basu, and B.K. Kahn, Model of equatorial scintillations from in situ measurements, Radio Science, 11, 821, 1976.

Burke, W.J., D.E. Donatelli, R.C. Sagalyn, and M.C. Kelley, Low density regions observed at high altitudes and their connection with equatorial spread-F, Planet Space Sci., 27, 593, 1979a.

Burke, W.J., R.C. Sagalyn, R.G. Rastogi, M. Ahmed, F.J. Rich, D.E. Donatelli, and P.J.L. Wildman, Postsunrise refilling of the low latitude topside ionosphere, J. Geophys. Res., 84, 4201, 1979b.

Burke, W.J., D.E. Donatelli, and R.C. Sagalyn, The longitudinal distribution of equatorial spread-F bubbles in the topside ionosphere, J. Geophys. Res., 85, 1335, 1980.

Dyson, P.L., and R.F. Benson, Topside sounder observations of equatorial bubbles, Geophys. Res. Lett., 5, 795, 1978.

Fejer, B.G., D.T. Farley, R.F. Woodman, and C. Calderon, Dependnece of F-region vertical drifts on season and solar cycle, J. Geophys. Res., 84, 5792, 1979.

Fejer, B.G., and M.C. Kelley, Ionospheric irregularities, Review of Geophysics and Space Science, 18, 401, 1980.

Goldberg, R.A., P.C. Kendall, and E.R. Schmerling, Geomagnetic control of the electron density in the F-region of the ionosphere, J. Geophys. Res., 69, 417, 1964.

Hanson W.B., and S. Sanatani, Large N_i gradients below the equatorial F-peak, J. Geophys. Res., 78, 1167, 1973.

Kadomtsev, B.B., Plasma Turbulence, Academic Press, New York, NY, 1965.

Kelley, M.C., G. Haerendel, H. Kappler, A. Valenzuela, B.B. Balsley, D.A. Carter, W.L. Ecklund, C.W. Carlson, B. Hausler, and R. Torbert, Evidence for a Rayleigh Taylor type instability and upwelling of depleted density regions during equatorial spread-F., Geophys. Res. Lett., 3, 448, 1976.

Kutiev, I., Ts.P. Dachev, and L. Sankov, Irregularities recorded by ion traps at about 1000 km height, Space Research, XVII, 483, 1976.

McClure, J.P., and W.B. Hanson, A catalog of ionospheric F-regions irregularity behavior based on OGO-6 retarding potential analyzer data, J. Geophys. Res., 78, 7431, 1973.

McClure, J.P., W.B. Hanson, and J.F. Hoffman, Plasma bubbles and irregularities in the equatorial ionosphere, J. Geophys. Res., 82, 2650, 1977.

Ossakow, S.L., and P.K. Chaturvedi, Morphological studies of rising spread-F bubbles, J. Geophys. Res. 83, 2085, 1978.

Ossakow, S.L., S.T. Zalesak, B.E. McDonald, and P.K. Chaturvedi, Nonlinear equatorial spread-F: dependence on altitude of the F-peak and bottomside background electron density gradient scale length, J. Geophys. Res., 84, 17, 1979.

Ott, E., Theory of Rayleigh Taylor bubbles in the equatorial ionosphere, J. Geophys. Res., 83, 2066, 1978.

Rao, B.C., Control of equatorial spread-F by the F-layer height, J. Atmos. Terr. Phys., 28, 1207, 1966.

Scannapieco, A.J., and S.L. Ossakow, Nonlinear equatorial spread-F, Geophys. Res. Lett., 3, 451, 1976.

Smiddy, M., R.C. Sagalyn, W.P. Sullivan, P. Anderson, and F. Rich, The topside ionosphere plasma monitor (SSIE) for the block 5D/Flight 2 DMSP satellite, AFGL-TR-78-0071, Air Force Geophys. Lab., Hanscom Air Force Base, MA, 1978.

VanZandt, T.E., W.L. Clark, and J.M. Warnock, Magnetic apex coordination: A magnetic coordinate system for the ionospheric F2 layer, J. Geophys. Res., 77, 2406, 1972.

Woodman, R.F., and C. La Hoz, Radar observations of F-region equatorial irregularities, J. Geophys. Res., 81, 5447, 1976.

REMOTE DETECTION OF THE MAXIMUM ALTITUDE OF EQUATORIAL IONOSPHERIC PLASMA BUBBLES

Robert F. Benson
Laboratory for Planetary Atmospheres
Goddard Space Flight Center
Greenbelt, MD 20771

ABSTRACT

The topside sounders on the Alouette 2 and the ISIS 1 satellites have been previously used to demonstrate the field-aligned nature of equatorial ionospheric plasma bubbles (Dyson and Benson, Geophys. Res. Lett., 5, 795-798, 1978). The present work is based on an investigation of nearly 200 post-sunset low altitude passes of these satellites near the dip equator in order to determine maximum bubble altitudes. These maximum altitudes are determined by calculating the apex altitude of the magnetic field line passing through the satellite when it is immersed in a bubble. Such calculations are performed only when ionospheric echoes are observed from the conjugate hemisphere. These echoes result from ducted HF sounder signals which are guided along field-aligned irregularities within the plasma depletion (bubble). The maximum bubble altitudes (which correspond to three longitude sectors centered on 0°, 75°W and 105°E) often exceed 1000 km but seldom 3000 km. They can exceed 2000 km in the early evening as well as after midnight. Events at such high altitudes are only observed in the American longitude sector. Most bubbles are observed between 20 and 02 hours local time. The electron density depletions within these field-aligned bubbles, as measured at the point of satellite encounter in the topside ionosphere, are generally less than a factor of 2 but can exceed a factor of 10.

INTRODUCTION

There are several lines of experimental evidence which indicate that large electron density N_e depletions in the equatorial ionosphere, commonly called equatorial bubbles, are distributed along the direction of the earth's magnetic field $\vec{B}$. Airborne investigations utilizing a combination of radio and optical techniques revealed large scale regions of decreased airglow intensity which have been interpreted as the bottomside signature of field-aligned equatorial plasma bubbles [Buchau et al., 1978; Weber et al., 1978; Weber et al., 1980]. The field aligned nature of these bubbles in the topside ionosphere was confirmed by the reception of ducted HF echoes from the conjugate hemisphere by the ISIS 1 topside sounder satellite when it was immersed in a region of locally depleted electron density [Dyson and Benson, 1978]. The field alignment of equatorial plasma bubbles in the topside ionosphere was also indicated by the plasma density probe on the polar orbiting ESRO-4 satellite where 21 near-conjugate pairs of plasma density depletions were identified [Heron and Dorling, 1979]. Additional evidence comes from models of guided propagation inside equatorial plasma bubbles which can explain VHF transequatorial propagation during the evening hours [Heron and McNamara, 1979; Heron, 1980]. Observations with the fully-steerable high-power ALTAIR radar on the Kwajalein Atoll in the Marshall Islands have provided further insight into the characteristics of magnetic field-aligned equatorial plasma bubbles [Tsunada, 1980b]. The correlation of poleward propagating patches of 6300Å airglow disturbances with the occurrence of spread F on bottomside ionograms has been interpreted by Sobral et al. [1980] as evidence of the vertical motion of field-aligned plasma bubbles over the magnetic equator.

Topside sounders can provide a wealth of information pertaining to equatorial plasma bubbles in addition to their field-aligned nature. Dyson and Benson [1978] where able to determine the N_e distribution along $\vec{B}$, to infer ion composition changes along $\vec{B}$ within the duct associated with the bubble, and to detect bubble asymmetries with respect to the magnetic equator. These conclusions were based on in-situ observations, i.e., obtained when the ISIS 1 sounder was immersed within a bubble. Muldrew [1980b] conducted a statistical investigation, based on the remote detection of bubbles from aspect sensitive scatter spread F signal returns, using data from the Alouette 1 topside sounder. This investigation determined the occurrence frequency of bubbles as a function of local time and longitude, the height dependence with local time, and typical bubble dimensions. Muldrew [1980a] conducted a statistical investigation on the occurrence of ducted echoes as observed by the ISIS 2 sounder. The results were presented as a function of L value and local time and were related to the study of equatorial bubbles. He concluded that ducts with $L \lesssim 1.2$ are associated with equatorial bubbles whereas those with $L \gtrsim 1.2$ are not. The importance of topside sounder data in the study of the highly disturbed nighttime equatorial ionosphere was recently emphasized by the recording of ISIS data during the "Wideband Equatorial" Defence Nuclear Agency campaigns at the Kwajalein Missile Range during the summers of 1977 and 1978 [Fulks, 1980]. Spread F signatures, characteristic of the initial phase of equatorial plumes, were found to occur most often on a particular L shell ($L \sim 1.12$). These remote spread F returns were found to be different from those observed at other times and locations and were considered to be possibly related to equatorial bubbles.

The present work is based on in-situ topside sounder bubble encounters. A bubble encounter is determined by the presence of a local spread F cloud plus ducted echoes from the conjugate hemisphere. A threshold level for the N_e depletion was not used as a criterion for a bubble encounter since one of the goals of the investigation was to determine the altitude where the density perturbation was essentially zero. The presence of local spread F was required in order to rule out the frequently occurring ducts which are observed independent of local spread F and are not considered to be associated with bubbles, e.g., see Figure 1 of Muldrew [1963]. The conjugate echo requirement was necessary in order to infer the maximum altitude of the field aligned density structure by determining the apex altitude of the associated magnetic field line. This maximum altitude will be presented as a function of local time. In addition, the N_e perturbation at the bubble encounter will be determined and will be related to the maximum altitude.

OBSERVATIONS

Anatomy of an event. The criteria used by Dyson and Benson [1978] for recognizing a bubble encounter were the presence of conjugate ducted echoes within an electron density depletion. A critical signature was the combination of ducted echoes from the local and conjugate hemispheres to form an epsilon shaped signal return on the ionogram. In the present investigation the encounter criteria were the presence of in-situ spread F and ducted echoes from the conjugate hemisphere. The combination of ducted echoes to form an ε signature was usually present but was not required if a definite conjugate ducted echo could be identified. Similarly, a depletion in N_e was observed but was not required.

An example of an equatorial plasma bubble ecounter by the ISIS 1 sounder is presented in Figure 1. On the left are portions of 6 ionograms; on the right are relevant orbital parameters, N_e along the satellite track and a schematic illustration to depict the possible structure encountered by the satellite. Key ionogram features are identified above ionogram a (H, 2H, 3H, etc., for the resonances at the electron cyclotron frequency f_H and its harmonics, N for the resonance at the plasma frequency f_N, and X for the extraordinary wave cutoff frequency f_x). Ionogram a reveals intense local spread F. No vertical (or ducted) ionospheric reflection traces are apparent even though the lack of interfering signals below about 9 MHz from ground-based transmitters indicates that an ionosphere capable of considerable ionospheric shielding existed below the satellite. This lack of echoes is attributed to extremely efficient scattering by the field-aligned N_e irregularities in the vicinity of the satellite which give rise to the local spread F. Ionograms b-e reveal ducted echoes, with varying amounts of spread F, from both the near and conjugate hemispheres. The near echo extends up to f_x. The conjugate echo also starts at f_x but at a virtual range of several thousand kilometers which initially decreases as the sounder frequency increases. Again, local spread F is present and vertical echoes are not observed on these ionograms. Vertical echoes are observed on ionogram f from both the ordinary and extraordinary waves down to the point where the (now remote) spread F cloud is encountered and a portion of a weak vertical echo is observed between 5 and 6 MHz, i.e., on the high frequency side of the spread F cloud. The ionogram between e and f (which is not shown) was similar to f but the virtual range to the spread F cloud was less (200 km to the low frequency edge) and no vertical echo was observed at frequencies beyond the spread F cloud maximum frequency. The

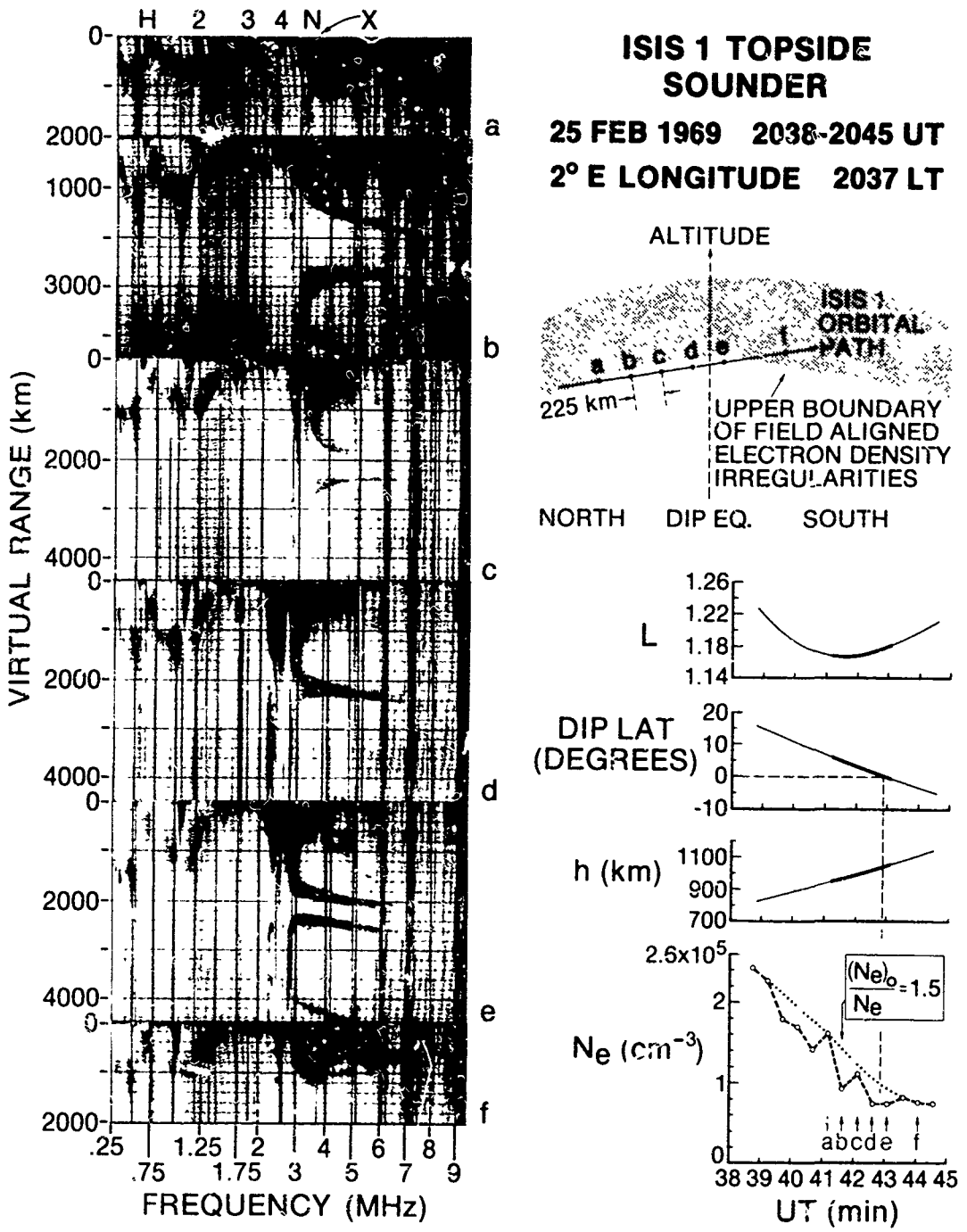


Figure 1. Ionograms recorded during an equatorial plasma bubble encounter by ISIS 1 (left), the electron density N_e along the satellite orbit as determined from these ionograms (lower right), pertinent orbital parameters (center right) and a schematic representation of the satellite's path through the top portion of the depleted region (top right).

strength of the ducted signals on b–e indicates the efficiency of wave ducting from an in-situ sounder. Similarly, the lack of vertical echoes on these same ionograms (even though ionospheric shielding implies the presence of a reflective ionosphere below the satellite) indicates the difficulty of transmitting signals across the field-aligned irregularities responsible for the spread F.

The electron density values plotted in the lower right of Figure 1 were determined from either (or both) f_N or f_x together with f_H. The values for each ionogram are connected with a dashed line and an estimate of the undisturbed background electron density $(N_e)_o$ by a dotted line. In this event the maximum electron density depletion, expressed in terms of the ratio $(N_e)_o/N_e$, was 1.5. The N_e structure shown in this figure reveals only the tip of the iceberg of density fluctuations since the sample points are spaced several hundred kilometers apart. An inspection of the ISIS 1 cylindrical electrostatic probe data for events such as this one reveals intense N_e fluctuations over scale lengths of tens of kilometers. The irregularities responsible for the ducted echoes on ionograms b-e correspond to an even smaller scale, i.e., N_e fluctuations of the order of a few percent over a distance of a few kilometers [Muldrew, 1963]. Thus it is not the "large" depletion of a factor of 1.5 that is responsible for the ducted echoes, rather it is the superimposed small scale structure. This point was also stressed by Dyson and Benson [1978] based on a consideration of the wide lateral extent of the large scale structure.

The schematic on the right hand side of Figure 1 illustrates that the satellite makes a nearly longitudinal pass through the top edge of the bubble. Because of the intense scattering, and the resulting lack of vertical echoes, very little information is available concerning the lower regions of the bubble. If the upper boundary is assumed to be relatively smoothly distributed along a magnetic field line and if errors in the magnetic field model are neglected, then the change in L value from the boundary encounter at ionogram a to the boundary encounter after ionogram e would imply a large upward component of vertical motion (500 m/sec). It is difficult to place much confidence on such an interpretation, however, since there is some evidence that the upper boundary may not be so well defined. In contrast to the relatively well defined boundary south of the dip equator, the ionograms recorded prior to the northern boundary crossing indicated a more complicated situation. While none of these ionograms revealed ducted echoes, some recorded remote spread F, some were engulfed in in-situ spread F and many were partially obscured by data drop-outs of the type discussed by Dyson and Benson [1978] in connection with their Figure 2. Thus the ISIS 1 sounder may have encountered the boundary at a time prior to the recording of ionogram a due to temporal and/or spatial variations in the bubble phenomena. Such an earlier crossing would correspond to a higher L value (than for ionogram a) which would reduce the deduced upward component of velocity. The subject of upward bubble velocity will be considered further in the next section based on the evaluation of a number of events to determine the distribution of the maximum bubble altitude with local time.

The field-aligned ducts associated with the event of Figure 1 show evidence of being asymmetrical with respect to the dip equator. Evidence for such a condition was also presented by Dyson and Benson [1978]. Ionogram c in Figure 1 is very similar to their Figure 5 in that the ducted conjugate hemisphere echo extends to a higher frequency than the ducted echo in the hemisphere of the satellite. In addition, the two ducted echoes merge to the same virtual range on ionogram d whereas the dip equator is crossed more than 100 km farther along in the orbit (closer to the recording of ionogram e). These observations suggest that at a given altitude the density was greater south of the dip equator than north of the dip equator during this ev[illegible].

Statistical results. Nearly 200 Alouette 2 and ISIS 1 satellite passes over the dip equator from 1968-1971 in 3 longitude sectors were inspected for the presence of equatorial plasma bubbles. The geographic latitudes and longitudes of the dip equator crossings for the satellite orbits used in this study are presented in Figure 2. The data were selected so as to have a low altitude ($\lesssim$ 1500 km) at the dip equator in the local time range from (approximately) sunset to sunrise. These constraints resulted in the data being primarily (80%) from ISIS 1. The satellite altitudes over the dip equator are plotted against the longitudes of the dip equator crossings in Figure 3. The satellite altitudes and longitudes at the dip equator crossings are plotted against local time in Figures 4 and 5, respectively. In Figures 3, 4, and 5 the open circles denote satellite passes where an equatorial bubble was encountered. From Figure 5 it is seen that bubbles were encountered more frequently in the African sector than in the American or Asian sectors. After discounting the orbits corresponding to local times prior to 18 hours in the American sector, so that the 3 data sets will correspond to approximately the same local time interval, bubbles were only encountered on about 9% of the passes over the American sector and 7% of the passes over the Asian sector whereas they were encountered on 20% of the passes over the African sector.

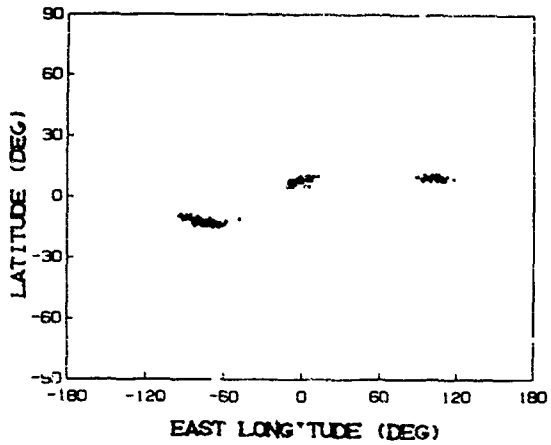


Figure 2. Geographic coordinates of the satellite dip equator crossings for all orbits used in the present study.

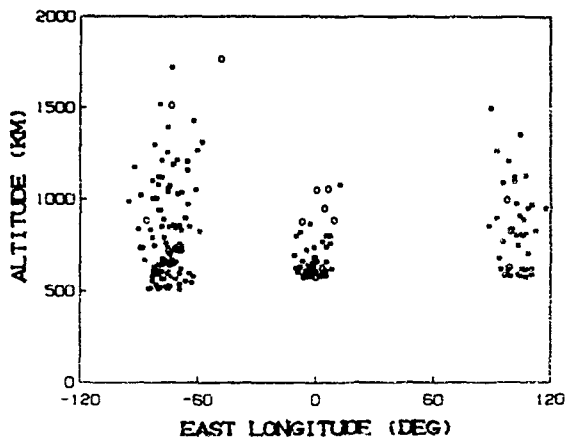


Figure 3. Altitude and geographic longitude of satellite dip equator crossings for orbits both with (circles) and without (asterisks) plasma bubble encounters.

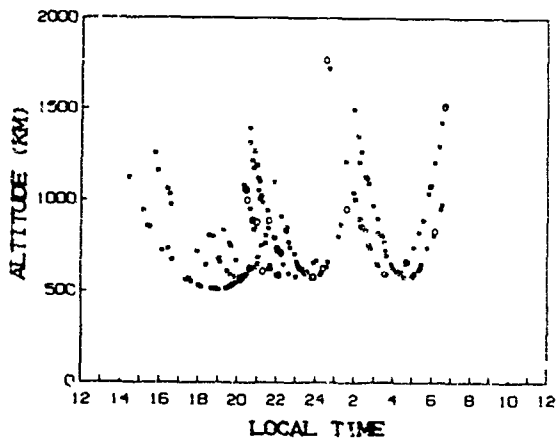


Figure 4. Altitude and local time of satellite dip equator crossings for orbits both with (circles) and without (asterisks) plasma bubble encounters.

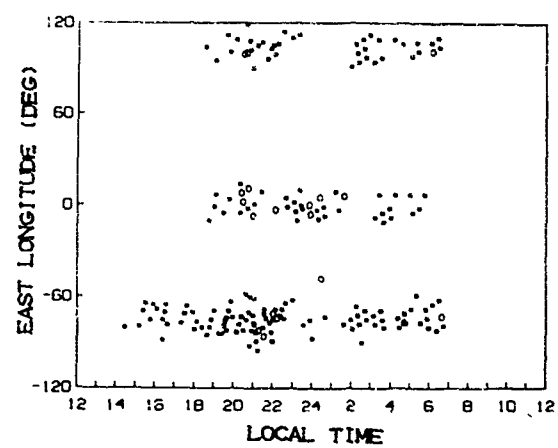


Figure 5. Geographic longitude and local time of satellite dip equator crossings for orbits both with (circles) and without (asterisks) plasma bubbles encounters.

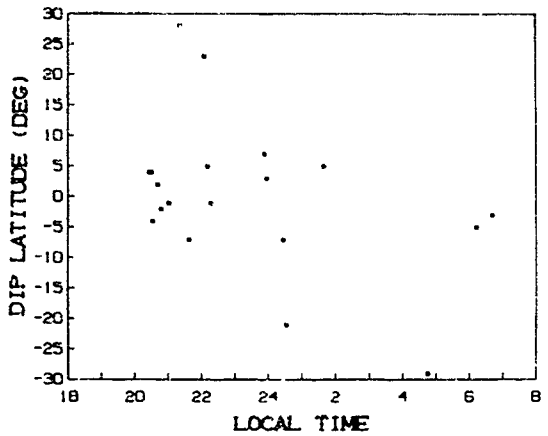


Figure 6. Dip latitude and local time corresponding to ISIS 1 plasma bubble encounters.

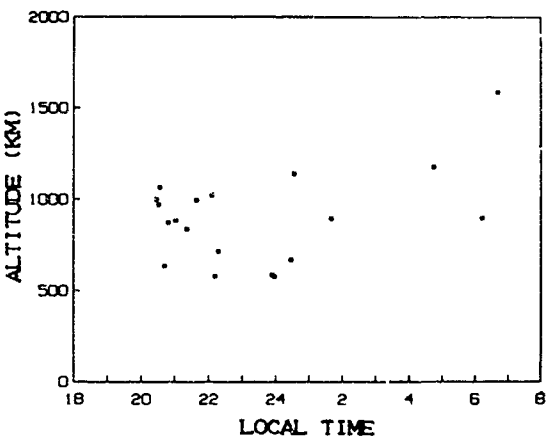


Figure 7. Altitude and local time corresponding to ISIS 1 plasma bubble encounters.

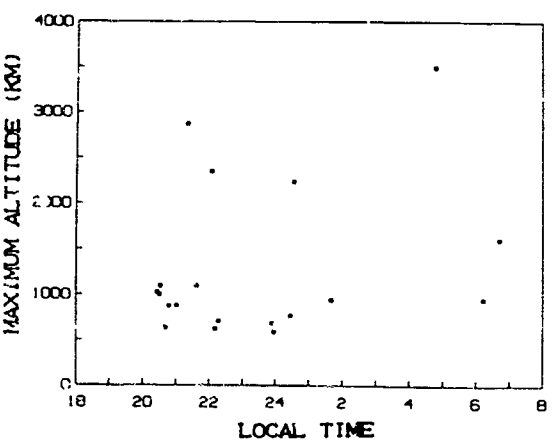


Figure 8. Projected plasma bubble maximum altitude over the dip equator h_{max} against local time.

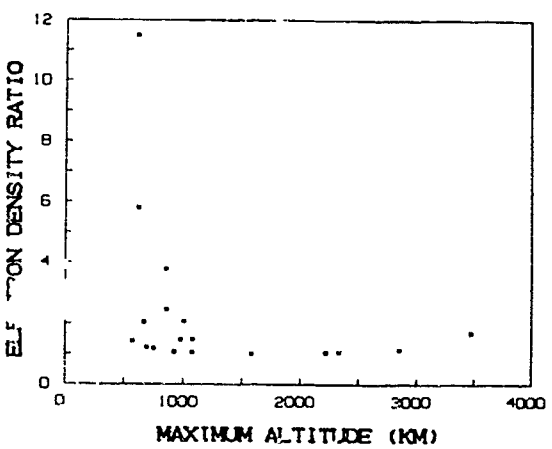


Figure 9. Ratio of the background electron density to the bubble electron density $(N_e)_o/N_e$ against h_{max}.

The dip latitudes and altitudes of the actual bubble encounters are plotted against local time in Figures 6 and 7. The longitudes of the bubble encounters are not presented here in this fashion because they are approximately the same as the longitudes of the dip equator crossings of the satellite as given in Figure 5 by the open circles. When several ionograms contain the signature of a bubble encounter, i.e., local spread F together with ducted conjugate echoes, the time corresponding to the ionogram recorded farthest from the dip equator was used to record the location in Figures 6 and 7. The event of Figure 1, for example, was identified by the time and position information of ionogram b. This procedure was used in order to tag the ducting path corresponding to the maximum altitude within the bubble structure at the dip equator. This maximum altitude h_{max} was calculated from the expression $h_{max} = (R_E + h)/\cos^2\theta - R_E$, based on a dipole field line, where θ is the dip latitude from Figure 6, h is the altitude from Figure 7 and R_E is the radius of the earth. A presentation of h_{max} vs. local time is given in Figure 8.

Figures 6, 7, and 8 indicate that while most events (80%) occurred between 2000 and 0200 local time, a few stragglers occurred near sunrise; while most events (80%) were encountered within 7^o of the dip equator, some were observed as far as 30^o from the dip equator; while most events (80%) were encountered at altitudes below 1100 km, some were encountered up to the maximum altitude sampled in this study (see Figure 4 also); and while most of the projected maximum bubble altitudes at the dip equator (70%) were $\lesssim$1100 km, values as high as 3500 km were obtained. These high values of h_{max} were obtained before local midnight as well as after midnight.

The density depletion within the bubble structure, as measured by the electron density ratio $(N_e)_o/N_e$, is plotted against h_{max} in Figure 9. Most of the depletions (80%) have $(N_e)_o/N_e \lesssim 2$. Larger depletions were only observed at lower projected maximum altitudes ($h_{max} \lesssim 900$ km), with the largest depletion (11.5) occurring at the lowest projected altitude ($h_{max} \approx 580$ km). The bubbles with $h_{max} \gtrsim 1100$ km had $(N_e)_o/N_e < 1.2$, except for the highest bubble (near 05 hours local time) where $(N_e)_o/N_e = 1.7$, and were observed in the -75^o longitude sector.

DISCUSSION

One of the most striking features in connection with the ionograms presented in Figure 1 is the intense scattering of radio signals associated with the bubble structure. The in-situ encounter indicates that the scattering is due to irregularities throughout the bubble rather than to irregularities confined to surface features. Such irregularities, which are maintained for thousands of kilometers along $\vec{B}$, are able to support the ducted propagation so evident in Figure 1 and first reported by Dyson and Benson [1978] in connection with equatorial plasma bubbles. Scattering throughout the bubble volume was assumed by Muldrew [1980b] in his investigation of bubbles observed in the Alouette 1 topside sounder data. He concluded that the cross sectional dimension of a typical field-aligned bubble observed in the 700 to 800 km altitude range was about 60 km. In a ground-based investigation using the Altair incoherent scatter radar, Tsunoda [1980a] concluded that scattering of signals from field-aligned irregularities took place over an altitude range of the order of 100 km associated with a topside bubble (with similar dimensions) just above the F peak. The present ISIS 1 results indicate that the dimensions of high altitude field-aligned bubbles ($h \gtrsim 1000$ km) are several hundred kilometers in cross section. This conclusion follows from an inspection of ionogram f in Figure 1 (and later ionograms not reproduced here) where the cloud like scatter signature is observed at a remote location and thus at higher frequencies (relative to f_x) where the difference between virtual range and true range becomes less important. These large field-aligned scattering volumes, high in the equatorial topside ionosphere, are likely to be a major factor in the commonly observed [e.g., Basu et al. 1980] intense VHF and gigahetrz equatorial scintillation. This scintillation is much more intense than that caused by ionization irregularities in either the bottomside of the equatorial ionosphere or the auroral zone [Kelley et al., 1980].

The observation that most bubbles occur between 2000 and 0200 local time (see Figures 6, 7, and 8) agrees with the Alouette 1 results of Muldrew [1980b]. The greatest frequency of occurrence of bubbles over the African sector and the least frequency of occurrence over the Asian sector (see Figure 5) is also consistent with his results. The present study, however, finds the frequency of occurrence in the American sector to be closer to the results obtained for the Asian sector rather than to the African sector as found by Muldrew. One must keep in mind the different selection criteria used in these two bubble studies based on topside sounder data. The requirement of in-situ bubble encounters in the present investigation would rule out most of Muldrew's events which included remote bubbles.

The topside sounder technique, however, will detect more in-situ bubbles than probe techniques due to the efficiency with which slight field-aligned electron density depletions can be detected by means of the reception of ducted echoes. Indeed, 80% of the bubbles detected in

the present study (Figure 9) would probably not have been detected in probe investigations such as those of Burke et al. [1979, 1980] where depletions are not considered unless $(N_e)_o/N_e \gtrsim 2$. The results of Figure 1 indicate how much more common small depletions are in the topside ionosphere than large depletions. In contrast to the situation often encountered in the vicinity of the F peak, where plasma depletions of several orders of magnitude can be observed [e.g., see Brinton et al., 1975; McClure et al., 1977], depletions of an order-of-magnitude are rare well into the topside ionosphere (see Figure 9). Similar results have been obtained from ion density measurements using the ISIS 1 spherical electrostatic analyzer where high altitude depletions were found to range from a factor of 2 to 8 [Burke, et al. 1979].

The distribution of projected maximum bubble altitudes at the dip equator with local time (Figure 8) has important implications for the concept of rising bubbles of depleted ionospheric plasma. This concept was initiated by the work of Kelley et al. [1976], Scannapieco and Ossakow [1976], and Woodman and LaHoz [1976] in order to explain nighttime equatorial spread F. Since these bubbles are formed in the lower F region after sunset in this model, the observations of bubbles prior to 2200 local time at altitudes above 2000 km are difficult to explain with reasonable bubble rise velocities. Burke et al. [1979] obtained a bubble height distribution with local time which did not include values at altitudes above 2000 km prior to 2400 hours and they claimed that observations in this space-time region would not be consistent with the rising bubble mechanism. The two cases observed in this region in the present study had $(N_e)_o/N_e \ll 2$ and would not have even been detected in the investigation of Burke et al. [1979]. The possibility that these high altitude bubbles with small electron density depletions (1.15 for the bubble with $h_{max} \approx 2900$ km or $L \approx 1.5$ and 1.08 for the one with $h_{max} \approx 2400$ km or $L \approx 1.4$ in Figure 8) are remnants from the previous day should be investigated [M. C. Kelly, private communication, 1981]. Another possibility is that these high altitude events are not due to rising bubbles. Muldrew [1980a] has suggested that medium frequency propagation ducts at altitudes corresponding to $L \gtrsim 1.2$ are formed in-situ and corotate with the earth. Muldrew attributes their formation to field-aligned currents resulting from a mismatch of conjugate potentials as first proposed by Cole [1971].

An important point concerning the high altitude bubbles ($h_{max} \gtrsim 1100$ km) of Figure 8 is that they all occur in the -75° longitude sector. This result is consistent with the DMSP and ISIS 1 observations of Burke et al. [1980] that topside plasma bubbles are found more often in the American longitude sector. They argued that the reduced value for $\vec{B}$ in this region makes it more likely for bubbles to rise well into the topside ionosphere. The present observations, which determine the maximum altitude a bubble has reached at a given local time, are consistent with this concept.

Acknowledgments. I am grateful to Diane Ziolkowski for assistance in the data analysis and computer programming. I am also grateful to Gary Page for helpful discussions and programming assistance. The data used in this investigation were obatined from the National Space Science Data Center at Goddard.

REFERENCES

Basu, S., S. Basu, J. P. Mullen, and A. Bushby, Long-term 1.5GHz amplitude scintillation measurements at the magnetic equator, Geophys. Res. Lett., 7, 259-262, 1980.

Brinton, H. C., H. G. Mayr, and G. P. Newton, Ion Composition in the nighttime equatorial F region: Implication for chemistry and dynamics (abstract), EOS Trans. American Geophys. Union, 56, 1038, 1975.

Buchau, J., E. J. Weber, and J. P. McClure, Radio and optical diagnostics applied to an isolated equatorial scintillation event, in Effect of the Ionosphere on Space and Terrestrial Systems, pp. 115-123, Edited by John M. Goodman, based on Ionosphere Effects Symposium held in Crystal City, Arlington, Va., Jan. 24-26, 1978, jointly sponsored by the Naval Research Laboratory and the Office of Naval Research, Washington, D.C., Government Printing Office, Washington, D.C., 1978.

Burke, W. J., D. E. Donatelli, and R. Sagalyn, The longitudinal distribution of equatorial spread F plasma bubbles in the topside ionosphere, J. Geophys. Res., 85, 1335-1340, 1980.

Burke, W. J., D. E. Donatelli, R. C. Sagalyn, and M. C. Kelly, Low density regions observed at high altitudes and their connection with equatorial spread F, Planet. Space Sci., 27, 593-601, 1979.

Cole, K. D., Formation of field-aligned irregularities in the magnetosphere, J. Atmos. Terr. Phys., 33, 741-750, 1971.

Dyson, P. L., and R. F. Benson, Topside sounder observations of equatorial bubbles, Geophys. Res. Lett., 5, 795-798, 1978.

Fulks, G. A., ISIS topside sounder data gathered on Kwajalein Atoll during the summers of 1977 and 1978, Defense Nuclear Agency Report 5339F, Prepared for Director, Defense Nuclear Agency, Washington, D.C. 20405, 31 October 1980.

Heron, M. L., Transequatorial propagation through equatorial bubbles - discrete events, Radio Sci., 15, 829-835, 1980.

Heron, M. L., and E. B. Dorling, Equatorial ionospheric plasma density bubbles observed by ESRO-4, Planet. Space Sci., 27, 1303-1310, 1979.

Heron, M. L., and L. F. McNamara, Transequatorial VHF propagation through equatorial plasma bubbles, Radio Sci., 14, 897-910, 1979.

Kelley, M. C., K. D. Baker, J. C. Ulwick, C. L. Rino, and M. J. Baron, Simultaneous rocket probe, scintillation, and incoherent scatter radar observations of irregularities in the auroral zone ionosphere, Radio Sci., 15, 491-505, 1980.

Kelley, M. C., G. Haerendal, H. Kappler, A. Valenzuela, B. B. Balsley, D. A. Carter, W. L. Ecklund, C. W. Carlson, B. Hausler, and R. Torbert, Evidence for a Rayleigh-Taylor type instability and upwelling of depleted density regions during equatorial spread F, Geophys. Res. Lett., 3, 448-450, 1976.

McClure, J. P., W. B. Hanson, J. H. Hoffman, Plasma bubbles and irregularities in the equatorial ionosphere, J. Geophys. Res., 82, 2650-2656, 1977.

Muldrew, D. B., The formation of ducts and spread F and the initiation of bubbles by field-aligned currents, J. Geophys. Res., 85, 613-625, 1980a.

Muldrew, D. B., Characteristics of ionospheric bubbles determined from aspect sensitive scatter spread F observed with Alouette 1, J. Geophys. Res., 85, 2115-2123, 1980b.

Muldrew, D. B., Radio propagation along magnetic field-aligned sheets of ionization observed by the Alouette topside sounder, J. Geophys. Res., 68, 5355-5370, 1963.

Scannapieco, A. J., and S. L. Ossakow, Nonlinear equatorial spread F, Geophys. Res. Lett., 3, 451-454, 1976.

Sobral, J. H. A., M. A. Abdu, C. J. Zamlutti, and I. S. Batista, Association between plasma bubble irregularities and airglow disturbances over Brazillian low latitudes, Geophys. Res. Lett., 7, 980-982, 1980.

Tsunoda, R. T., On the spatial relationship of 1-m equatorial spread F irregularities and plasma bubbles, J. Geophys. Res., 85, 185-190, 1980a.

Tsunoda, R. T., Magnetic-field-aligned characteristics of plasma bubbles in the nighttime equatorial ionosphere, J. Atmos. Terr. Phys., 42, 743-752, 1980b.

Weber, E. J., J. Buchau, and J. G. Moore, Airborne studies of equatorial F layer ionospheric irregularities, J. Geophys. Res., 85, 4631-4641, 1980.

Weber, E. J., J. Buchau, R. H. Eather, and S. B. Mende, North-south aligned equatorial airglow depletions, J. Geophys. Res., 83, 712-716, 1978.

Woodman, R. F., and C. La Hoz, Radar Observations of F region equatorial irregularities, J. Geophys. Res., 81, 5447-5466, 1976.

AMPLITUDE AND PHASE SCINTILLATION MEASUREMENTS USING THE GLOBAL POSITIONING SYSTEM

by

C. L. Rino and M. D. Cousins
SRI International
Menlo Park, California 94025

J. A. Klobuchar
Air Force Geophysics Laboratory
Hanscom Air Force Base, Massachusetts 07131

ABSTRACT

The Global Positioning System (GPS) will ultimately include 18 to 24 satellites in 12-hour orbits, such that at least four satellites are simultaneously visible from any point on the earth. The satellites transmit a 10.23-MHz code with a period of exactly one week for precise time-of-arrival measurements. The system operates at L-band to minimize ionospheric effects. Nonetheless, an appropriately designed receiver can Doppler track and reconstruct the "compressed" carrier with adequate sensitivity to extract phase scintillation data. The phase data can be processed to predict the scintillation at lower frequencies where the perturbations are much more severe. Thus, the global coverage of the GPS system can be exploited for scintillation monitoring.

A dual-frequency, Doppler-tracking GPS receiver with a specially designed interface for scintillation measurements has been developed. We shall present the preliminary results from operations at the Pacific and Atlantic equatorial stations and will discuss the potential of the GPS system for global scintillation monitoring.

I INTRODUCTION

The Global Positioning System (GPS) will ultimately consist of 18 to 24 satellites, each transmitting a unique 7-day-long PRN code at a 10.23 MHz rate, together with orbital elements and time marks from which a user can determine the satellite's position at the time of transmission (GPS time). By simultaneously delay and Doppler tracking the signals from four satellites, users can determine their position and velocity in real time. (The fourth satellite is required to eliminate the user's clock offset from GPS time.) The GPS constellation will be configured so that at least four satellites are visible from any point on earth at all times (see "Navigation," Vol. 25, No. 2, 1978, entire issue).

The signal propagation time is given by the formula

$$\tau_d = \rho/c + 1.345 \frac{N_T(10^{16}\ \text{el/m}^2)}{[f(\text{GHz})]^2} \quad , \tag{1}$$

where ρ is the range to the satellite, N_T is the total electron content (TEC), and f is the frequency. In the units indicated, the ionospheric delay is given in nanoseconds, i.e., the ionospheric delay is 1.345 ns per 10^{16} el/m^2 at 1 GHz. The corresponding carrier phase perturbation in radians is

$$\delta\phi = k\rho - 8.45 \frac{N_T(10^{16}\ \text{el/m}^2)}{f(\text{GHz})} \quad , \tag{2}$$

where $k = 2\pi f/c$. Note that aside from a sign, Eq. (1) is obtained from Eq. (2) after dividing by $2\pi f$.

Typical ionospheric delay errors exceed several tens of nanoseconds, with corresponding Doppler errors approaching 10 Hz. To eliminate these ionospheric biases, the GPS satellites transmit two identically coded signals (L_1 at 1575.42 MHz and L_2 at 1227.6 MHz). The delay difference

$$\Delta\tau \triangleq \tau_d^{(1)} - \tau_d^{(2)}$$

$$= 1.345 \frac{N_T}{f_1^2} [1 - (f_1/f_2)^2]$$

$$= 0.35\ N_T\ (\text{ns}) \tag{3}$$

isolates the ionospheric term whereby it can be eliminated from Eq. (1). Similarly,

$$\Delta\phi \triangleq \delta\phi^{(1)} - \delta\phi^{(2)}\ (f_1/f_2)$$

$$= 8.45 \frac{N_T}{f_1} [1 - (f_1/f_2)^2]$$

$$= 1.184\ N_T\ (\text{rad}) \tag{4}$$

isolates the ionospheric contribution to the instantaneous phase. In both Eq. (3) and Eq. (4), N_T is measured in units of 10^{16} el/m^2 and frequency is in gigahertz.

A commercially available GPS receiver provides a τ_d estimate every 1.5 s, with an rms error of 2 to 3 ns. From Eq. (3), therefore, absolute TEC measurements with an uncertainty of less than 10^{16} el/m^2 and a time resolution of one minute are readily obtained. The relative carrier phase can be measured with an rms error between 0.1 and 0.2 radians. This is adequate to detect the presence of weak amplitude scintillation at VHF. However, the performance can be considerably improved by using a high-gain antenna.

With the exception of the region near the geomagnetic equator and possibly the polar cap, L-band amplitude scintillation is rarely observed. Extensive data from the Wideband Satellite have shown, however, that measurable L-band phase scintillation always accompanies UHF or even VHF amplitude scintillation (Fremouw et al., 1979). It has also been shown that all the important parameters that are used to predict system performance (e.g., fade depth, coherence time, coherence bandwidth) can be computed from L-band phase and amplitude scintillation data (Rino, 1979 a,b; Rino and Owen, 1980; Rino and Owen, 1981; Rino et al., 1981).

A state of the art, dual-channel GPS receiver can, therefore, provide TEC and scintillation data that are useful for both scientific and engineering applications. Indeed, the global coverage of the GPS constellation make it uniquely suited for near real time ionospheric forecasting. To demonstrate these concepts, the Air Force Geophysics Laboratory (AFGL) had funded SRI International to procure a dual-frequency GPS receiver, and develop a microcomputer control and data acquisition system. The system is intended to be a prototype GPS ionospheric monitor station--at least insofar as the functions are concerned.

In this paper we report the results of preliminary data taking compaigns at Kwajalein, Marshall Islands (9.4 lat, 167.5 lon, 9° dip) during August 1980. Because our initial antenna system did not have a ground plane, the data were severely contaminated by multipath. The data were recovered by taking advantage of the fact that the multipath error behaves like a random excess delay component

and is, therefore, cancelled in the differencing operations defined by Eq. (3) and Eq. (4). More recently, an improved antenna system has been developed that appears to have eliminated the multipath contamination.

Our experience with the system to date leads us to conclude that the use of GPS for ionospheric monitoring is technically feasible, and it can greatly enhance our ability to perform ionospheric forecasting. The receiver and data acquisition system are described in Section II. The data are summarized in Section III. Our conclusions and recommendations are presented in Section IV.

II THE RECEIVER AND DATA ACQUISITION SYSTEM

The GPS receiver Model 5010 used in these experiments was designed and built by Stanford Telecommunications, Inc. (STI). The 5010 receiver accepts L_1 and L_2 signals from a low-noise preamplifier fed by an omni-directional antenna. Using an initial Doppler estimate (within 500 Hz), the receiver will acquire and track the clear acquisition code, decode the 50-bps data stream and extract the handover word, which provides a preliminary setting for the p-code. The receiver then acquires and tracks the L_1 and L_2 p-codes and carriers. At typical signal levels, the rms delay and phase jitter are less than 2 ns and 0.2 rad, respectively.

The navigation data, the receiver status, and an accumulated carrier-cycle count are communicated via a general purpose interface bus (GPIB), which also accepts receiver control commands and data. The L_1 and L_2 AGC levels, the outputs from number-controlled oscillators that track the L_1 and L_2 carriers, and the 1.5-second epic pulses from the L_1 and L_2 p-codes are also presented at various receiver output ports.

The GPS data acquisition system (DAS) was designed and built by SRI International. It consists of two HP-5370A universal-time-interval counters to measure the L_1 and L_2 "pseudo range" delay, an HP-2621-P terminal for the operator interface, a Kennedy 9600 tape drive with a Model 9218 formatter to digitally record the data, an HP-7402 chart recorder to display the data, and a central microcomputer to perform the control, interface, and data acquisition functions. A functional diagram of the system is shown in Figure 1.

The microcomputer is built around a Mostek Z80-A microprocessor and various input/output interfaces. A Ziatech ZT-80 is used to communicate with the HP-5370A counters and the 5010 GPIB. Analog data eter the computer via a Burr-Brown MF 8408/MP 8416 "multibus" compatible data acquisition system, which contains a multi-channel 12-bit A/D converter, and two 12-bit D/A converters.

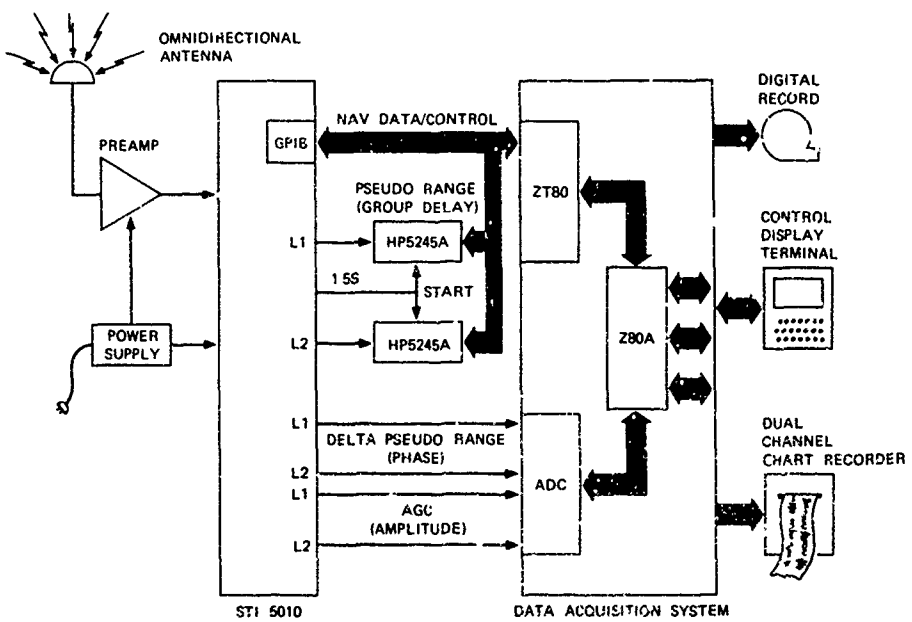


FIGURE 1 FUNCTIONAL DIAGRAM OF GPS IONOSPHERIC MONITORING STATION

The L_1 and L_2 AGC and accumulated Doppler channels are sampled at a 25-Hz rate--thereby generating 600 samples for each 6-s data frame. These data are written on a 9-track digital tape, together with the 300-bit navigation data frame, the accumulated L_1 and L_2 Doppler counts, four 1.5-s pseudo-range-delay numbers each from the L_1 and L_2 p-codes plus "housekeeping" and status information.

Preliminary off-line processing includes: (1) converting the raw AGC samples to signal intensity, (2) reformatting the navigation data, and (3) converting the Doppler count to continuous carrier phase. As it is recorded by the computer, the integrated Doppler count is given by the formula

$$I_k = \mathrm{mod}(8125t_k - f_d t_k, 4096) \quad , \tag{5}$$

where f_d is the Doppler shift as defined by the time derivative of Eq. (2), and t_k is the time of the kth sample in seconds. Because $t_k - t_{k-1} = 0.04$ s, f_d is nearly constant, and $f_d t_k$ is a good estimate of the instantaneous carrier phase as given by Eq. (2) directly. It is straightforward, therefore, to recover a continuous phase record from I_k.

Ultimately, we hope to perform all these routine manipulations in the microcomputer as the data are acquired, as well as make various time and delay corrections. The recorded data will then consist of L_1 and L_2 signal intensity (or amplitude) and phase, corrected pseudo range, IEC, and orbital parameters to determine the satellite position. For ionospheric monitoring applications, scintillation summary parameters could also be provided for real-time transmission to an ionospheric forecasting center.

III DATA ANALYSIS

The configuration of the initial 6 GPS satellites is such that only one satellite was visible from Kwajalein during the night-time period when the scintillation activity peaks. During the August 1980 observations, the satellite (GPS 8 or NDS 4) rose from the Southeast around 2250 local time and achieved its maximum elevation of 70° to 75° around 0330. Thus, most of the scintillation activity was already in its decay phase.

Data were recorded on 12 consecutive evenings beginning 21 August 1980. During seven of these evening data runs, the scintillation was severe enough that the GPS 5010 receiver could not maintain lock on the signal during the early portion of the pass. When the satellite was above ~ 30°, however, the scintillation was not severe enough to cause the receiver to break lock--although significant scintillation was observed as late as 0300 on Day 235. Only one evening showed sufficiently little activity that it could be regarded as quiet.

After reconstructing the amplitude and phase data (see Section II), the data were detrended as described in Fremouw et al., (1979). The detrender is a recursive digital high-pass filter that removes frequency components above a specified cutoff. In the intensity data, these variations result from antenna pattern variations and changing range to the satellite. In the phase data, a very large trend is induced by the $k\rho$ term in Eq. (2). In principal, this term can be removed analytically by using the satellite ephemeris data. Because the rate of change of ρ is comparatively small, it is more convenient to set the detrender cutoff low enough to remove the contribution of $k\rho$ in the detrending operation.

Processed amplitude data from a period that has two isolated patches of L-band scintillation is shown in Figure 2(a,b). For each patch, the peak value of the S_4 index measured over a 1-min interval is indicated. The original data were sampled at a 50-Hz rate, but filtered to 10 Hz to remove noise. The detrend filter cutoff was set at 1 min. Scintillation at L-band (of this type and intensity) was commonly observed during the Wideband Satellite operations at Kwajalein (Livingston et al., 1980).

The corresponding phase scintillation data derived from the "delta pseudo range" (DPR) count defined by Eq. (3) is shown in Figure 2(c,d). In the initial configuration of the 5010 receiver, the DPR count was only available every 6 s. As a consequence, it was necessary to use a much longer 300-s detrend interval. The most prominent feature in the data is the very large but short-lived phase excursions.

We have now established that these perturbations are due to multipath reflections from nearby reflectors. A PNR code will automatically reject echoes delayed more than one chip interval (97.57 ns or 29.32 m for the GPS p-code). Even so, a straightforward computation of excess path delay for perfect reflectors within 30 m of the antenna can account for the observed excess path-delay changes of 1 to 20 ns (10 to 200 rad) over time periods of one min. The detailed structure

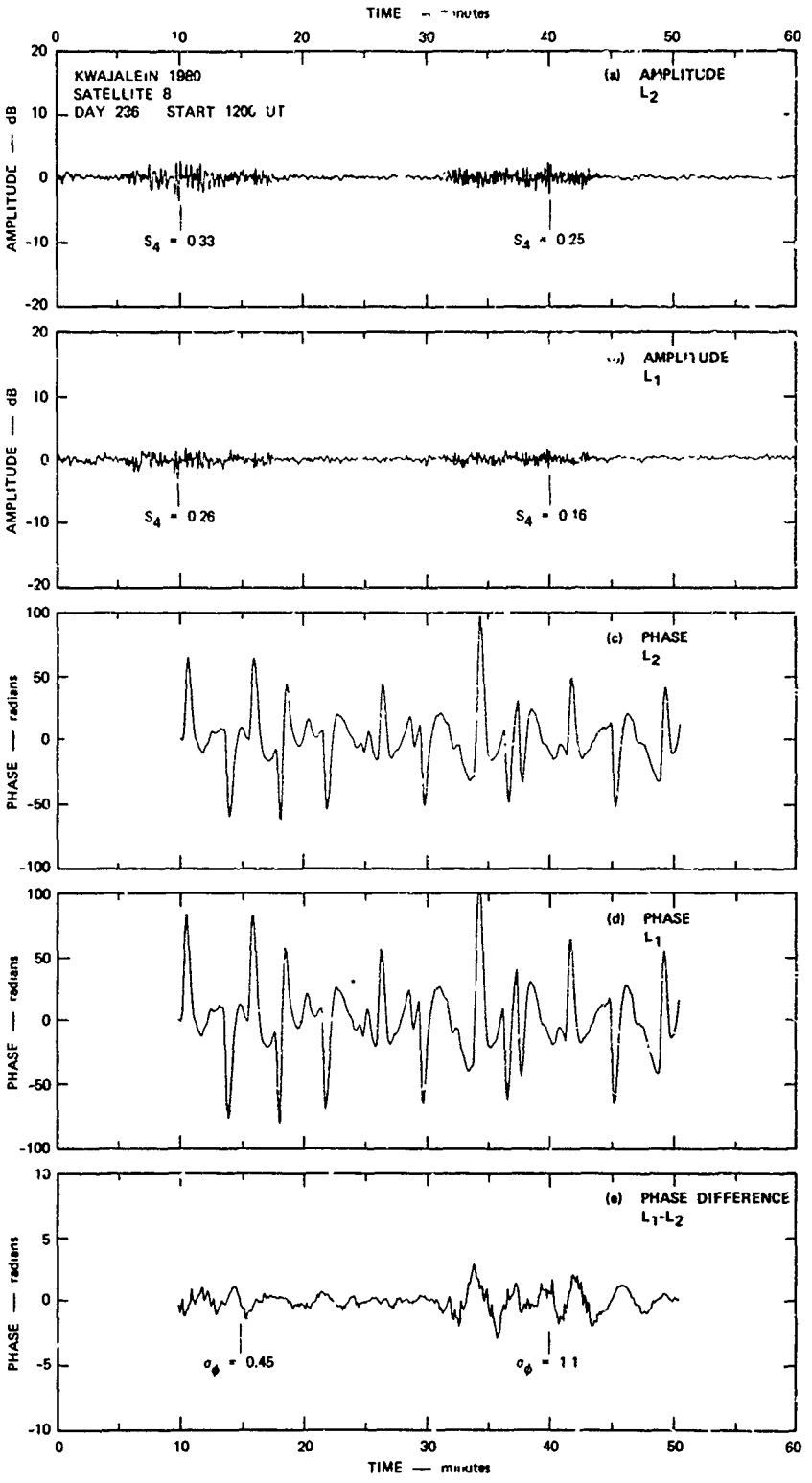


FIGURE 2 EXAMPLE OF MODERATE L-BAND SCINTILLATION PATCHES

of the pseudo-range, and delta-pseudo-range delay response to the multipath echoes is influenced by the receiver dynamics--but, it is readily shown from the data that the phase errors are linearly proportional to frequency.

Now, although such multipath contamination would severely impair the intended function of the GPS reciver, the fact that it scales linearly with frequency provides a means of eliminating it from the scintillation component, which scales inversely with frequency. The result of applying the difference operation defined by Eq. (4), to the reconstructed L_1 and L_2 delta-pseudo-range counts, is shown in Figure 2(e). The phase scintillation enhancements are now the prominent features in the data.

Figure 3 shows the amplitude and phase data for a more disturbed period. It should be kept in mind that the peak phase excursions and the rms phase level are critically dependent on the detrend interval, and, therefore, do not relate directly to the intensity scintillation level [Rino, 1979a].

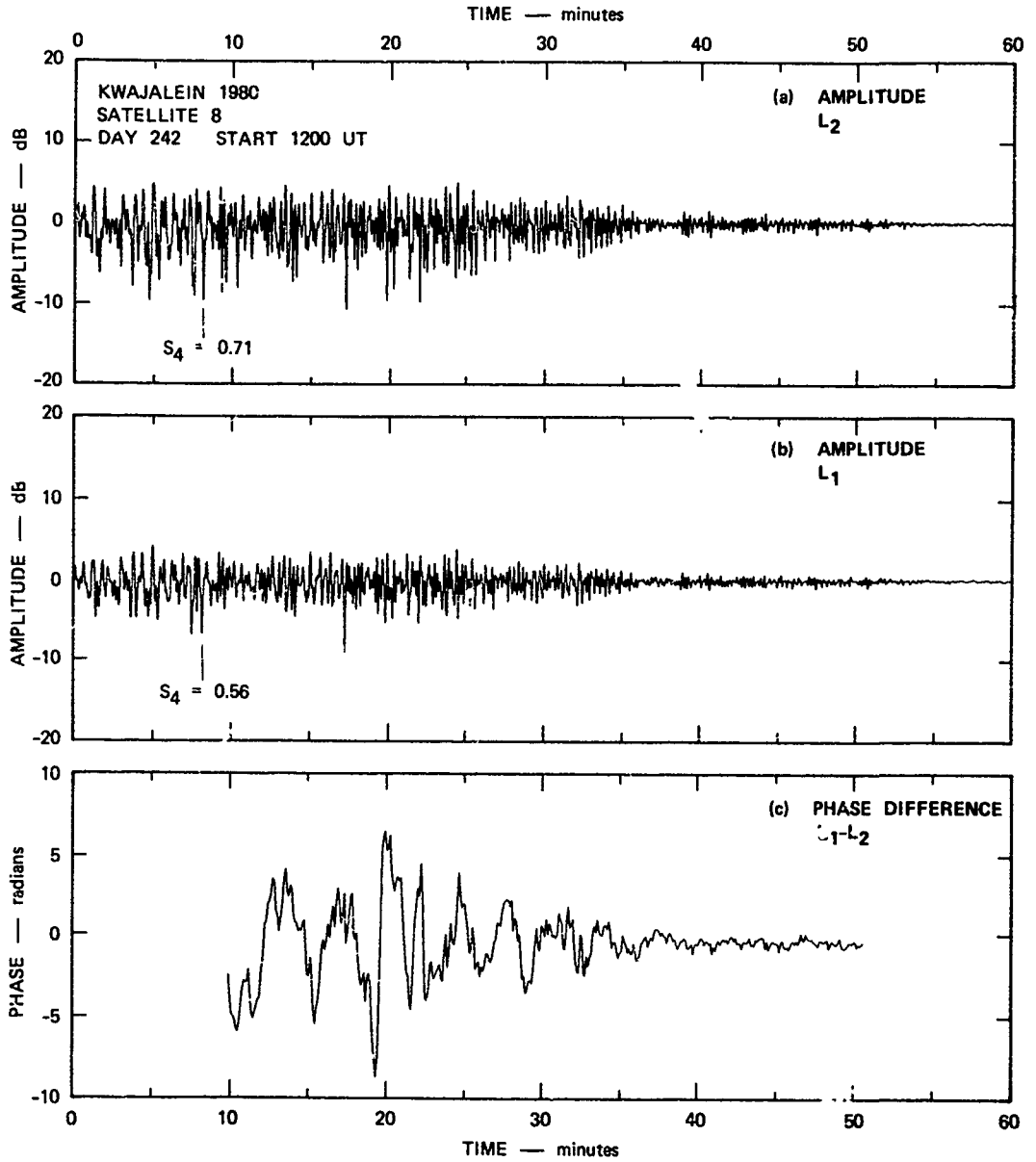


FIGURE 3 EXAMPLE OF STRONG L-BAND SCINTILLATION AND DECAY

The multipath contamination precludes a direct test of the frequency dependence of the rms phase. In Figure 4, however, a scatter diagram of the L_1 and L_2 S_4 indices for the amplitude data shown in Figure 3(a,b) is shown.

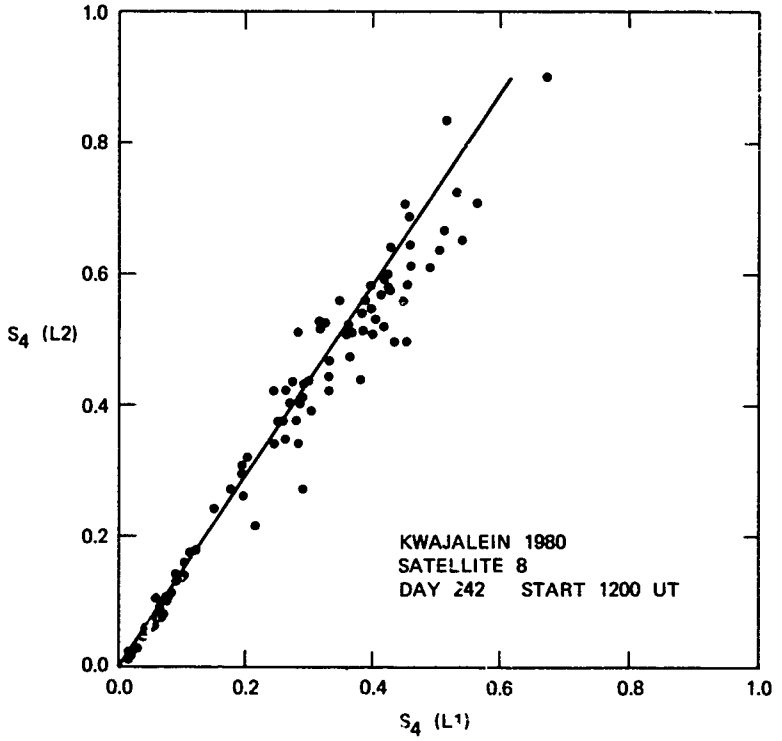


FIGURE 4 SCATTER DIAGRAM OF L_1-L_2 S_4 VALUES FOR AMPLITUDE DATA SHOWN IN FIGURE 2. Straight line is $S_4 \propto \lambda^{1.5}$

Under conditions of weak scatter, S_4 scales with wavelength as $\lambda^{(p+3)/4}$, where p is the power-law index of the phase spectrum (see, for example, Rino and Owen, 1980). The solid line in Figure 4 corresponds to p = 3. The fact that the data points fall slightly below the curve is consistent with earlier findings from the Wideband Satellite experiment--namely that the spectral p index is somewhat less than three, and varies systematically with perturbation strength [Livingston et al., 1981].

The preliminary data presented in this section are representative of the first amplitude and phase scintillation measurements from the GPS system. The results demonstrate the feasibility of using conventional GPS equipment for amplitude and phase scintillation measurements. Indeed, the phase structure can be processed to extract useful parameters, even when the corresponding amplitude scintillation level is lost in the background noise. The noise limitation can, of course, be improved considerably by using a high-gain antenna.

As discussed in Section I, to eliminate the multipath problem, the antenna has since been modified to accommodate a ground plane, which removes the back lobes and substantially reduces the antenna gain near the horizon. The receiver has also been modified to provide the delta-pseudo-range count from which the phase data are derived at a 25-Hz rate.

During January and February 1981, the modified system was operated at Ascension Island in the South Atlantic. A sample of detrended amplitude and phase data are shown in Figure 5. It can be seen that the multipath has been eliminated, and that the phase data have sufficient resolution to clearly show the phase noise. The Ascension Island data are being analyzed in detail to extract all the usual scintillation parameters, but these results will be presented at a later time.

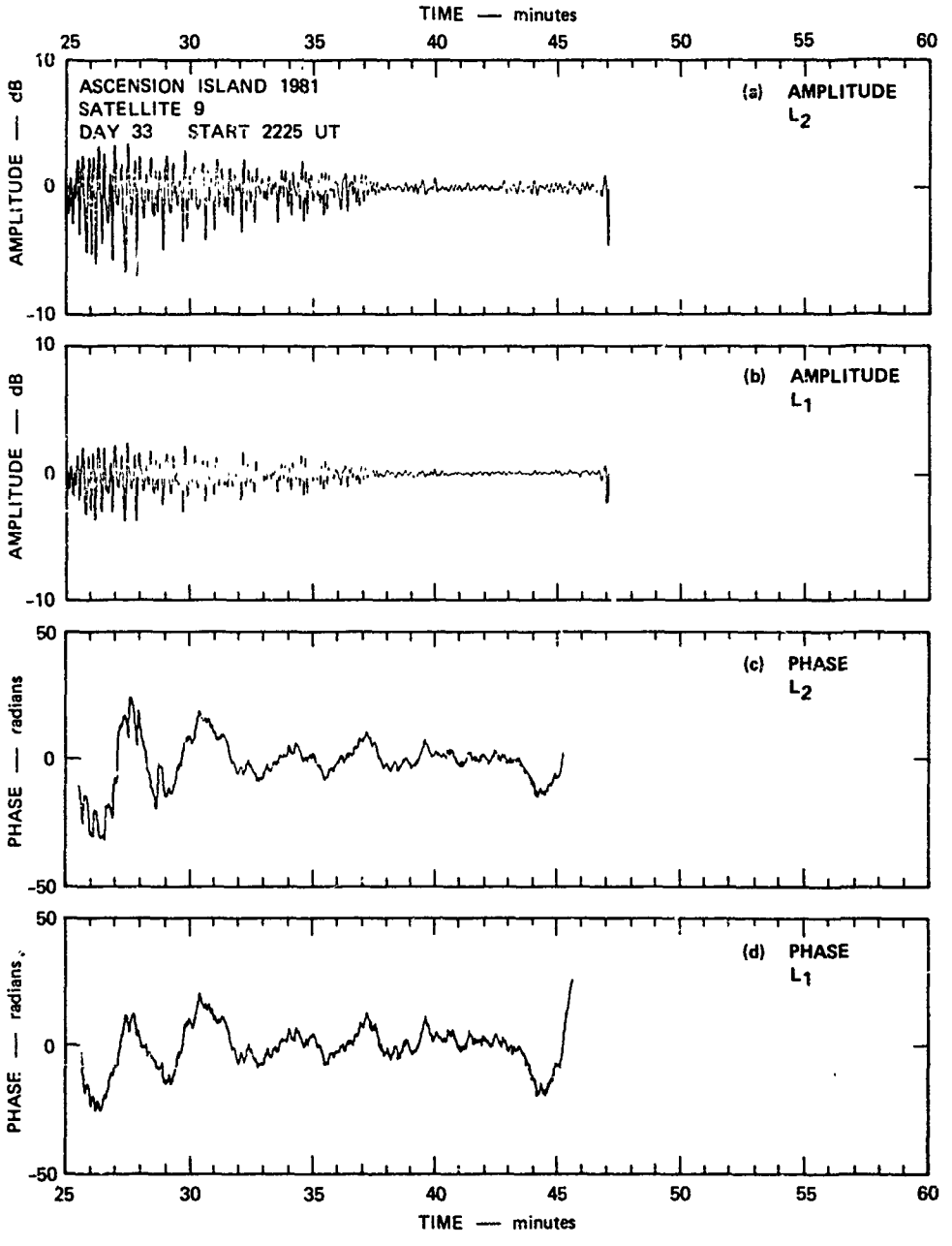


FIGURE 5 SAMPLE DATA OF DATA TAKEN AT ASCENSION ISLAND DURING JANUARY 1981, USING IMPROVED ANTENNA AND DAC

IV DISCUSSION

In this paper, we have presented the first amplitude and phase scintillation measurements using GPS satellites as a signal source. The AFGL objectives in supporting this work were to (1) demonstrate that conventional GPS receivers can be used to extract amplitude- and phase-scintillation data, and (2) develop a functional prototype of a GPS ionospheric monitoring station. The phase data are particularly important because the phase perturbations at L-band can be measured even when the corresponding amplitude scintillation is below the noise level. Thus, the global monitoring capabilities of GPS can be fully exploited.

The only serious problem encountered was multipath contamination, which can be eliminated by properly shielding the antenna. Indeed, this must be done if the GPS receiver is to perform its intended position-locating function. There are, however, a number of changes in the receiver and data acquisition system that should be made to improve signal acquisition and data processing efficiency.

The STI 5010 receiver, for example, does not maintain lock when $S_4 \gtrsim 0.8$ primarily because the receiver automatically reverts to an unaided search mode if the signal is below threshold for a preset period of time. Reacquisition would be greatly improved if the receiver "remembered" the delay and Doppler at the time of outage. As discussed in Section II, moreover, most of the off-line data analysis operations could be done in the DAC microcomputer.

The data recently acquired during the Ascension Island campaign are undergoing detailed analysis to demonstrate our ability to predict scintillation structure at lower frequencies. The group delay data are also being analyzed to extract absolute total electron content estimates. The results will be reported in detail in later publications.

ACKNOWLEDGEMENTS

The authors gratefully acknowledge the assistance of Robert Livingston and Nancy Walker of SRI International in developing the data analysis software.

REFERENCES

Fremouw, E. J., R. L. Leadabrand, R. C. Livingston, M. D. Cousins, C. L. Rino, B. C. Fair, and R. A. Long, "Early Results from the DNA Wideband Satellite Experiment--Complex-Signal Scintillation," Radio Science, Vol. 13, p 167, 1978.

Livingston, R. C., C. L. Rino, J. P. McClure, and W. B. Hanson, "Spectral Characteristics of Medium-Scale Equatorial F-Region Irregularities," J. Geophys. Res., in press, 1981.

Rino, C. L., "A Power Law Phase Screen Model for Ionospheric Scintillation. 1. Weak Scatter," Radio Science, Vol. 14, No. 6, pp 1135-1145, November-December 1979.

Rino, C. L., "A Power Law Phase Screen Model for Ionospheric Scintillation. 2. Strong Scatter," Radio Science, Vol. 14, No. 6, pp 1147-1155, November-December 1979.

Rino, C. L., V. H. Gonzalez, and A. R. Hessing, "Coherence Bandwidth Loss in Transionospheric Radio Propagation," Radio Science, in press, 1981.

Rino, C. L., and J. Owen, "The Time Structure of Transionospheric Radiowave Scintillation," Radio Science, Vol. 15, No. 3, pp 479-489, 1980.

Rino, C. L., and J. Owen, "On the Temporal Coherence Loss of Strongly Scintillating Signals," Radio Science, Vol. 16, No. 1, pp 31-33, January-February 1981.

IONOSPHERIC EFFECTS ON HF OVER-THE-HORIZON RADAR

Dennis B. Trizna and James M. Headrick
Radar Techniques Branch, Radar Division
Naval Research Laboratory

ABSTRACT

High frequency radar operates in the band of frequencies between 3 and 30 MHz, and as such, uses the refractive properties of the ionosphere to achieve long range coverage as far as 3000 km for a once refracted signal. The ionospheric path quality for such radar coverage depends very strongly upon ionospheric disturbances or the lack thereof, and upon knowledge of the ionospheric modes available. Originally developed for aircraft detection, more recent applications toward ship detection and remote sensing of ocean surface conditions impose fairly severe constraints on Doppler spectrum quality, and the attendant ionospheric path analyses. A general review of HF radar work at the Naval Research Laboratory was published recently by Headrick and Skolnik (1). In this paper, some of the constraints imposed by the ionosphere are discussed, including limitations due to details of ionospheric structure under undisturbed conditions, as well as problems imposed by a variety of types of ionospheric disturbances. Examples of effects on radar displays are shown for several of such cases.

INTRODUCTION

The development of high frequency (HF) over-the-horizon (OTH) radar for purposes of detection of targets at very long ranges depended upon two important points: 1) the short term stability of the ionosphere over time periods comparable or longer than integration times required to allow separation of targets from clutter in the Doppler spectrum; 2) the ability to develop high dynamic range hardware through the entire radar and processing chain. High dynamic range here refers to the ability to detect a very small signal at one Doppler frequency in the presence of one or more high level signals at different Doppler frequencies; i.e., the entire system must be linear over the entire amplitude range in which one expects to find signals of interest, as much as 90 decibels. This second requirement is necessary because for slow moving targets, the clutter from the sea or land cannot be simply filtered since its spectral bandwidth includes frequencies which are the same as the targets in which one is interested. With the development of the hardware capability, one must then contend with ionospheric effects which can occur over a similar dynamic range of amplitudes. These effects can be present under qualitatively "good" propagation conditions, and if properly understood, can be dealt within the processing of radar data As propagation conditions become worse, for a variety of reasons, one finds that different types of ionospheric disturbances will have different effects upon the radar Doppler spectrum, and hence upon the radar displays. If one recognizes and understands the type or source of the disturbance, one can quite often modify one's operating parameters to contend with the disturbance, or at least minimize its effects over a significant region of Doppler-range space available.

In order to demonstrate the effects of several types of ionospheric disturbances, it is worthwhile to spend some time describing the typical radar displays used with the HF radar, and then identify the degradation in quality for different types of disturbed ionosphere. The HF radar is not a scanning radar in the sense that a microwave radar is, with a moving antenna which rotates to cover the azimuthal dimension. Operating with a typical one degree beamwidth, at the several hundred to several thousand kilometers in range covered by an HF OTH radar, the beam is sufficiently wide to allow a few tens of kilometers wide a cell. A similar size is typical for the range dimension as well. Although this seems like quite a large area upon first consideration, the relative uncertainty in position of the target, determined by the ratio of the cell length to the range of detection is of the same order as for microwave radars. However, the clutter created by such a large area is much greater than the cross section of a typical target of interest, and one must employ Doppler processing of the radar signal to separate the two, in distinction to the microwave case in which the clutter and target are more nearly equal. Doppler processing requires that the transmitted signal be phase coherent on a pulse to pulse basis, that is, the phase of the signal be the same from one pulse to the next. Rather than display the returned signal for each pulse, a series of samples of pulse returns for each of a set of time delays is stored in a memory, and the Fourier transform of the series is taken. Although the complex transform is useful, typically the absolute magnitude is taken to determine a power spectrum, called the Doppler spectrum: signal amplitude versus Doppler frequency, which can be translated to cross section versus target velocity. These spectra are then stored for each of the series of range-time delays, thereby determining a three dimensional matrix of amplitude-Doppler-range.

Research radars typically present the three-dimensional matrix of information two dimensions at a time, primarily for quantitative reasons, although color displays would be ideally

suited for qualitative viewing. Of the three possible two-dimensional projections, the most informative display to consider first is the amplitude-Doppler spectrum for a given range bin, and example of which is shown in Fig. 1. It represents the spectral content for all scattering surfaces existing within the rather large volume resolution cell for a given time delay Naval HF radar applications typically involve illumination of the sea surface, and Doppler spectra of sea scatter are primarily encountered. Scatter from the sea surface is characterized by two very strong returns, displaced either side of zero Doppler by equal fixed amounts, called Bragg lines The amount of Doppler shift is determined by the phase velocity of the ocean waves responsible for the scatter, those with wavelengths one half the radar wavelength, the Bragg resonant waves. The magnitude of the shift in Hertz is given roughly by 0 102 times the square root of the radar frequency in megahertz. This Bragg scatter mechanism is the surface analogue of volume scatter responsible for so-called incoherent scatter from refractive index fluctuations of free electrons within the ionosphere. Second order effects also occur in the surface scatter case due to double scatter from pairs of ocean waves satisfying the conservation of energy and momentum equations defining Bragg scatter, and is typically 25 to 50 decibels down from the first order Bragg scatter, observable here because of the high dynamic range of the instrumentation. All of the scattered power is contained in a region roughly half a Hertz either side of zero Doppler, although the exact spectral spread is determined by the radar operating frequency. Figure 1 shows an example of highest quality sea scatter collected using the most stable E-layer mode of propagation, and includes ground backscatter at zero Doppler as well. Spectral degradation from this example will occur due to ionospheric disturbances of various types

A second useful display of the three dimensional matrix is that of Doppler-range, thresholded for some fixed amplitude.

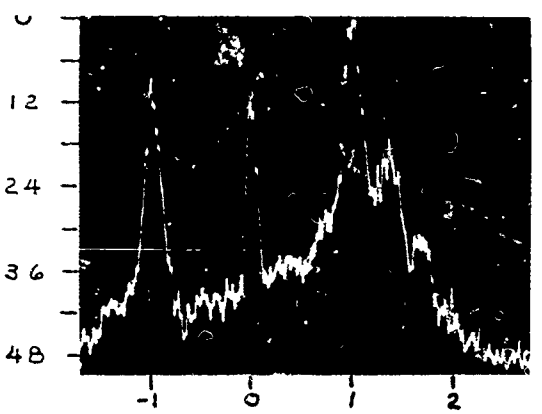


Figure 1. An amplitude-Doppler power spectrum for a given delay time is shown. This represents one of the three two-dimensional projections of a processed amplitude-Doppler-range matrix of radar data. This example shows a typical high quality spectrum collected using a stable E-layer propagation mode with an HF radar. The amplitude scale is in arbitrary decibel units, while the frequency is in normalized Bragg frequency units described in the text. The two peaks at ±1 are the first order Bragg lines, with land backscatter at zero Doppler, and a calibration signal at about 2.2 units Other low level signal around zero Doppler and outside the right Bragg line represent second order sea scatter also described in the text

The display is a binary one, black if a spectral point lies below the threshold and white if it lies above. If a set of amplitude-Doppler spectra of Fig. 1 were collected for a series of consecutive ranges to form a Doppler-range display, thresholded a few dB below the lower of the two Bragg lines, the resulting display would consist of a pair of broad lines centered about zero Doppler, extending throughout all ranges. Typically, the Doppler spectra vary from one range to the next, due to changing sea conditions, with one Bragg line becoming greater than the next, and perhaps both growing in amplitude in a like manner relative to earlier ranges. The resulting Doppler-range display for such a more realistic case would show two range-variably broadened lines with changing widths, as is shown in Fig. 2a. As one lowers the threshold below the level of the second order scatter amplitudes around zero Doppler, the sea clutter appears as one broad swath around zero Doppler as a function of range, as in Fig. 2b. As one continues to lower the threshold nearer the noise floor, the black region of the display begins to speckle as the noise begins to show, as in Fig. 2c This last threshold level is typically useful for finding low level targets near the noise floor which are far from the clutter return in Doppler. Figure 2a might be used to assist in determining propagation path quality and for remote sensing applications

A third display to consider is that of amplitude-time for a given Doppler frequency. Although not shown here, it is useful for determining the fading characteristics of a target or clutter and aiding again in propagation analyses. It is simply a plot of the spectral output of one Doppler frequency as a function of time.

A fourth display using the three dimensional matrix of data projected into a fourth dimension, time, is the Doppler-time display In appearance it is very similar to the Doppler-range display. It is constructed by taking one arbitrary range interval from the Doppler-range display, and projecting several consecutive samples of this display in tandem, so that the horizontal axis represents time rather than range, again thresholded for a given amplitude. It is very useful for establishing tracks of target Doppler histories. Traveling ionospheric disturbances are also very apparent on such a display, as the clutter return exhibits an oscillatory behavior with time. With several different range extents simultaneously displayed in such a manner, one can determine propagation speeds of such disturbances as they propagate through the region of coverage. Figure 3 shows an example of the Doppler-time display for a case of a traveling ionospheric disturbance, probably created by passage of a gravity wave through the ionosphere. With this introduction to the displays used most often to analyze HF radar data, we now procede to examples of different effects that the ionosphere can have in degradation of radar data from the high quality examples which have been demonstrated. These shall be divided into examples of two types: the first group will consider cases which occur for a typical undisturbed ionosphere and which are due in some part to the design of the radar, in particular, the transmit and receive antennas; the second group will consist of cases in which various types of disturbances of the ionosphere impact the spectral quality of the data.

Included in examples of ionospheric effects on HF radar which are due to the undisturbed ionosphere are the following: natural f-region turbulence; multipath, due to vertical layer structure of the ionosphere; blanketing and non-blanketing sporadic-E; and the mid-latitude trough. Examples of effects of the disturbed ionosphere include aurora; meteors; traveling ionospheric disturbances; sunrise and sunset at ionospheric altitudes; spread-F, and equatorial ionospheric irregularities created after local sunset.

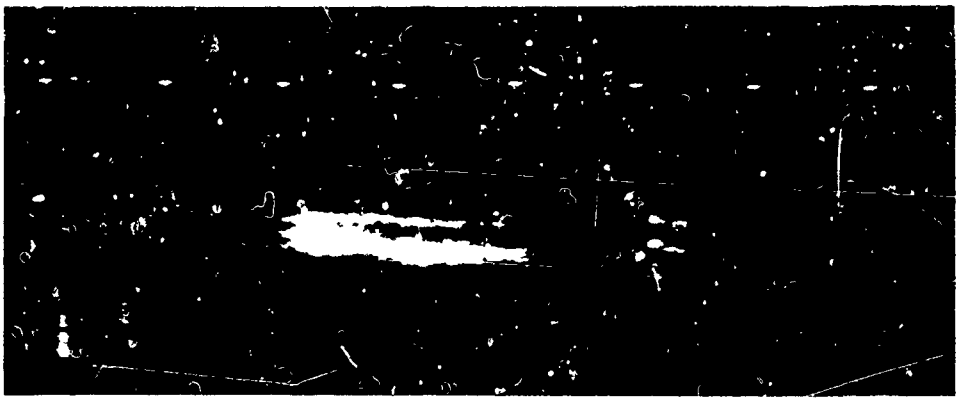


Figure 2a. This is a Doppler-range display, for a threshold chosen far above the noise level, near the peaks of the clutter. It represents the second two dimensional projection of the three dimensional matrix of processed radar data. Is is useful for determining path quality and for remote sensing applications The two swaths represent the two Bragg line returns from the sea surface, the broader swath indicating the higher amplitude Bragg line The ovals at the top of the figure are calibration signals set at 100 n mi and each 450 n mi thereafter. The entire Doppler width shown is roughly 8 Hertz

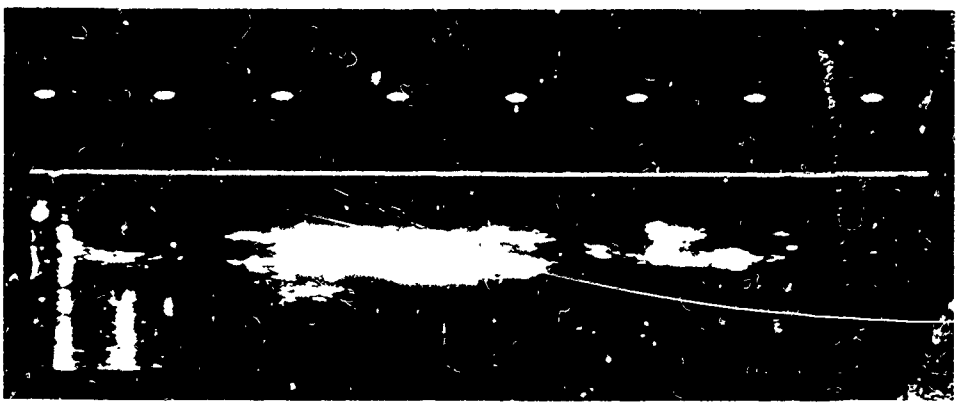


Figure 2b This is a second version of 2a, with a threshold 12 db nearer the noise floor than 2a, so that second order scatter from the sea surface is now apparent between the Bragg lines of first order sea scatter. Second hop scatter is seen at the sixth calibration marker at 2350 n mi The very narrow returns seen between the second and third calibration markers at about 750 n mi is probably due to sporadic E coverage because of the narrowness of the lines. Some meteor returns are seen at the shortest ranges as primarily negative Doppler returns, indicating receding velocities The narrow range swath between the clutter and the calibration signals is a 60 cycle hum line.

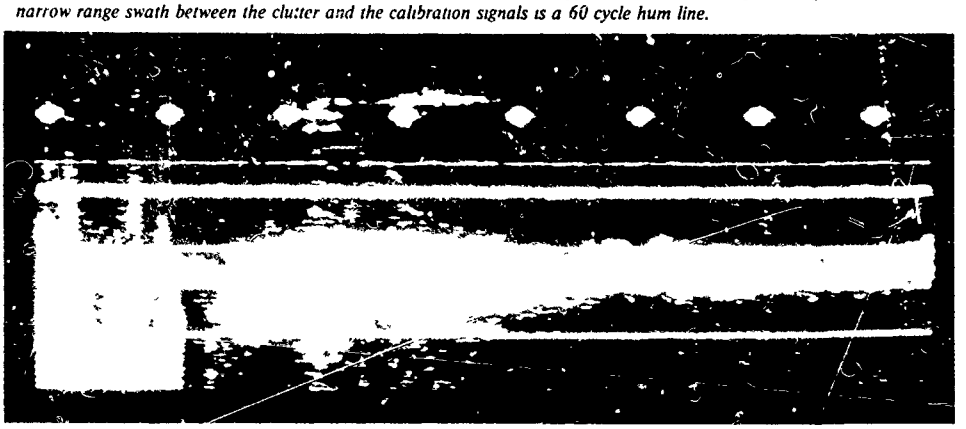


Figure 2c. A third version of 2a is shown for a threshold set very near the noise level, showing up as the speckled area. This version of the display is useful for detection of targets near the noise and away from the clutter As this is exactly the same data as displayed in 2a and 2b, lower level meteor returns are now more apparent. Harmonics of the 60 cycle hum lines are causing the additional swaths on each side of zero Doppler

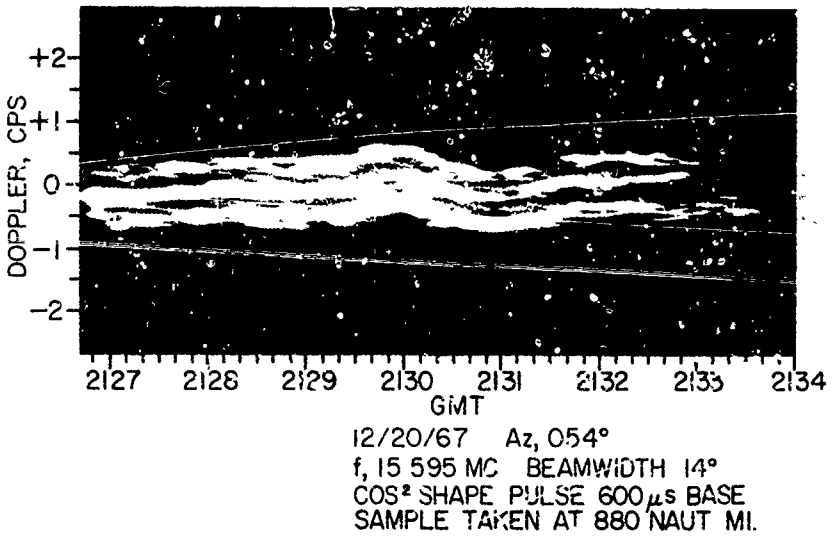


Figure 3. A different display from the two dimensional projections of the data matrix is that of the Doppler time history. It is created by taking a small range extent of the Doppler-range display of Figure 2, and displaying it in consecutive fashion for a series of successive coherent integration periods, time running left to right, with axis scales indicated. This display shows a traveling ionospheric disturbance beginning at 21:29:30 UT as a slight Doppler oscillation.

THE UNDISTURBED IONOSPHERE

Before giving examples of spectral degradation of HF Doppler backscatter spectra, it is worthwhile to devote a few words to a physical interpretation of radar backscatter, particularly how it differs from the case of ionospheric propagation of a point-to-point one way propagation path. An example of the latter under very good propagation conditions is shown in Fig. 4, for a path between Cyprus and England. Thirty two spectra of 3 2 second integration time were averaged to produce the signal spectrum shown The main signal is at −40 Hertz Doppler; its image due to *I* and *Q* intermodulation is at +40 Hertz; the term at zero Doppler is a d.c. offset in the processor. As is seen, the dynamic range which the ionosphere allows for point-to-point paths is very high, of the order of 75 dB. The distribution in Doppler is Ricean for short integration times of the order of 3.2 seconds. a very strong spectral peak with much smaller spectrally spread shoulders some 70 dB below. These latter are the result of scatter paths from angles off the great circle path between the transmitter and receiver, as well as via higher and lower paths in elevation angle. The primary path is very clean, yet upon increasing the integration time it will split into two signals, the result of ordinary and extraordinary ray multipath effects. This is due to the fact that a linearly polarized wave splits into a right and a left handed circularly polarized wave due to the earth's magnetic field. Each of these in turn traverses slightly different paths between points, with different phase paths and phase change rates, creating a very slight Doppler frequency difference in the signal received via the two paths.

F Region Turbulence

The case of radar sea clutter propagated via an ionospheric path is a more complicated problem, in that it represents a convolution of many one way paths A wave front is transmitted

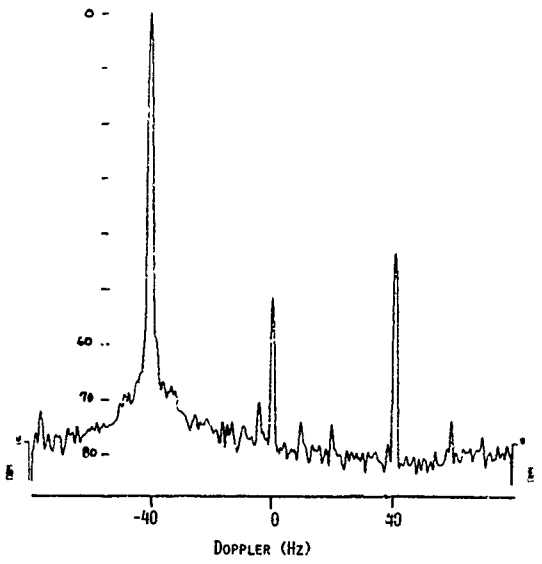


Figure 4. This is an amplitude-Doppler display of a one way path signal from Cyprus to England via an F-layer mode. Thirty-two successive 3.2 second conerent integration periods were averaged to produce the display. The strongest signal at -40 Hertz is the main signal, with the other two strong signals at zero and +40 Hertz processing effects, d.c. offset and the I&Q phase nose image, respectively. This is an indication of the high dynamic range signal the ionosphere will allow under good conditions. Minor scatter and multi-paths are shown some 70 db below the peak of the main signal as broad shoulders.

which is phase coherent across its face, and illuminates the ionosphere over an area determined by the pulse length and the antenna beam width. For a one degree beamwidth and scatter from 2000 km range, the ionosphere is illuminated at 1000 km, resulting in cell of the order of 15 km wide. Pulse lengths appropriate to this dimension are typical, resulting in a square reflection area 15 km on a side. This is much larger than the scale sizes of F-region turbulence so that the transmitted wave encounters phase changes which vary across the phase front of the wave. Assuming for the moment that the earth's surface does not impose any additional phase shift, the process is repeated in the return path back to the radar. If the ionosphere were not time variable, the resulting return signal would present a vector sum of these various paths to produce a very narrow spectral backscatter return, with no indication of the phase structure the ionosphere imposed upon the return wave. However, turbulence at F-region altitudes, and far less so at E-region altitudes, is time variable, resulting in a time rate of phase change, and therefore Doppler broadening of the backscatter signal. The statistics of such broadening are not Ricean as defined earlier. Rather the spectral peak tends to break up into several nearly equal amplitude contributions with Doppler processing of 100 sec integration times and longer. The paths responsible for these contributions are available throughout the entire integration period since they have spectral widths of the order of the Doppler filter determined by the integration time. If the paths were present for just a small fraction of the integration time, their contribution to the spectrum would be broad with a width the reciprocal of the time they were present. The statistics of this broadening due to what shall be identified as F-region turbulence broadening have not been considered in detail regarding their diurnal and seasonal behavior.

Multipath Due To Ionospheric Layers

The next level of severity of spectral degradation under normal ionospheric conditions to be considered is due to multipath because of ionospheric layering. With a wide antenna pattern in elevation angle, as is the case with most HF radars which do not have much height extent in their antenna structures, energy transmitted as a function of elevation angle is ideally expected to scale to range is a uniform manner. The lowest elevation angles will transmit energy to the farthest ranges, the highest that do not penetrate the ionosphere will transmit to the shortest ranges, assuming a mirror ionosphere at a fixed altitude. In reality, the various layers of the ionosphere can cause energy transmitted at a higher elevation angle to be time delayed and refracted in a non-uniform manner with elevation angle. The net effect of this behavior is to create group path delays sufficient to create an additional time delay beyond that expected for a mirror ionosphere. The returned energy is therefore superimposed in time delay upon energy which was reflected from a much longer ground range, sometimes with a different Doppler shift because of a different phase change rate along the different path traversed. The resultant spectrum appears as the superposition of two spectra shifted relative to one another. This type of problem is particularly prevalent with transmit-receive antenna combinations which have an elevation angle pattern which is very broad, with significant gain at higher elevation angles, such as vertical monopoles. Two solutions are available to this problem: one, confine the energy to very low take off angles with appropriate antenna elements and with a somewhat shorter total range coverage available, variable by changing operating frequencies; second, create an antenna system which can be steered in elevation angle, by adding vertical structure to it as well, such as horizontally polarized dipoles placed along a tower.

Even with these attempts at easing the difficulty, the problem can be a severe one, requiring path analysis using ray tracing techniques (2) or virtual height equivalents and simple reflection (3) to interpret clutter returns. In either case, ionospheric profiles which vary with latitude and longitude and the radar's antenna patterns are used as inputs. The ray tracing then indicates the paths each ray for a given elevation angle actually take, in terms of the total delay time experienced, the actual incremental path through the ionosphere including exit angle, angle to magnetic field lines along the path, total phase path relative to that transmitted, transmission losses, and final signal amplitude as a function of time delay. The virtual height-simple reflection analysis uses similar inputs, but just outputs signal amplitude versus range, with a label of the ionospheric layer traversed for each amplitude as a function of time delay Both of these techniques have been developed to a very high degree of accuracy over the past several years, and the greatest uncertainty today is in modeling the transmission losses encountered in propagation through the D-region of the ionosphere An example of the type of latitude-longitude variable ionospheric profile which is used as input into these propagation programs is shown in Fig. 5. Lines of plasma frequency contours are shown as a function of time delay or range for a given radar bearing. Such data sets have been established as median predicted values on a grid basis over the entire globe, for each month of the year with a three hour increment, with a solar activity index input as a scaling parameter for modifying electron densities With one or more real time ionograms available as well, the mean value table can be approximately scaled for yet a better description. An example of the simple reflection analysis outputs of amplitude versus time delay and propagation modes versus elevation angle and time delay are shown in Figs. 6a and 6b, with the amplitude contribution for different modes indicated for each time delay. Such analyses provide a very powerful tool in providing path analysis for HF radar operation.

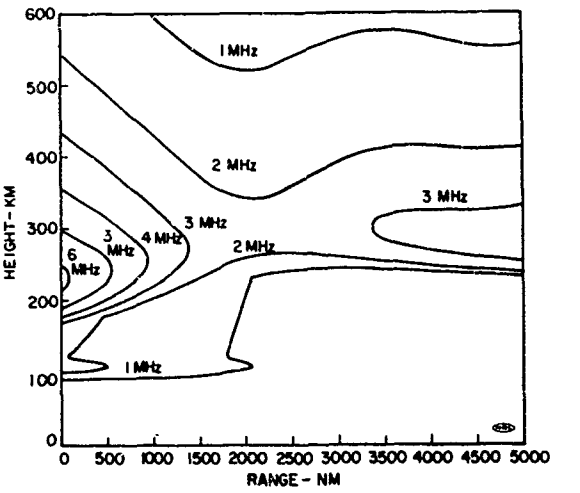


Figure 5. An example is shown of one of the inputs used for ray-trace and path analysis software with HF radar data: a two-dimensional projection of a three dimensional ionosphere, along a ray path for some unspecified radar bearing. The models are based upon long term averages of data and can be modified by changing a solar index to increase or decrease plasma densities.

Sporadic E-Layer

A sometimes bothersome element of the undisturbed ionosphere, perhaps ought to be associated with the slightly disturbed ionosphere as it is a function of weak solar activity, and

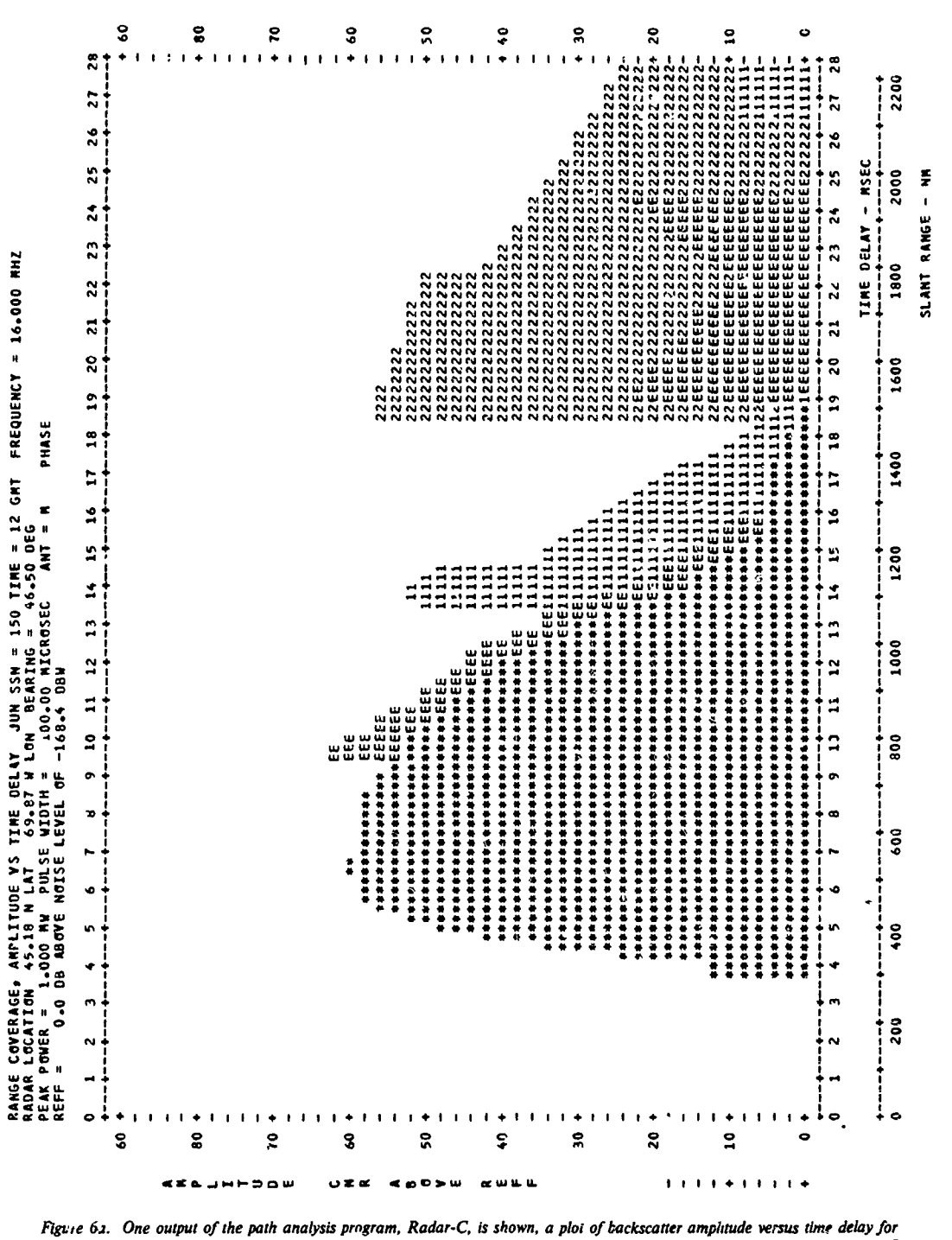


Figure 62. One output of the path analysis program, Radar-C, is shown, a plot of backscatter amplitude versus time delay for the radar antenna system and transmit power of the USAF HF Experimental Radar System. The asterisk refers to sporadic E modes; E, to normal E modes; 1 and 2, to F_1 and F_2 modes, respectively. For some ranges, two modes are available with different amplitudes. For a different antenna vertical pattern, the relative amplitudes of these contributions would be different for a given range delay as discussed in the text.

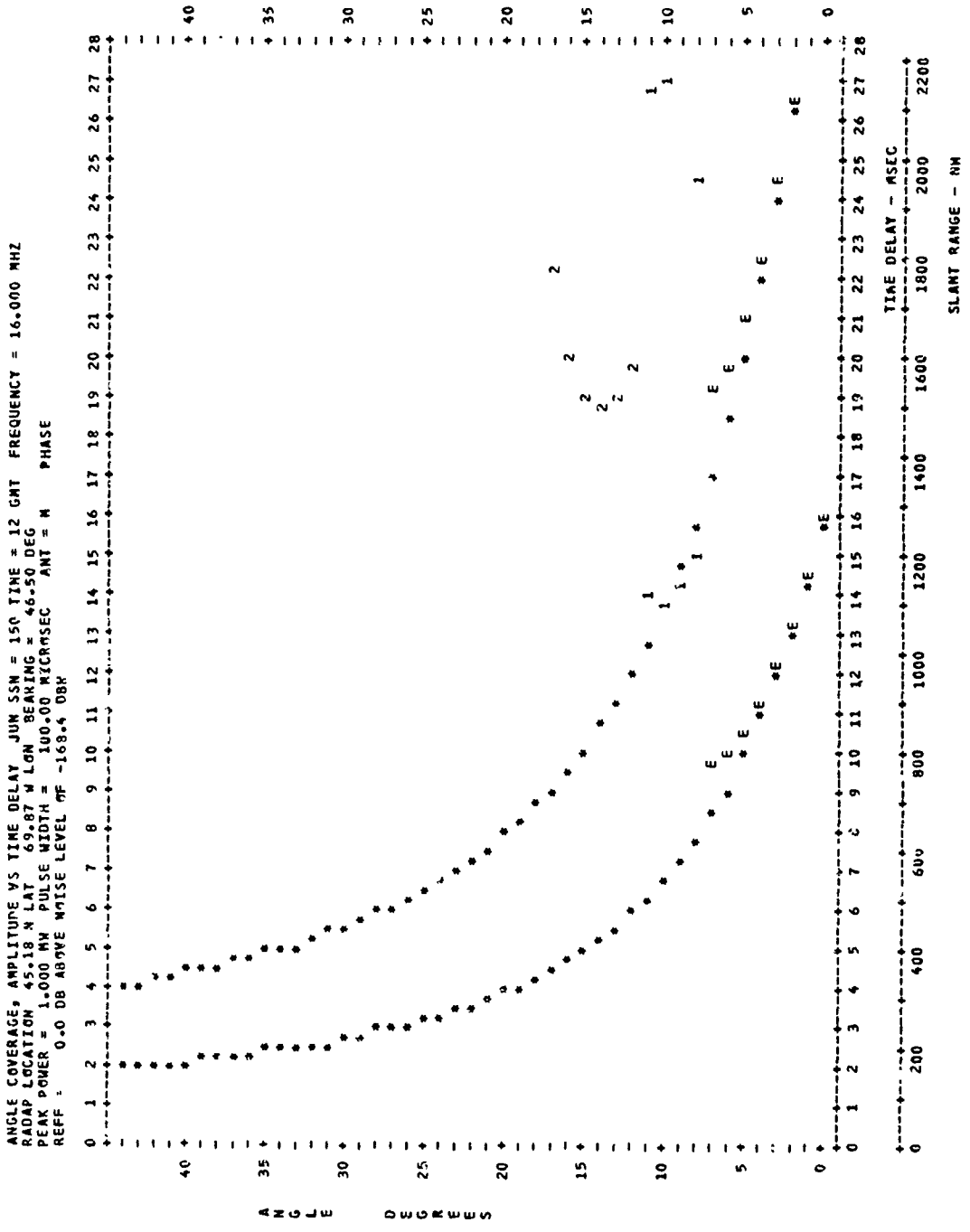


Figure 6b. A second output of the program is shown. For the same radar parameters input, the coverage modes are indicated for each elevation angle with the same annotation as in 6a. The resultant time delays for the different modes are indicated. For some elevation angles, two modes may be available, with the mode strength determined by the antenna elevation angle pattern, and indicated in 6a. Both 6a and 6b should be used in parallel for path analysis.

that is the sporadic E-layer. Some words are in order about the normal E-layer first, since it is quite often dismissed as an absorbing and partially refracting layer, but not recognized as responsible for full refraction as a mode of propagation in itself. This can be the image created by operating experimental radars with broad elevation angle coverage in which the F-layer modes dominate and overshadow E-layer modes. Experiments have been conducted at NRL using the MADRE HF radar, which has elevation angle flexibility, exclusively using E-layer mode propagation modes (4). The antenna consists of two rows of horizontally polarized dipoles one row above the other, which can be fed in phase or 180 degrees out of phase relative to one another. In one case a very strong lobe is created at an elevation angle at 6 degrees with a minor lobe at 18 degrees; in the other case the strongest lobe appears first at 13 degrees with a second at 25 degrees. In the first case range extent can be confined to 2,000 km maximum for some choice of operating frequencies, early morning operations, and for eastward looks. Propagation paths are achieved as predicted for a single E-layer ionosphere, with very crisp spectral quality of backscatter. In the second case an F-layer mode is apparent, with clutter received from beyond that expected for E-layer coverage only, in addition to short range. In this case the quality of clutter for the shortest range is very high, whereas the mid and longer ranges indicate some ionospheric turbulence of the sort one might expect for F-region modes. Hence, for radars with broad transmit-receive pattern which allow an even greater spread in elevation angle than MADRE, the E-layer mode would be generally overwhelmed by high angle F-layer modes and therefore not judged as significant as a high quality path ever. Therefore, antenna elements with low takeoff angles are preferred for unambiguous range coverage which allows separation of E and F-layer modes, and, in particular, use of the high quality normal E-layer modes. This would be the case for radars with remote sensing applications as an example, when multi-path in range presents confused data. For target detection applications, continuous coverage with range is desired, regardless of the mode of propagation, as long as path analysis techniques are employed for accurate range determination.

Sporadic E-layer is a term derived from an ionosonde or a radar oblique backscatter record. From an ionogram, it is characterized by strong reflection of radio energy from a fixed low altitude, over a frequency range which can extend beyond the F-layer maximum plasma frequency. Within this frequency range, it can disappear and reappear several times, in sporadic fashion. For blanketing conditions, it will remain strong with increasing radio frequency without disappearing, and with very little energy penetrating to F-region altitudes. On an oblique backscatter record, energy is confined to 2,000 km, with very weak or no one hop scatter via F-region modes, for blanketing condition. From the radar standpoint, it can vary dramatically in terms of the intensity of its effect. A prominent mechanism for its production is thought to be wind shear at E-region altitudes. It extends over a much narrower altitude range than normal E, is far more intense in terms of the electron density, and is typically patchy in its horizontal structure. As a result, the radar energy which propagates through the sporadic E-layer is variable inversely as the strength of the sporadic E. Path analysis and prediction models account for its presence by modeling it as purely reflective over the areas over which it exists, and accounting for its intensity by a transmission coefficient, which allows an appropriate fraction of the energy to propagate through it unrefracted except by the normal E-layer and high altitude F layers. Under its most severe conditions, the blanketing case, the transmission coefficient is zero and the layer becomes virtually a mirror. Under such conditions, sporadic-E becomes very apparent to radars using broad elevation angle coverage since there is a sharp cutoff in one hop propagation at the predicted 2,000 km, with no coverage available via F-layer modes to longer ranges. When it is present to a severe degree it is very debilitating operationally for any radar. In any case, the understanding and prediction of sporadic-E is important to HF radar operation and should receive appropriate priority in ionospheric research.

The Mid-Latitude Trough

Another ionospheric anomoly which which exists under normal undisturbed conditions and which should be considered in path prediction and analysis is the mid-latitude trough, a region of electron density depletion associated with the auroral oval. It is the boundary between two regions in latitude which have different F-region ionization production mechanisms during nighttime hours, particle precipitation in the auroral region versus the normal nighttime F-region. A width of roughly 500 km is associated with it, and it can be quite detrimental to HF radar operation. Radar energy which propagates into the trough region will experience a weaker refraction there, with the primary effect being a bending of the path of the radar energy off the intended great circle path, for cases when the energy is incident at angles not perpendicular to the trough. This will introduce obvious uncertainties into the position of the area intended to be covered by the radar, and has been proposed as the reason for unexpected results of some radar experiments in the past. Knowledge of the structure of the trough and its temporal behavior should be considered as very important to the operation of HF radars operating in the trough regions.

THE DISTURBED IONOSPHERE

Radar Aurora

In addition to the structural anomolies of the undisturbed ionosphere, there are various types of disturbances of the ionosphere which can have deleterious effects on radar Doppler spectra. The first of these to consider is radar aurora. While optical aurora is characterized by emission of energy at optical frequencies due to ionization of molecules by solar particle precipitation along magnetic field lines, radar aurora is defined as radar scatter from field aligned irregularities in the ionosphere due to the particle ionization. The radar energy scattered from the aurora can have a very wide range of chracteristics, depending upon the type of transmission format used by the radar. We shall treat in detail only the case of a simple pulse radar, as opposed to pulse compression radars of FM-chirp pulse or FM-CW formats, the effects of which can be predicted. Typical returns from the radar aurora are confined to a finite range extent based upon the physical volume occupied by the field aligned irregularities responsible for the scatter, typically fifty to a few hundred kilometers at most. The Doppler characteristics can be quite variable, in spread or bandwidth of frequencies scattered as well as the central frequency of the frequency band. Because ambiguities may arise both in Doppler and in the range dimension, it is worthwhile discussing the idea of radar ambiguity.

A pulsed radar transmits pulses at regular intervals in time. If the interpulse period transforms, for example, to a thousand miles in range delay and a target is observed at a hundred miles, there is an ambiguity as to whether this range is truly a hundred miles or eleven hundred miles, or any multiple of a thousand miles plus one hundred. One could double the interpulse period, doubling the unambiguous range to two thousand miles, and so on, but this would have negative effects upon the Doppler dimension. That is, the reciprocal of interpulse period is the pulse repetition frequency (PRF), and this number, in Hertz, is the maximum unambiguous Doppler

bandwidth that one can achieve. Hence if one were to change from a ten Hertz PRF to five Hertz to improve on range ambiguity, one suffers a degradation in the Doppler ambiguity by a factor of two. One does not know then whether the returned signal is at a 2.5, 5, 7.5, 10, etc, Hertz Doppler offset, with appropriate uncertainties in target speed. A one hundred millisecond interpulse period for a ten Hertz PRF translates to an unambiguous Doppler of ten Hertz and an unambiguous range of 15,000 kilometers. Any change in the PRF scales the Doppler frequency and range in an inverse fashion from these values

Radar aurora may have Doppler spreads of a few to many tens of Hertz, centered at anywhere from zero to several hundred Hertz. Very little work has been done at HF to determine the Doppler characteristics of the radar aurora. An example of one auroral return is shown in Fig 7. For the relatively low PRF which was used, the auroral return fills all Doppler space and is confined in range to an amount which is not determinable from the figure If one is interested in targets at a particular range which is not the range of the auroral return, and the aurora is range ambiguous, one can usually choose a PRF which is just slightly different rather than an integral multiple of the original, so the ambiguous range of the aurora lies at some other range than that of interest This of course depends upon the siting of the radar. For cases when the radar is at a northern latitude and the return is not range ambiguous, so that PRF changing is not an effective technique, one may employ adaptive antenna processing techniques Here one attempts to place a null in the direction of the strongest radar auroral return by appropriate phasing of individual antenna outputs As it is not always possible to handle multiple returns this way, or perhaps difficult to suppress the auroral return with enough dynamic range to allow target detection, the radar aurora presents one of the most difficult ionospheric disturbances to deal with for most HF radars. In the case of a radar with a non-simple pulse, using either a type of compressed pulse or an FM-CW format, the problem is compounded in that the auroral return will spread over a larger range and Doppler extent than for a simple pulse radar, with an appropriately smaller amplitude, such that the integrated backscattered energy over Doppler and range is the same as for the pulsed case. Properties of the radar aurora which are important are their Doppler and range characteristics, whether or not these characteristics can be classified into different groups, the frequency dependence of their cross section, their behavior with time (diurnal seasonal, and solar cycle), and knowledge of their structural sizes for adaptive beam forming applications Work in these areas would certainly be useful for HF radar applications

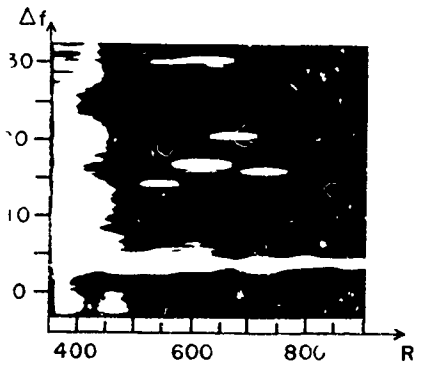


Figure 7. A Doppler-range display is shown in which broad band auroral returns are seen at 400 n mi range, extending through the entire unambiguous Doppler bandwidth, as well as appearing in all range bins but restricted in Doppler spread around zero. A clutter filter was used here so the region right at zero Doppler is free of targets and aurora. The PRF has been chosen so as to place the range ambiguous aurora at a folded range which does not affect the range cells which contain the targets of interest.

Meteor Returns

Meteors entering the atmosphere produce ionized trails which can present a very high backscatter cross section to an HF radar The doppler bandwidths are typically sufficiently high so that the entire unambiguous Doppler bandwidth is filled for typical PRFs employed, due to the high deceleration rate of the ionizing particle Figure 8 shows an example of a Doppler-time history during strong meteor activity. Because the trail

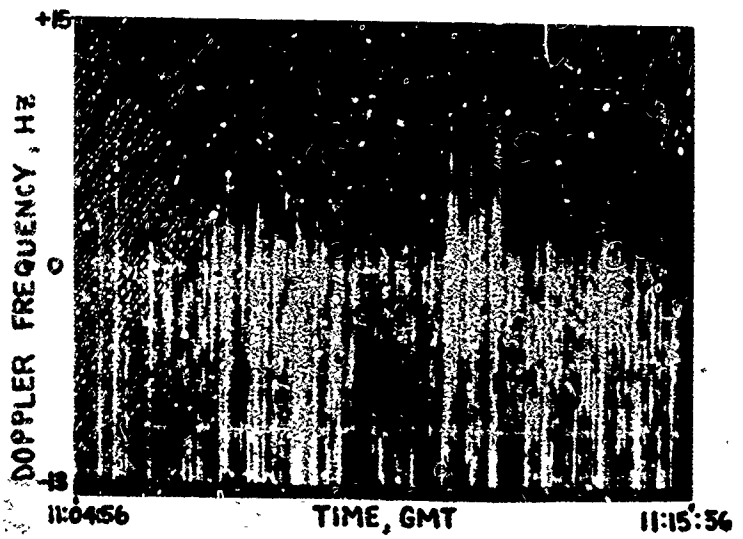


Figure 8 A Doppler-time display is shown for a range bin in which meteor trail returns are present The negative Doppler region has more scattered energy which is typical for an eastward look. For the short integration times employed, one can discern tracks at zero and minus ten Hertz, low level processing artifacts in this case, in spite of the meteor echoes. Such a display is almost imperative for identifying targets in the presence of meteor echoes.

lasts for a very short time, some improvement in target detection on a Doppler-time history can be achieved by using very short integration times, so that there are longer periods of noise free spectra during which the target can be detected above the normal noise floor. Such is the case for the data of Fig. 8, for which two simulated targets are observed to be visible in the presence of a multiplicity of meteor returns which fill all Doppler bins during those periods they are present In addition to the relatively high amplitude returns which occur as the meteor trail is initially created, the residual ionization can also provide a very low level background clutter return which has a lifetime extending over several integration periods. The Doppler spread now is very narrow after deceleration, and can be offset up to a few Hertz in Doppler due to the winds local to the ionization Such an event is generally due to a relatively large meteor and is relatively rare.

Traveling Ionospheric Disturbances

Traveling ionospheric disturbances (TID's) due to gravity waves propagating through the illuminated portion of the ionosphere will cause an undulation in the Doppler-time history of clutter or a target return, as was shown in Fig 3 These can be produced by solar disturbances, earthquakes, very heavy storm activity, and nuclear explosions in the atmosphere. Detection of the TID associated with testing in the Pacific was accomplished by NRL personnel (5) using the MADRE radar. Under the most severe of conditions, the undulations can be virtually continuous, although this condition is very rare

Sunrise and Sunset Effects

During periods of sunrise and sunset at ionospheric altitudes, the ionization undergoes so severe a transition that radar operation becomes difficult. At sunset the MUF drops precipitously, so that if a relatively high radar frequency was being used relative to the MUF, the backscatter disappears completely For lower radar frequencies, the radar backscatter undergoes a severe shift in Doppler, effective slant ranges change, and the spectral quality of the clutter becomes very poor. Such a scenario may last up to the order of half an hour As stability is re-established, the backscatter return may oscillate in Doppler as due to a TID, until the new stable conditions are set. At sunrise the reverse pattern takes place with time. A detailed study of this behavior has not been conducted, but would prove very useful for prediction of such outage times to radar operation

Spread-F

Ionospheric conditions referred to as spread-F, due to the characteristic spread ionogram echo from the F region, and associated with so-called "flutter fading" of communications at HF, also affects HF radar scatter in a similar way Although no examples are presented here, the effects of spread-F are to present a very broad clutter return with several multiple peaks, not unlike multi-path effects, or in the extreme, auroral like echoes from the F-region. For this reason, unless a real time vertical ionogram is available, spread-F is not easily identifiable on the radar display Very little work has been done to recognize radar spread-F since it is primarily a night time phenomenon. Typically, little experimental work has been conducted at night, although perhaps radars such as the Air Force Experimental Radar System (ERS) will encounter the effect more because of its projected around the clock operations.

Equatorial Long Range Returns

A very interesting and unusual radar observation was that of scatter from equatorial irregularities associated with sunset using the NRL MADRE radar several years ago (6). The scatter was observed as a very broadly Doppler spread return across a a relatively broad range spread The radar PRF was reduced to determine the range ambiguity and it was found that the scatter was coming from a region over North Africa. The data are shown in Fig 9. The top portion is a plot of non-processed returned power versus range, a typical A-scope display as used with non-coherent radars. The normal clutter is observed at ranges extending from about 500 to 2,000 nautical miles on the display, somewhat spiky in nature as is usually observed from sea clutter. The return from the equatorial region is seen as a much smoother return with range extending from 5,000 to 6,000 nautical miles. The clutter from the subsequent pulse is seen beginning at the far right of the display. A montage of Doppler-range displays are shown for several consecutive times in the lower display, accounting for the slight

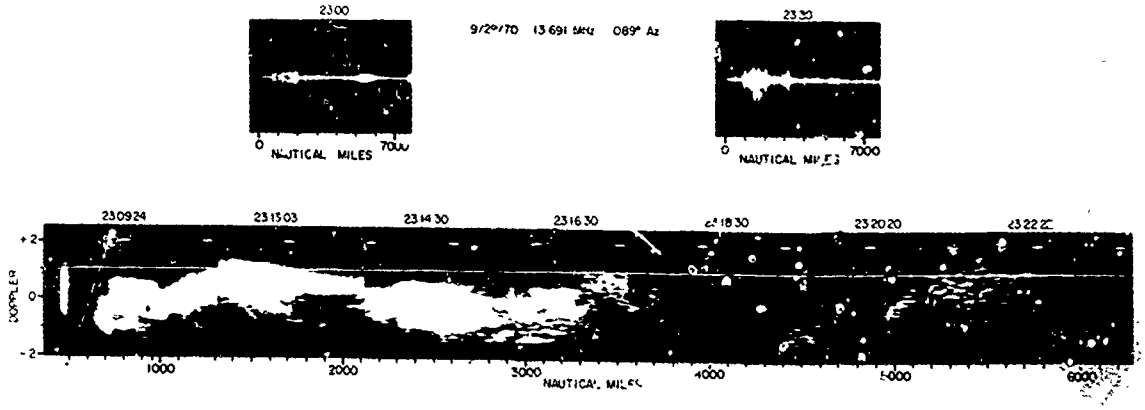


Figure 9. An example of radar scatter from equatorial structure is seen on this extended Doppler-range display. The equatorial scatter is seen as a relatively broad Doppler return compared to the clutter, in two patches centered at 4500 n mi, and again from 5,000 to 5,800 n mi. The unprocessed time-averaged amplitude-range presentations in the upper left and right, collected at two different times, indicate the total power scattered relative to the clutter.

discontinuities in range. The bending of the clutter is quite typical for this time of day (Universal times employed), and is due to a slowly rising ionosphere at some altitudes. The scatter from the equatorial region is observed to have a much broader spread in Doppler frequencies, and made up of a large number of discrete returns as opposed to the more continuous sea clutter. The succession of returns at 2 Hertz is a series of calibration signals, and represents what a large ship at high radial speed might look like on the display It is noted that the calibration markers near the equatorial scatter are much more difficult to detect among the spread Doppler returns Professor Jean Delloue, of the University of Paris, has done extensive work on forward propagation effects of the equatorial scatter phenomenon (private communication), and agrees with the interpretation of the radar phenomenon being due to earth backscatter via field-alligned irregularities. He has reported that the active volume lies in the region of the setting sun at ionospheric heights, and the scatter region can actually be followed along in azimuth. Using a transmit site in Europe and a receive site in Africa, both steerable, he was able to follow the movement of the irregularities and maintain a propagation path using frequencies above the MUF, by steering the arrays accordingly The rate of steering indicated that the scattering volume was following the ionospheric sunset and was repeatable on a regular basis. In many cases, the quality of the path was sufficient for intelligible voice communications This would agree with the results of Figure 9, which indicates a bandwidth suitable for audio intelligibility THe effect can be present throughout the night once it is established

SUMMARY

We have attempted to present some of the effects the disturbed and the undisturbed ionosphere can have on high frequency over-the-horizon radar operation A series of examples of the effects on radar displays were given, as they give an idea of the severity of each of the problems. Because of the large number of such ionospheric effects, we have chosen to only demonstrate them rather than to attempt to reference literature on related research on the phenomena It is felt that the researchers in the various fields encountered would profit by showing the potential for using HF radar as an experimental tool in further studying the mechanisms responsible for the effects, and we invite correspondence regarding co-operative experiments in such endeavors

REFERENCES

1 Headrick, J M and Skolnik, M I., "Over-The-Horizon Radar in the HF Band", Proc.IEEE, Vol. 62, pp. 664-673, 1974.

2 Jones, R.M., and Stephenson, J J., "A Versatile Three-Dimensional Ray Tracing Computer Program for Radio Waves in the Ionosphere", Office of Telecommunications Report 75-76, (PB248856), 1975.

3 Thomason, J.F., Skaggs, G.A , and Lloyd, J L., "A Global Ionospheric Model", NRL Report 8321, 1979.

4 Trizna, D.B , Preliminary results of HF Radar Tracking of Ocean Eddies", (in preparation).

5 Davis, J R. and Headrick, J M., "A Comparison of High Altitude Nuclear Explosion Effects in the E-Layer with Variations in Magnetic Field Strength", J.G.R., 69, pp 911-916, 1964.

6 Headrick, W C and Headrick, J M , (unpublished results).

MEASUREMENT OF SEA AND ICE BACKSCATTER REFLECTIVITY USING AN OTH RADAR SYSTEM

William F. Ring and Gary S. Sales

Rome Air Development Center
Deputy for Electronic Technology
Propagation Branch
Electromagnetic Sciences Division
Hanscom AFB, Massachusetts 01731

Introduction

The characteristics of radio waves that are back scattered from the earth's surface have interested observers for many years because of their potential to yield significant information about the scattering surface. An example would be the recognition of coastal regions or islands by the different scattering characteristics of land and water. The technique has been pursued to develop a method of remote sensing of sea states (Ref 1) including wave heights and wind direction by studying the dominant peaks in the observed Doppler spectrum. The interest in this study is to quantify the surface reflectivity of sea water, sea ice and the Greenland ice cap so they may serve as standards or references for calibrating the sensitivity of a High Frequency (HF), Over-the-Horizon radar when there are no aircraft targets of opportunity to serve as calibration sources. With these calibrated surfaces the system performance could be assessed at any time or place.

Data and Analysis

The source of data for this study was the Polar Fox II experimental backscatter radar which was operated in northern Maine by the Raytheon Company under Air Force sponsorship for a twelve month period in 1971-1972. The portion of the coverage area that is pertinent to this study is shown in Figure 1. The experiment used a high power pulsed radar (800 kW peak) operated at 30 pulses per second with a 10 kHz chirp bandwidth. Radar parameters are listed in Table 1.

TABLE I

Radar Parameters	
Location:	Caribou, Maine. 47°N 68°W
Frequency:	Selectable in the band 6 to 26 MHz
Power:	64 kW avg. 800 kW peak
Range Resolution:	15 km
Receive Beamwidth:	7 degrees at 12 MHz
Azimuth:	11 steps from -30°T to +60°T
Measurement:	One minute/Freq/Beam Position/Hour

The 49° azimuth was selected as a source of sea water returns at all ranges at all times of the year. The 15° and 23° azimuths provided sea ice returns at the shorter ranges (1600, 1800 and 2100 km) during selected intervals of the winter and spring months. The determination of whether the scattering area consisted of sea ice or water was based on ice analysis charts for the period of operation prepared by the USN FLEWEAFAC, SUITLAND MD. The 15° and 23° azimuths were also the source for the Greenland ice cap scattering surface at the longer ranges.

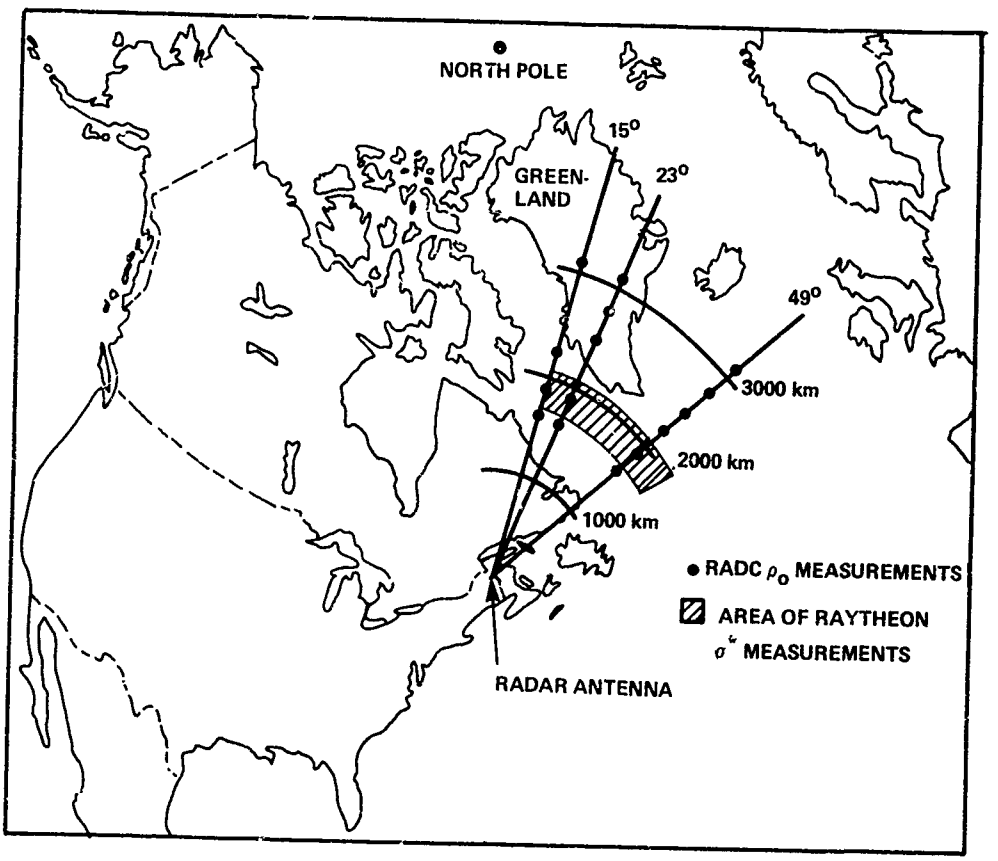


FIG. 1 COVERAGE OF THE POLAR FOX II RADAR

By sampling and comparing data at the same ranges on all azimuths included in the study, an attempt was made to eliminate any dependence on elevation angle. The original Polar Fox II experiment required that this radar be located so as to include the auroral oval in its coverage area but for purposes of this reflectivity study it is preferable to select data such that they are not affected by auroral propagation anomalies. This consideration dictated that the data be limited to the mid-day, 17-19 UT, period (13-15 local time in the coverage area).

The data selected for this study were limited to F-region propagation modes in order to insure confidence in our description of the propagation loss and antenna gain factors involved in the analysis of the received signals. For the same reason it was decided to avoid the sometimes anomalous propagation conditions associated with the summer day time by limiting the data base to March, April, September, October, and November. Radar data were available for a total of thirty days over the five month period.

The analysis begins with the measurement of the received backscatter power and then with the receiver calibration, antenna gain patterns, transmitter power and the assumed ionospheric parameters for estimating propagation angles and absorption losses, the surface reflectivity is computed using the radar equation.

$$\rho_o = \frac{(4\pi)^3 R^4 (L^2)^2}{P_T G_T G_R \lambda^2 \tau (R\Phi)} P_R \quad (m^2/m^2)$$

R = range to the reflecting surface

P_R = received backscatter power

L = ionospheric loss (one leg)

P_T = transmitted power

G_T = transmit antenna gain

G_R = receive antenna gain

λ = wavelength of observing frequency

τ = pulse width

ϕ = azimuthal beamwidth

The ionospheric loss term is derived using the method of George and Bradley (Ref. 2). Antenna gain factors were obtained from measurements performed in a cooperative program by Raytheon, Lincoln Laboratory (Ref. 3) and the Rome Air Development Center (Ref. 4).

To insure tht the selected backscatter data was propagated by a one hop F(1F) mode it was necessary to review manually the range/Doppler characteristics of the received signals. The received signal for a given azimuth measurement is displayed in the format shown in Figure 2. Here the Doppler spectrum over ± 15 Hz and 45 km range steps is presented by printing the magnitude of the received power (-dBW) in the appropriate range/Doppler location. This provides a pictorial overview of the received power in a two dimensional format. To be accepted as 1F-propagated surface clutter the energy had to be concentrated in the narrow band centered on 0 Hz and be part of a continuum of signals over a range interval appropriate for F mode coverage. In the example shown, clutter from ionospheric irregularity produced measurable returned power at all Doppler frequencies in the range interval from 900 to 1300 km (outside the area of interest for this study). Ground clutter returns start around 2100 km and continue to 3400 km.

Additional criteria were imposed on the data selection which can best be described by visualizing the received power at 0 Hz Doppler, as a function of range as shown in Figure 3. To be included in the data base used for this study, a received power sample had to occur at a range where the ground backscatter signal was fully developed, i.e. not on the leading edge and more than 10 dB above the noise level.

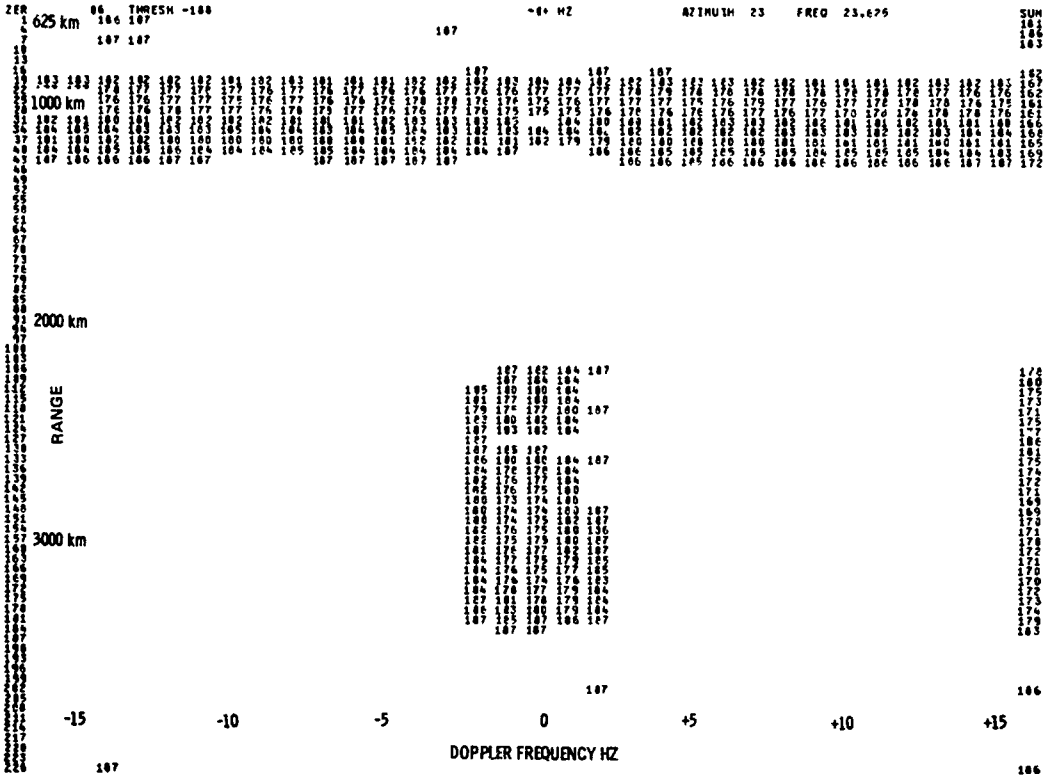


FIG. 2 ONE MINUTE SAMPLE OF RADAR DATA

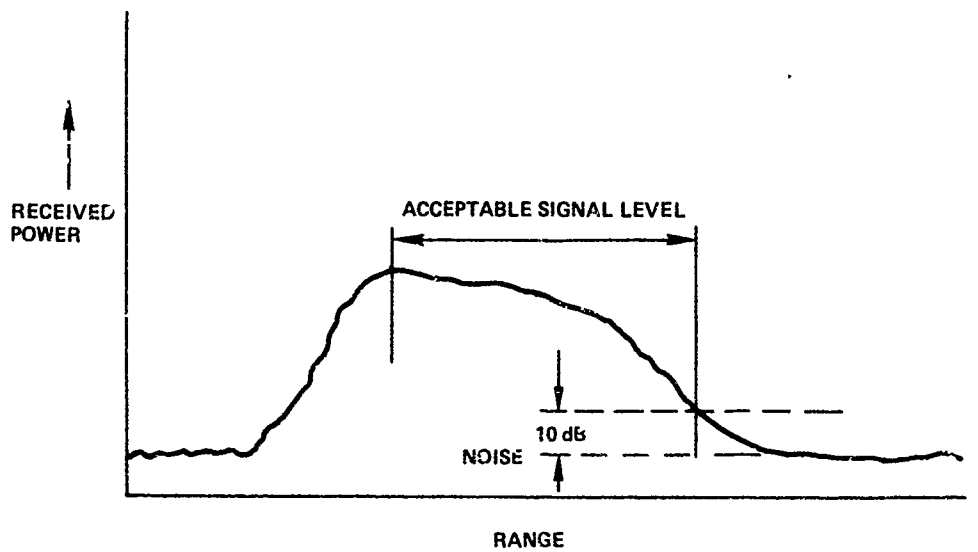


FIG. 3 TYPICAL GROUND CLUTTER SIGNAL

Backscatter reflectivities. The values shown in Table 2 were determined using the relationship for surface refelctivity described above. These values represent the upper decile of the data available from ninety hours of observations. The spread of the individual measurements below the upper decile was consistently of the order of 20 dB independent of the scattering medium, including the Greenland ice cap. The upper decile value is cited so that observed variability below that level can be explained as the result of various loss factors.

TABLE 2. Upper Decile Values of Surface Reflectivity $dB(m^2/m^2)$

Frequency	12MHz	15MHz	19MHz	24MHz
Surface				
Sea Water	-22	-24.5	-20	-23
Sea Ice	-34	-39		
Ice Cap			-34	-37

Temporal variations. In an attempt to measure temporal structure of the variations in reflectivity, the signals measured one minute apart from adjacent azimuth beams (9^0 corresponds to 300 km transverse distance at the range of interest) were compared. Under these conditions the scatter of the data points was substantially the same as previously measured. Unfortunately, it is not possible to sample more frequently than once per hour at exactly the same azimuth. To the extent that this 9^0 azimuthal change can be considered small and since the motionless ice cap shows the same variability, these one minute fluctuations appear to be characteristic of the ionospheric propagation path and loss factors. Some of the major contributors to the spread in these measurements would include polarization, absorption and F-region irregularities. These factors, particularly absorption (auroral) and F-region irregularities, are certainly enhanced by the proximity to the auroral zone, even in daytime.

Azimuthal variations. If all equipment and propagation factors had been properly taken into account, it is reasonable to expect a constant reflectivity for a given surface type (e.g. sea water) independent of observing azimuth. This did not prove to be true when reflectivities at azimuths of 15^0, 23^0 and 49^0 were compared using data from September when sea water was present in the 1600-2100 km range segment. In order to understand this unexpected result the Raytheon analysis of the Polar Fox II data which produced a reflectivity value σ^*, for every 75 km range interval was examined. One significant difference is that the σ^* values do not contain an ionospheric loss term and thus, what is measured is ρ_o with loss or $\rho_o/(L_2)^2$. These σ^* values, taken over the range interval from 1600 to 2050 km and at azimuths of 15, 23, 31, 40, 49 and 60 degrees provided a sizeable data base for this investigation of the azimuthal dependence of reflectivity. For a given range the ratios of reflectivity for different azimuths were compared and the results showed the same azimuthal variation as the ρ_o values. There is good agreement between the Raytheon analysis and the work presented here.

Because the azimuthal variation of reflectivities is strikingly similar to the azimuthal antenna patterns a careful review was made to insure that the antenna measurement results were reasonable and had been correctly incorporated in the reflectivity calculations. The validity of the antenna considerations having been confirmed there remains an unexplained but apparently real azimuthal latitudinal variation of sea water backscatter reflectivity as shown in Figure 4. Assuming constancy of the reflective properties of the sea and accepting the validity of the antenna pattern measurements, it is necessary to compensate for this observed azimuthal variation. Although the source of this assumed propagation loss mechanism, which increases in a northerly direction is presently undetermined, the 12 and 15 MHz reflectivity values have been corrected for it and these corrected values are shown in Table 3.

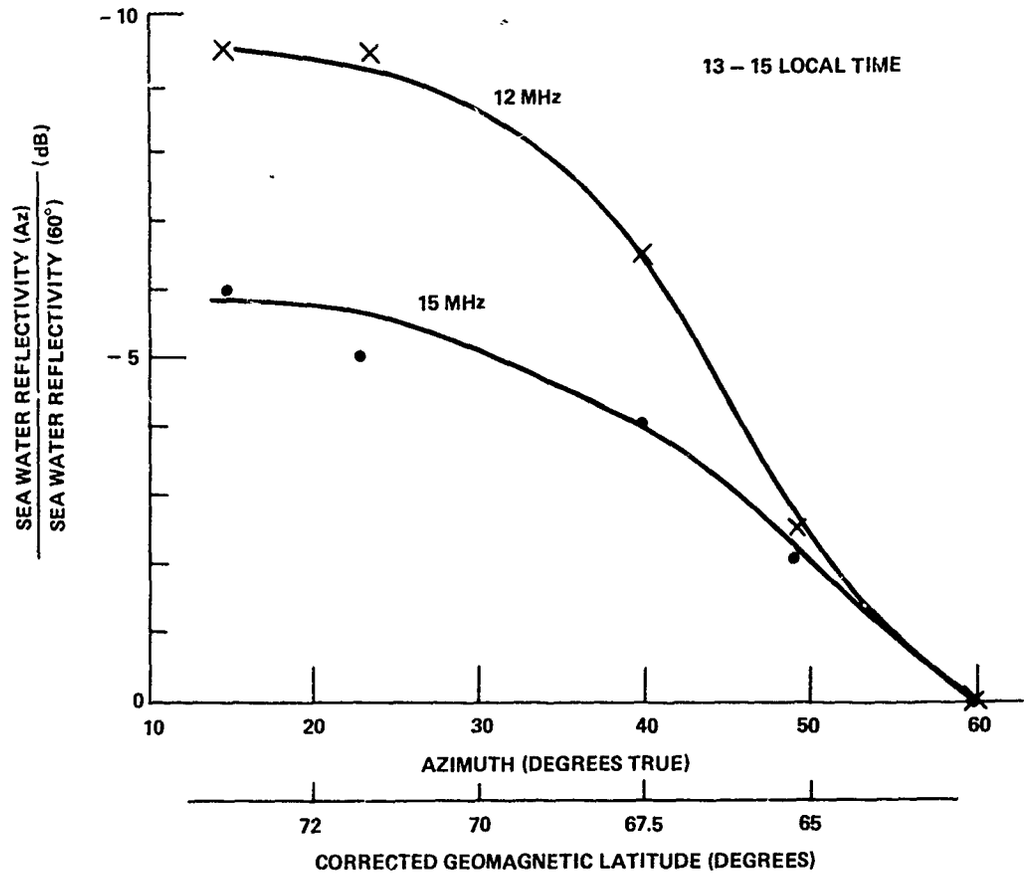


FIG. 4 AZIMUTHAL DEPENDENCE OF SEA WATER REFLECTIVITY

TABLE 3. Corrected Upper Decile Values of Surface Reflectivity $dB(m^2/m^2)$

Frequency	12 MHz	15 MHz	19 MHz	24 MHz
Surface				
Sea Water	-17	-20	-20	-23
Sea Ice	-25	-34		
Ice Cap			-31	-35

It can be seen from Figure 5 that the 12 and 15 MHz absorption values are in good agreement with a frequency dependence of f^{-2} and this relationship was used to correct the 19 and and 24 MHz reflectivity values in Table 3. Since the sea water reflectivity values at 19 and 24 MHz come only from the 49 degree azimuth these values are unchanged.

Discussion

The upper decile value of reflectivity of sea water approaches -17 dB m^2/m^2 with little variation with frequency over the range of 12-24 MHz in these observations. The upper decile value is cited so that observed variability below that level can be explained as the effect of the various loss factors. Theory (Ref. 5) predicts that the first order scattering coefficient should be approximately -17 dB, independent of sea state or frequency, in this region of the HF band. The independence of reflectivity with sea state is attributed to the fact that the waves effective in the scattering process are half the radar wavelength, or 6.5 to 12.5 meters for this study, and waves of these lengths are almost always present and fully developed to their maximum height on the open ocean in the North Atlantic.

Because of the location of the sea ice and ice cap regions, only narrow bands could be used for illumination so inferences as to frequency dependence are impossible. It is clear that the radar scattering cross section of both these media are significantly below that of the open sea.

Perhaps the most interesting result of this study is the azimuthal or latitudinal variation. If it is assumed that this variation is due to auroral absorption an inconsistency arises when the data is compared to the measurements of Hartz et al. (Ref. 6) or predictive models by Foppiano (Ref. 7) and Basler (Ref. 8). These measurements and models do not show a consistent increase in absorption from 65 to 72 degrees geomagnetic latitude as reflected in the present data. The thoroughness of the antenna measurement effort and the reasonableness of the antenna pattern results require, however, the postulation of some loss (ionospheric absorption or scattering) mechanism.

Conclusion

The radar scattering cross section of open sea water, sea ice or the Greenland ice cap are sufficiently predictable to be useful in calibrating the system sensitivity of an OTH radar. Questions that remain to be answered concern the variance in these backscattered signals as a function of averaging time and the geographical area illuminated.

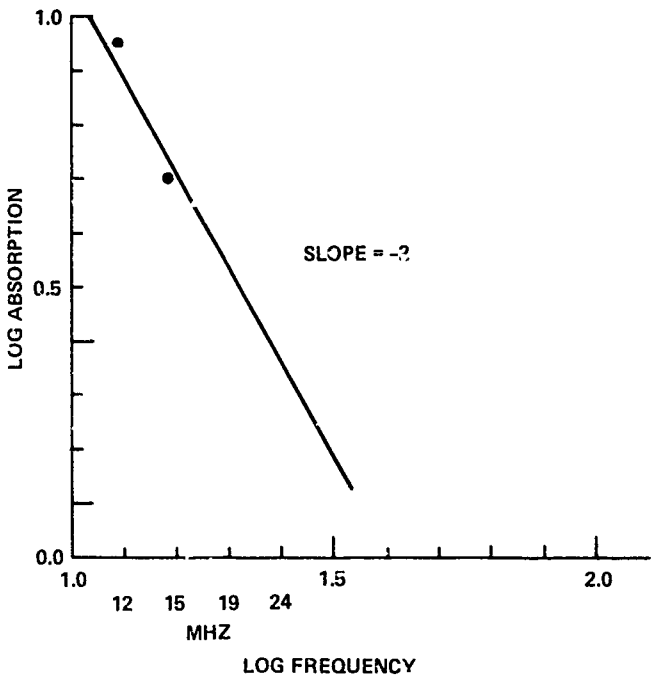


FIG. 5 FREQUENCY DEPENDENCE OF AZIMUTHAL (LATITUDINAL) ABSORPTION

References

1. Sea Backscatter at HF: Interpretation and utilization of the echo. Barrick, D. E., Headrick, J. M., Bogle R.W. and Crombie, D. D., Proceedings of the IEEE, Volume 62, No. 2, June 1974.

2. A new method of predicting the ionospheric absorption of high frequency waves at oblique incidence. George, P. L. and Bradley, P.A., ITU Telecommunication Journal, Geneva, May 1974.

3. Propagation Measurements Program (Polar Fox II) Antenna Patterns and Gains. Nichols, B.E., Lincoln Laboratory Project Report PMP-6, July 1973, AD 896-788L.

4. Antenna Pattern Measurements Project 1765. Parry, J. L., RADC-TR-73-32, March 1973, AD 909-089L.

5. First Order Theory and Analysis of MF/HF/VHF Scatter from the Sea. Barrick, D.E., IEEE Trans. Antennas and Propagation, AP-20, 2-10. January 1972.

6. A study of auroral absorption at 30 MHz. Hartz, T.R., Montbriand, L.E., and Vogan, E.L., Canadian Journal of Physics 41:581 (1963)

7. A New Method for Predicting the Auroral Absorption of HF Sky Waves. Foppiano, A., CCIR, IWP 6/1, Documents 3 and 10 (1975).

8. Chatanika Model of the High-Latitude Ionosphere for Application to HF Propagation Prediction, Vondrak, R. R., Smith, G., Hatfield, V. E., Tsunoda, R.T., Frank, V.R., and Perreault, P.D., RADC-TR-78-7 (1978), AD AO 53154.

ADAPTIVE UTILIZATION OF HF PATHS: A WAY TO COPE WITH THE IONOSPHERIC LIMITATIONS AFFECTING MID-LATITUDE AND TRANSAURORAL SHORT-WAVE LINKS

Ashok K. Gupta
The Catholic University of America
Washington, D.C. 20064

and

Mario D. Grossi
Harvard-Smithsonian Center for Astrophysics
Cambridge, Massachusetts 02138

ABSTRACT

Recent advances in HF path characterization as well as in microprocessor technology, coding schemes and communication equipment, make it possible to conceive approaches of adaptive utilization of HF propagation that promise a substantial improvement upon present-day HF circuit performance.

The paper discusses a new adaptive approach based on path sounding, channel probing, MFSK signalling and waveform diversity. The sounding and probing functions provide the initial information for the working of the adaptive control loops in a two-way, high data rate, adaptive exchange between the two terminals of the link, with feedback acknowledgement.

It is shown that the proposed method makes it possible to transmit a data rate of 24 Kbit/sec with an error rate of 10^{-4} (therefore suitable for high-quality digital voice) in a single hop HF ionospheric path established in mid-latitude or across the aurora belt.

1. INTRODUCTION

HF propagation paths are time-spread and frequency spread channels, and are characterized by severe variability in the time domain of all their properties, inclusive of path losses. Improvements over a present-day HF link's performance in terms of circuit reliability, data rate, and error rate can be achieved only through the use of adaptive schemes that in principle appear to be able of coping with the channel variability, if the necessary penalty in terms of equipment complexity is willingly accepted.

This paper illustrates an adaptive method that makes it possible to transmit a data rate of 24 Kbit/sec with an error rate of 10^{-4} (therefore suitable for high quality digital voice), in a single hop ionospheric path established in mid-latitude or across the aurora belt. The proposed adaptive approach is based on path sounding, channel probing, MFSK signalling and waveform diversity. The real-time oblique ionospheric sounding and probing functions provide the initial information for the working of the adaptive control loops in a two-way, high data rate, adaptive exchange between the two terminals of the link, with feedback acknowledgement.

For the sake of illustration, we have assumed in this paper that the link has 1125 to 3375 carriers, respectively for a mid-latitude and a transauroral path, available in the HF band between 3MHz and 30MHz. A sounding scan lasts 100 to 160 seconds and is repeated every 300 to 480 seconds. The master station of the link, where the sounding xmitter is located, also generates the waveform for channel probing and includes a complete terminal for two way digital communications. During the pauses of emissions, measurements of noise and interference levels are performed at both the master and the slave station of the link, for use by decision-making microprocessors and control units. At each terminal, the transmitting and receiving facility could have separate units for sounding/probing and for communicating or these functions could be performed by the same equipment in different modes of operation. In the latter case, the equipment at the two terminals could be identical and the assignment of the master and slave roles could be dictated by operational requirements.

By processing the data obtained by sounding and probing, it will be possible to select automatically the group of frequencies to be used for communicating. At each sounding cycle, information about the frequency selection and about the waveform to be employed is exchanged between terminals and used locally to achieve adaptivity. During the next sounding scan (performed at a rate of one every 5 to 8 minutes) the group of frequencies selected for communication are excluded from the sounding frequency plan. Instead, information on the changing status of the group of communicating frequencies is obtained from measurements performed on the coded waveform that is part of the

communication bit stream.

In Section 2, we illustrate the characterization of HF ionospheric paths, while in Section 3, we discuss path sounding and channel probing. Section 4 illustrates a communication approach applied to mid-latitude and transauroral HF links, in order to counteract effectively ionospheric limitations. Section 5 and 6 briefly discuss, respectively, on an application of M-ary transmission to one-way communications and on an application of adaptivity to a meteor burst channel.

2. HF IONOSPHERIC PATHS CHARACTERIZATION

The HF ionospheric channel, by virtue of energy propagating over many paths of different propagation delays, exhibits time dispersion (or frequency selective fading), and by virtue of time variations of individuals phase-path lengths, exhibits frequency dispersion (or time-selective fading). To represent this time and frequency selective fading, various mathematical models can be formulated to identify its measurable parameters. These modes (or characterization) are also useful in selecting measurement techniques for the channel parameters. The selection of channel parameters itself depends upon the particular application and upon the receiver structure. The general principles involved in randomly modelling time-variant channels are well known [1,2,3,4,5]. These composite filter models. The two types of models may be distinguished initially. A "multipath transmission view" is suggested by identifiable mechanisms or paths of propagation between input and output terminals. The composite filter view is a view whereby channel output is related to input by means of mathematical operation involving a suitably defined system function.

In multipath view, the channel is described in terms of specular path, diffuse path or quasi-specular path. The output of a given medium may actually consist of a linear combination of non-diffuse and diffuse components, and different combinations may be necessary to describe the output of the same median at different times. For example the sum of a few quasi-specular paths, each of which fluctuates slowly relative to the fluctuations of the resultant signal, is modeled by a sample function of a Gaussian process plus a specular component. Such a combination may be encountered over HF links.

The representation of a channel output in terms of a linear combination of non-diffuse and diffuse signal models with stationary statistics transforms the characterization of the channel into the specification of:

1. a multipath structure,
2. the relative intensities, and average delay and doppler differences of the various distinguishable paths.
3. the statistical characteristics of individual path delays and delay spread, and path doppler shifts. Doppler spread and spectral skewness.
4. the constitution and properties of the fine structure of each distinguishable path [i.e., its possible decomposition into the sum of non-diffuse (specular or quasi-specular) and diffuse components] including the relative characteristics of the components, and the statistical properties of their parameters.

When modeled by a "composite filter," the variable multipath nature of the propagation process over HF channels cause a number of effects of great importance in the evaluation of the signal transmission performance of the "equivalent" filter.

First, the group delay differences among the various paths cause the overall channel-composite filter attenuation and delay characteristics to vary with frequency. If the delay-spread of the significant paths is not a very small fraction of the reciprocal of the bandwidth occupied by the signal, the various components of the signal will experience non-uniform attenuation and delay, which results in signal waveform distortion. The characteristics of this distortion will be random for a randomly time-variant linear channel.

Second, the fluctuations in the relative characteristics of the various paths cause each frequency component in the signal to acquire a modulation of envelope and phase (and, hence, frequency). This multipath-induced modulation causes the signal level to fluctuate up and down, which raises the possibility of signal outages or dropouts, each resulting from a "fade" or drop of the received-signal strength below the threshold of acceptable performance in the presence of independent additive disturbances. In addition, a non-zero delay spread among the various paths limits the bandwidth over which the fluctuations experience by signal components at different frequencies will maintain the necessary degree of mutual coherence to keep the resultant signal distortion below tolerable bounds.

Finally, the fluctuation rate induced by the channel in a particular parameter of the carrier sets a non-zero limit on the frequency content in the baseband spectrum that can be used to modulate that particular carrier parameter if interference between the channel fluctuations and the desired signal baseband waveform is to remain negligible.

If the "composite" filter model is used, the time, frequency and/or statistical behavior of a channel-characteristic system function offers the basis for the definition of the gross parameters for describing the gross aspects of the channel response. The system functions for the Randomly Time-Variant Linear (RTVL) Channels have been studied extensively in the past in the time domain [e.g. sampling model (3)] or in frequency domain [e.g. paired-echo model (1)]. In time-domain approximation of the system function, there is no explicit quantitative indication of how well the result approximates the frequency-domain characteristics. In practical applications, one is usually concerned with how well the actual filter characteristics can be approximated by a given number of paths, or "taps." The paired-echo model was proposed to minimize the number of taps required to approximate the system function in frequency domain.

In all signal models, the fading characteristics of the channel is incorporated by varying the tap or path coefficient, corresponding to the fading rate or the fading bandwidth of the channel. Recently a complex filtering technique (5) has been developed to simulate Asymmetrical Doppler-shifted Narrowband Gaussian (ADS-NBG) channels. With this technique, it became possible to account for some additional channel parameters not considered before, such as the non-zero centroid and the skewness of fading spectrum (or of the spectrum of the extended channel comprising the fading channel and receiver processing filters). Such spectral shapes are known to occur in troposcatter channels (5) and in the HF transauroral channels, of direct interest here (6). The reasons for this non-zero Doppler shift and for the asymmetry in the spectrum of the baseband process are:

1. Doppler shift (see Figure 1). This may be due to "offset" in the tuning of the link's oscillators, to time changes in the refractive/scattering properties of the medium, etc.

2. Asymmetry in the spectrum (see Figure 2). This may be due to the geometrical asymmetrics in the link (off-great-circles-path superimposition to the main path), asymmetrics in the receiver filters (RF, IF or baseband), etc.

Theoretical interpretations developed thus far account only for the Doppler shift due to the frequency offset in the local oscillator at the receiver. No analytical description had been worked out before in order to account for the shift due to propagation medium and for the asymmetry in the received baseband spectrum. Figure 3 shows a typical representation available in the literature for the received spectrum which is clearly symmetrical.

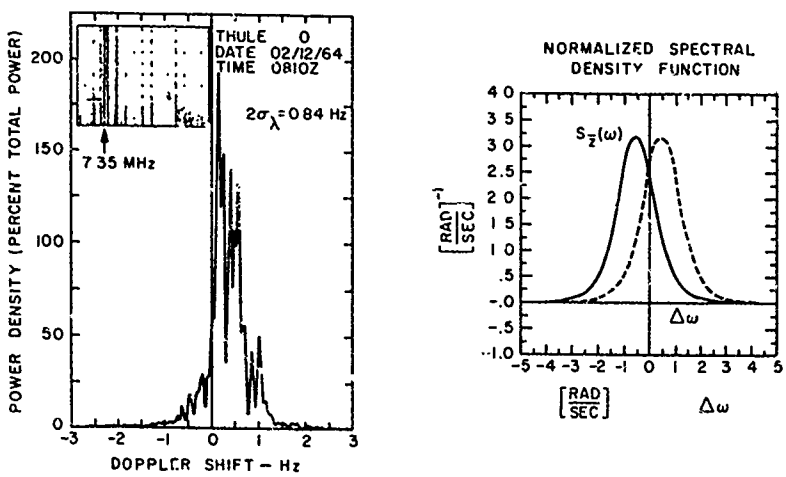


Figure 1. Doppler-Shifted Spectrum in a Transauroral Path and an Example of Computer Simulation

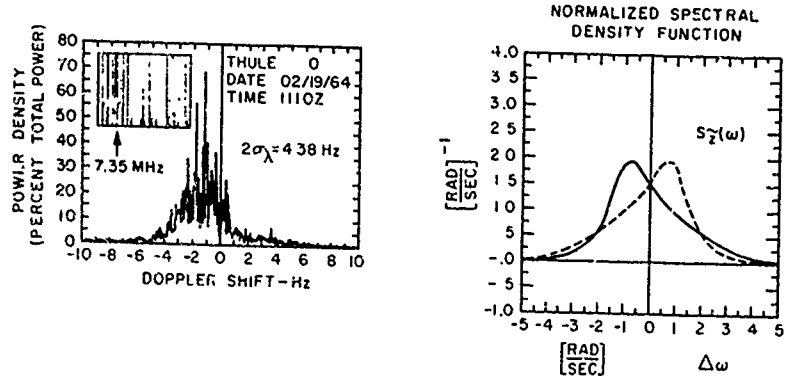


Figure 2.

Example of Skewed Asymmetrical Spectrum in Transaurcral Path and a Computer Simulation.

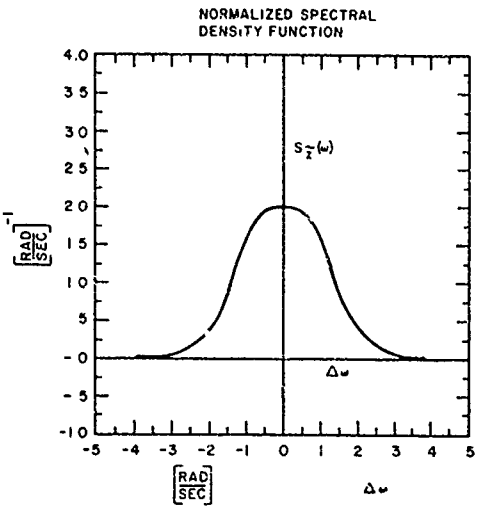


Figure 3.

Typical Symmetrical Spectra Considered Thus Far in the Literature.

This spectrum is then shifted by the "spectral-shift unit" to simulate the frequency offset in the receiver oscillators. In (5) an analytical representation of the channel has been developed that fully accounts for all aspects of the two phenomenon mentioned above, thereby removing the limitations in the present channel simulators. The effects of non-zero centroid and skewness of the fading spectrum on the binary-error probabilities of incoherent and differentially coherent phase-reversal matched filter receivers employing post-detection diversity combining has been investigated in (7). The results of this investigation to the HF links is a topic of continued research.

In practical applications, the input to the channel is time-limited and the output spectrum is confined by a bandpass filter. Therefore a discrete channel model is also of interest and has been developed in (8) where the time-frequency plane model is derived from a scattering function. The signal design approach discussed in this paper (see Section 4) is based upon this discrete channel model. With this model, it also became possible to demonstrate the various shapes of the possible delay-Doppler occupancy patterns. If the input to a channel is confined by a time gate to the time interval $0 < t < T$ and the output spectrum is confined by a bandpass filter to the frequency interval $-\frac{W}{2} < w < \frac{W}{2}$, then the number of coefficients of significant amplitudes can be determined by how many $\frac{1}{2}$ rectangles of dimension $\frac{1}{TW}$ can be fit into the delay-Doppler occupancy pattern. If S_A is the area of one such region, then the number of coefficients of significant amplitude or diversity $\mathcal{Z}$ can be expressed as

$$\mathcal{Z} \approx TW\ S_A \quad . \tag{1}$$

In the signal design approach discussed here, $TW = 1$ is assumed and therefore

$$\mathcal{Z} \approx S_A \quad . \tag{2}$$

3. PATH SOUNDING AND CHANNEL PROBING

3.1 Transmission Parameters and Their Measurements

In light of the discussion in Section 2, the gross transmission parameters, important in characterizing the effects of RTVL medium upon broad classes of signals are (1): the outage (drop-out or fade out); the dealy soread; the coherence bandwidth; the diversity bandwidth; the spectral moments of the Doppler spectrum (Doppler-shift, Doppler spread, spectral skewness etc.): the diversity time; the instantaneous channel function; the second-order channel functions. These parameters are not all independent. For example, the coherence bandwidth is inversely proportional to the delay spread.

The measure of communication channels is important in digital communications because high-speed digital data transmission requires considerable knowledge and equalization of channel characteristics. It is known (3) that the problem of the measurement of system functions of random time-invariant channels differs from the corresponding problem for time-variant channels in that, even in the absence of noise, the random system function may be unmeasurable. Kailath introduced a channel parameter called a "spread factor" as the measurability criterior. This parameter is the produce of B_{max}, the maximum rate of variation of the system in H_z, and L_{max}, the maximum multipath spread of the channel in seconds. According to Kailath, the system functions of a linear channel can not be measured if the "rectangular spread factor" of the channel S_R ($= B_{max}\, L_{max}$) > 1 and if no further information than B_{max}, L_{max} is known about the channel. However, he was careful to point out that additional channel knowledge would generally allow exact channel measurement even though $B_{max}\, L_{max} > 1$. Based upon the discrete representations of the channel corresponding to input-time and output bandwith constraints (Section 2), Bello (8) proposed the region of the non-zero delay-Dopper occupancy pattern being less than unity as the less stringent measurability criterion. This new channel parameter is called the area spread factor of the channel S_A. Another method of evaluating the "spread factor" of the channel is to define the spread factor as the ratio of the bandwidth of the fast fluctuations and the correlation bandwidth of the channel being measured. Fortunately except for certain esoteric radio channels (for example, the orbital radio channels), the spread factor is less than unity and thus the measurement techniques described below are generally quite useful.

These measurement techniques for a random dispersive channel are analyzed in (9) in three levels of increasing complexity:

a. the measurement of multipath spread and Doppler spread, and Doppler shift and spectral skewness.

b. the measurement of second-oder channel functions,

c. the measurement of instantaneous channel functions.

For the parameters in (a), measurement techniques used are based upon differentiation, level-crossing and correlation (10). For (b), the techniques used are correlation technique multitone technique, pulse-pair and chip techniques [see ref. (7)]. For the measurement of the instantaneous values of the channel functions, the cross-correlation, multitone and pulse pair techniques are used [see ref. (7)].

Since the signal design approach discussed in Section 4 only requires the measurement of Doppler spread and multipath spread, we limit ourselves to the probing of these two channel parameters.

3.2 General Remarks

It would be important to know accurately the time variability and the related statistics of such fundamental parameters as path losses, noise and interference spectral density, multipath spread and doppler spread. Unfortunately these are known only in particular cases so that a reliable experimental investigation on the properties above is thus far an unfulfilled requirement. In general, we can say that these ionospheric channels exhibit time fadings that are important in determining the design of the signal and, in addition, show long-term variations due to large-scale fluctuations of the medium. Such slow effects have a time constant significantly greater than 5 to 10 minutes, an interval of time selected (as will be seen in the following sections) as the basic sounding/probing periodicity. Adaptive approaches to the communication problem are required to circumvent this long-term variability in propagation conditions.

In this paper, we make the usual distinction between path sounding and channel probing, with the former devoted to the measurement of path losses and of noise and interference levels, and with the latter devoted to measurement of such parameters as multipath spread and Doppler spread. The following criteria were adopted:

1. The link is assumed reciprocal, except for the noise and interference levels at each terminal. Therefore, the decision on the frequencies to be used (this decision is the output of the sounding operation) is based on the measurement of interference and noise both at the master and at the slave station, and on the one-way measurement of the path losses between the two.

2. Processing of multipath spread and Doppler spread (the basic operation of the channel probing activity) is performed at the slave station and the results are transmitted back to the master station, for use in the final selection of the frequencies to be used in communicating.

3. Channel probing is to be undertaken only at the best frequencies put in evidence by the path sounding, in order to shorten the overall cycle sounding/probing.

As far as the estimate is concerned of the path loss and of the time/frequency spreads, we would like, ideally, to provide a reliable estimate of all the parameters indicated in Figure 4 for each path of interest in our study. They are the quantities B_d, B_0, B_m, $S(\xi,\nu)$. Actually it would be sufficient to simplify the scattering function to a group of N gaussoides, with $B_0 = 0$, $B_d = 0$, and to reduce therefore the scattering function to the analytical expression

$$S(\xi,\nu) = \sum_{i=1}^{N} P_i (2\pi B L_i)^{-1} \exp \left\{ -\frac{1}{2} \left[\frac{(r - r_1)^2}{L_i^2} + \left(\frac{f}{B} \right)^2 \right] \right\} \qquad (3)$$

In this formula, the parameter N represents the number of paths in the structure r_i and L_i are the mean delay and the multipath spread, B is the Doppler spread of the path, and P_i denotes the relative strength of the i^{th} path.

Further simplification can be achieved by representing the scattering function $S(\xi,\nu)$ as a single gaussoid, whose amplitude is a function of the path losses and whose widths L_{tot} and B_{tot} are respectively the total time spread and the total Doppler spread.

Because of the time variability of the path, we also require the knowledge of the statistical properties of all these parameters, so that we can compute their median values and plot the curves that provide the percentage of the time (y-axis) during which path losses, the multipath spread and the Doppler spread exceed the value of the abscissa (x-axis). Unfortunately only fragmentary data exist that are usable to this end.

First, let's review the case of a one-hop HF path in mid-latitude. Davies (11) gives a step-by-step procedure that can be easily followed and leads to a reliable estimate of the path losses. For a path with a length 2.5 megameters working at such a frequency f that 0.85 MUF4000 < f < MUF-4000, a typical value of path losses (inclusive of absorption losses and of antenna gains) is 130 dB.

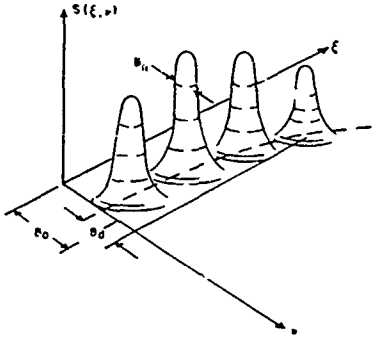


Figure 4.
Multimodal Scattering Function.

Few theoretical formulations and even fewer experimental data are available on the multipath spread and Doppler spread of single-hop HF paths in mid-latitude. Bailey (12) investigated the HF multipath spread phenomenon and its dependence upon operating frequency (specifically, upon its ratio to the MUF). He also included in his analysis the dependence of multipath spread on path length, link location, local time as well as season. This analysis shows that in a 2.5 megameter path, when using a working frequency of 0.9-MUF4000, the time spread L = 100 microsec. Reliable statistics on the time spread are available. Concerning now the Doppler spread of a mid-latitude one-hop HF path, we can estimate a median value of 0.1 Hz. For this parameter too, a reliable statistic is an unfilfill illed requirement.

Let's review now the case of a transauroral path. Lomax (13), Shaver et al. (14), and Shepherd and Lomax (15) have reported on experimental measurements of high-latitude HF propagation character-

istics, specifically for the frequency of 7.366 MHz. These authors found that winter propagation (inclusive of "normal" and "off-path" rays) is characterized by a time spread of approximately 1200 microseconds and by a Doppler spread of approximately 1 Hz. Summer propagation shows Doppler spread up to 20 Hz, although most of the signal energy was received within 0.5 Hz from the carrier. Data on transauroral-path losses at HF are scant, although these are certainly larger than in mid-latitude paths. In our study we have assumed two values for this parameter, 150 dB and 165 dB. An estimate of the percentage of time during which these values will be exceeded is, however, lacking.

As far as noise and interference is concerned, their levels at the two terminals of the link will not be the same, therefore these measurements must be performed at both stations.

Receiver front-end noise will be negligible at HF with respect to background noise (atmospheric and galactic) and with respect to man-made interference. Table 1 provides an estimate in dB above KTB of the atmospheric and galactic noise level at 20 MHz, for various seasons and hours of the day.

Table 1. Atmospheric and Galactic Noise (db above KTB)

	Hours	20 MHz	
		Atmo.	Galactic
Winter	0000-0400	22	22
	0400-0800	24	22
	0800-1200	28	22
	1200-1600	42	22
	1600-2000	37	22
	2000-2400	30	22
Spring	0000-0400	30	22
	0400-0800	25	22
	0800-1200	34	22
	1200-1600	45	22
	1600-2000	37	22
	2000-2400	39	22
Summer	0000-0400	22	22
	0400-0800	25	22
	0800-1200	25	22
	1200-1600	36	22
	1600-2000	34	22
	2000-2400	30	22
Autumn	0000-0400	25	22
	0400-0800	20	22
	0800-1200	36	22
	1200-1600	38	22
	1600-2000	40	22
	2000-2400	36	22

Interference into the receivers of the link from nearby transmitters, or from co-channel emissions, either at close range or at a distance, as well as man-made noise will be the predominant factors in establishing the overall signal-to-(noise + interference) ratio at a given carrier of the HF link. As it will be discussed later on in this paper, the adaptive scheme that we propose will use a waveform characterized by the presence of numerous spectral lines and the noise + interference level at each one of them will be verified in order to exclude the most interfered ones.

After review of CCIR Report No. 65, we have adopted in our analysis two values for the level of noise + interference; +30 dB and +45 dB above KTB.

Table 2 summarizes the channel properties that we have assumed as educated guesses for our study of adaptive HF propagation path utilization.

Table 2. HF Channel Properties

	Midlatitude Path	Transauroral Path
Multipath Spread	100 μsec	1500 μsec
Doppler Spread	0.1 Hz	10 Hz
Path Losses	130 dB	{150 dB 165 dB
Noise + Interference (above KTB)	{30 dB 45 dB	{30 dB 45 dB

3.3 Path Sounding

Path sounding has the scope of measuring path losses at an adequate number of spot frequencies in the band of interest (3 MHz to 30 MHz) and of measuring at the same time noise and interference levels, at the same frequencies and at both ends of the link.

Table 3 gives the parameters of the proposed sounding scan. The master station radiates sequentially 1125 to 3750 carriers to cover the 3-MHz to 30-MHz band, in a time interval 100 to 160 seconds (88 milliseconds to 47 milliseconds per carrier). Of the two numbers given above for each sounding parameter, the first applies to a mid-latitude path, the second to a transauroral path. The scan is repeated every 5 to 8 minutes.

Once a set of frequencies has been chosen for communicating, it is automatically excluded from next sounding cycle. However, information on the channel status for each one of the frequencies thus excluded from sounding and probing is still updated once every 5 to 8 minutes by measurements performed on the communication waveform. Frequency switching is preceded by a "tone" of notification and takes place even while communications go on, for the case in which the channel deteriorates and another set of frequencies is found more suitable for carrying out the communications.

The block diagram in Figure 5 has been worked out for Link Configuration 1 (two-way adaptive link) and illustrates the various functions of the two terminals of the link. Here, one-way sounding and probing is achieved from the master station to the slave station. Acknowledgement is from the slave to the master station. Finally, communication is a two-way exchange between the stations.

Table 3. Sounding/Probing Scan Parameters

	Midlatitude Path	Transauroral Path
Band covered	3 MHz-30 MHz	3 MHz-30 MHz
Number of spot frequencies	1125	3375
Separation between two adjacent spot frequencies	24 KHz	8 KHz*
Sounding scan time	100 seconds	160 seconds
Rate of sounding scan repetition	one every 300 seconds	one every 480 seconds
Dwelling time per spot frequency	88 millisec	47 millisec
Nominal bandwidth of sounding receiver	24 KHz	8 KHz
Width of sounding pulse	41.5 microsec	125 microsec
Pulse repetition frequency	100 pps	100 pps
Pulses per dwelling time	8 pulses	4 pulses
PRF duty cycle	10^{-3}	$1.25\ 10^{-2}$
Noise and interference measurements integration time, for each spot frequency	53 millisec	47 millisec
Overall noise and interference measurement time	60 seconds	160 seconds

*This value is chosen because 8 KHz is the bandwidth of the signal waveform selected for the transauroral link. The path coherent bandwidth is only 666 Hz.

Mid-latitude link

Step 1 - 100 seconds devoted to sounding operation.

Step 2 - 60 seconds devoted to measurement of noise and interference at both terminals of the link.

Step 3 - 20 seconds devoted to computations, taking into account the need of accumulating at a single terminal (the slave station) the information pertaining to noise and interference at both terminals. During this step, the microprocessor at the slave station selects the frequencies and designates them to the master station.

Step 4 - 100 seconds devoted to channel probing, to be performed only at the frequencies designated by Step 3.

Step 5 - 20 seconds devoted to computations, acknowledgement and information exchange between the two terminals, in order to perform the final selection of frequencies to be used in communications, by taking into account the data on multipath spread and Doppler spread.

Step 6 - The two terminals are now ready to initiate communications. The frequencies finally adopted for communications are excluded from next sounding/probing cycle (one every 300 seconds), although they continue to be monitored by measurements on the modulation waveform.

Transuroral link

Step 1 - 160 seconds

Step 2 - 160 seconds

Step 3 - 20 seconds

Step 4 - 120 seconds

Step 5 - 20 seconds

Step 6 - The two terminals are now ready to initiate communications. The sounding/probing cycle is repeated every 8 minutes (480 seconds).

Communications are therefore inhibited only in the first 300 (or 480 seconds) of link operation. After this initial adaptive adjustments of the link's terminals, any readjustment is performed without requiring a discontinuation of communications.

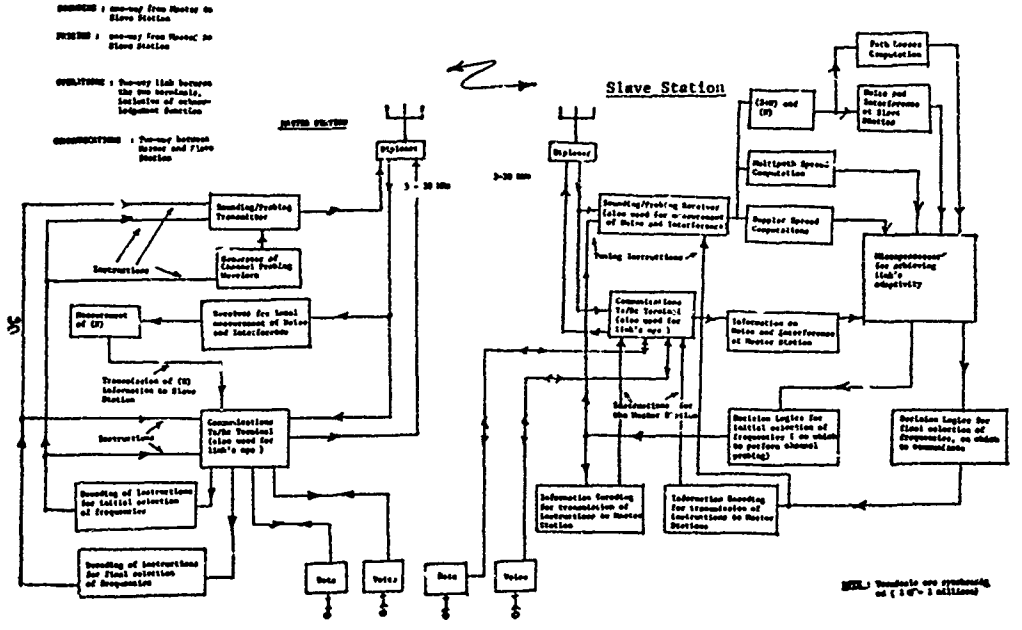


Figure 5. Simplified Block Diagram of Link's Terminals

3.4 Channel Probing

The channel probing is aimed at gathering information on the time and frequency dispersive effects of the HF propagation, after the path sounding has determined path losses and noise plus interference levels at the available spectral lines, and has identified the frequencies promising enough to be worthy of the channel-probing effort. All these functions are slowly varying functions, so that one sample very 5 to 8 minutes is adequate.

The measurement of multipath spread and of the Doppler spread can be achieved with a variety of methods, either based on the direct measurement of these two quantities or on indirect measurements such as the ones based on the fact that, at a given frequency, the reciprocal of the Doppler spread gives the e.m. wave fading period or that the reciprocal of the multipath spread, at a given instant in time, gives the frequency interval within which carriers fade coherently. Because the amount of time required to process the information on the dispersive properties of each channel is not trivial, we propose to perform these measurements only for those frequencies for which path sounding has indicated acceptable path losses and affordable noise and interference levels.

Measurement of the Doppler-spread is based on the fact that, when a CW tone is transmitted, the received process is narrowbnad Gaussian due to doppler spreading characteristics of the channel. Since the Doppler-spread is the rms bandwidth (second central moment) of the Doppler spectrum of the complex envelope of the narrow-band Gaussian process, the methods summarized in (16) can be used. These methods utilizes the inphase and quadrature components, $\alpha(t)$ and $\beta(b)$, of the complex envelope of the received process.

The inphase and quadrature components can be determined by multiplying the received carrier by both a local carrier and a 90° shifted local carrier at the same frequency as the received carrier (or as near to the same frequency as possible) and then extracting the low-frequency components. Strictly speaking, Doppler spread is independent of the Doppler-shift, thus precise knowledge of the received carrier frequency is not necessary. However, as the local carrier frequency departs from the received carrier frequency, the extracted $\alpha(t)$ and $\beta(t)$ increase in bandwidth, necessitating larger bandwidth filters and passing more noise. Thus, from the point of view of maximizing signal-to-noise ratio, it is desirable to keep the local carrier frequency as near as possible to the received signal frequency. If the transmitter and receiver oscillators are not phase coherent, the Doppler-spread can be measured by the use of only the envelope or more generally any well-behaved nonlinear functions of the envelope of the received carrier (17). Fading rates have also been determined by measuring the average number of times/unit time the envelope of the received carrier crossed a specified level, or by computing the zero crossings of the inphase (or in quadrature) component of the narrowband Gaussian process [see ref. of Rice in (16)].

The multipath spread parameter is a measure of the dispersion in path delays suffered by a process propagated through a random channel. From a strictly mathematical point of view, the multipath spread measurement problem is entirely analogus (dual) to the Doppler-spread measurement problem. The processing is now done on the complex amplitude spectrum of the received transient, similar to the methods discussed for the measurement of Doppler spread.

The techniques discussed earlier for the instantaneous measurement of Doppler and multipath spread require either the extraction of complex envelopes or envelopes of the received carriers. In (18), FM and SSB techniques have been discussed which simultaneously measure both the parameters from the envelopes of the received carriers. In SSB (single side band) technique, two carriers separated by F Hz are transmitted and from the received detected envelopes, represented by E_t (f+F) and E_t (f), the Doppler spread and multipath spread are measured simultaneously as shown in Figure 6.

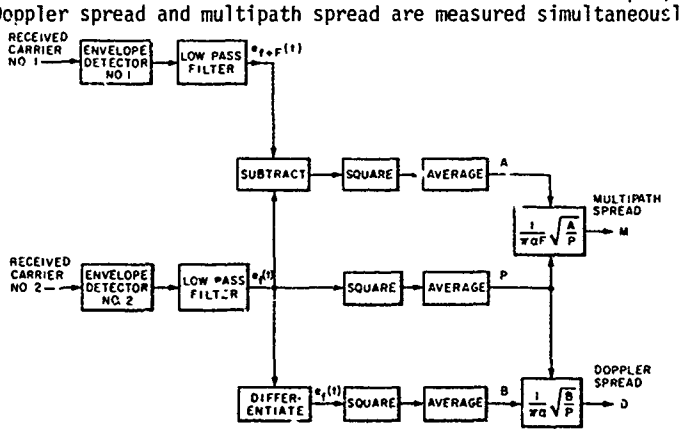


Figure 6.

Simultaneous Measurement of Doppler Spread and Multipath Spread Using Envelopes Only by the SSB Techniques.

4. COMMUNICATION METHOD, ADAPTIVITY AND DIVERSITY

4.1 Discussion of Diversity and Adaptivity

The signal design approach adopted in this paper makes it possible to provide a data rate of 24 kbits/sec with an error probability of 10^{-4}, therefore suitable for high quality digital voice. The system is adaptive in the sense that, at any given time, it makes use of sounding information and of the probing information to translate the transmitted spectrum around the carrier which yields the best signal-to-noise ratio. The use of relatively wideband, aside from allowing the transmission of a high data rate, provides inherent diversity and therefore protection against fading. Furthermore, since the transmitted energy is spread over a considerable range of frequencies, the interfering effects of the communication waveform on other systems operating in the same frequency interval will be obviously reduced.

The communication method under discussion [see ref. (19) for details] is based on dividing the time-frequency plane in intervals of duration $\frac{1}{B}$ (B is the Doppler spread) in time axis and in intervals of duration $\frac{1}{L}$ (L is the multipath spread) in frequency axis. Each rectangle represents a portion of the time-frequency plane with approximately correlated fading. This method is applied to two cases. Case 1 corresponds to a mid-latitude link and Case 2 corresponds to a transauroral link. The pertinent properties of these links are summarized in Table 2. The diversity situation is very different in the two cases. In Case 1, the basic diversity cell is a rectangle of size $\frac{1}{L} - \frac{1}{B} = 10^4$ Hz x 10 sec and in Case 2, the cell size is 667 Hz x 10^{-1} sec.

To start with, we use FSK modulation. We will then extend this analysis to MFSK signalling. We assume that the two transmitted FSK signals (marks and space) are composed of n chips, arranged in time sequence, in parallel, or in some arbitrary series-parallel combination. We also assume TW = 1 for the radiated waveform. Let E be the total energy received in all n chips. The diversity per chip $\mathcal{Z}$ is given by the expressions contained in Table 4. If the mark (space) contains n chips, the total signal diversity is $\mathcal{Z}n$.

If	Then
BT < 1 and LW < 1	$\mathcal{Z} = 1$
BT < 1 and LW > 1	$\mathcal{Z} = LW$
BT > 1 and LW < 1	$\mathcal{Z} = BT$
BT > 1 and LW > 1	$\mathcal{Z} = BL$

Table 4.
Expressions to be Used in the Computation of the Diversity Z

In the case of a binary alphabet, the probability of error is given by

$$Pe \cong 0.2\, e^{-\frac{1}{2}\frac{E}{N_0}A(y)} \tag{4}$$

where N_0 is the noise-power density (watts/wz) and the A(y) is efficiency function plotted in Figure 7. Pe is minimum when A(y) is maximum which occurs at $y \cong 0.35$, where $A(y) \cong 0.3$.

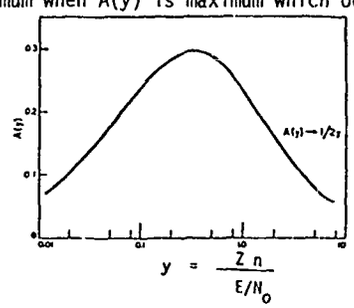


Figure 7.
Efficiency Function A(y).

For Case 1, in order to achieve $Pe \cong 10^{-4}$, from (4) with A(y) = 0.3, $\frac{E}{N_0} \cong 55$.

$$\frac{E}{N_0} \cong 55 \tag{5}$$

Since $y = \frac{\mathcal{Z}n}{E/N_0} \cong 0.35$, we have

$$\mathcal{Z}n = 0.35 \times 55 \cong 20 \tag{6}$$

In order to find n, we must determine BL + LW so that we can find $\mathcal{Z}$ from Table 4. Since R = 24 kbits/sec,

$$T \cong \frac{1}{R} = 4.15\ \mu\text{sec and } W = \frac{1}{T} = 24\ \text{kHz} \tag{7}$$

Thus

$$BT = 0.1 \times 41.5 \times 10^{-6} < 1 \quad \text{and} \quad LW = 10^{-4} \times 24 \times 10^{3} > 1 \tag{8}$$

Therefore from Table 4,

$$\not{Z} = LW = 2.4 \tag{9}$$

From (6) and (9), one obtains

$$n = \frac{20}{\not{Z}} \cong 9 \tag{10}$$

This diversity, required to counteract fading, can be achieved for instance by signalling simultaneously on nine carriers, each with a 24-kHz bandwidth. Minimum separation between two adjacent carriers is also about 24 kHz.

Counteracting fading is not, however, the only probelm that we have to solve in order to reliably transmit at 24 kbit/sec data rate. In fact, our basic bit length is 41.5 μsec, while the time spread of the channel is 100μsec. Consequently, we need a number of groups of frequencies (each with nine lines) as computed from

$$\frac{T + L}{T} = 1 + \frac{100}{41.5} \cong 4 \tag{11}$$

Thus one can compute the total bandwidth occupancy of the link as

$$\underset{\substack{\text{mark \&}\\ \text{space}}}{\underline{2}} \times \underset{\substack{\text{carriers}\\ \text{per group}}}{\underline{9}} \times \underset{\text{groups}}{\underline{4}} \times \underset{\substack{\text{bandwidth}\\ \text{per carrier}}}{\underline{24\ \text{kHz}}} = 1.728\ \text{MHz} \tag{12}$$

Total number of carriers used = 2 x 9 x 4 = 72.

The above figures apply to simplex communications and must be multiplied by 2 to achieve duplex connection. To estimate the required transmitted power for the two cases of noise and interference levels, one has

$$N_0 = 10^{3}\,(kT) = 4.14 \times 10^{-18}\ \text{watts/Hz}$$

The required signal power is therefore

$$P_2 = \frac{55\ N_0}{T} = 5.48 \times 10^{-12}\ \text{watts}$$

Since path losses for Case 1 are 130 dB, inclusive of antenna gains, the transmitted power is

$$P_T = 10^{13}\ Pr = 55\ \text{watts} \tag{13}$$

Under the assumption that noise and interference amount is +45 dB above kTB,

$$P_f = 1.74\ \text{kwatt} \tag{14}$$

For Case 2 (see Table 2), we have

$$B_T = 1041.5 \times 10^{-6} < 1 \text{ and } LW = 1.5 \times 10^{-3} \times 24 \times 10^{3} > 1$$

Therefore $\not{Z} = LW = 24\ 1.5 = 36$. Table 5 illustrates the transmitter power requirements in this case. Since $\not{Z}n = 20$, a single chip (n=1) is more than enough. The number of carriers per group is decreased from Case 1 to Case 2, from 9 to 1; however, there is a severe deterioration in intersymbol interference. Since $\frac{T+L}{T} = 1 + \frac{1500}{41.5} \cong 38$, we have 38 independent group of frequencies. The estimate of total bandwidth occupancy is

$$2 \times 1 \times 38 \times 24\ \text{kHz} = 1.82\ \text{MHz} \tag{15}$$

with 2 x 1 x 38 = 76 number of carriers.

Path Losses	Noise and Interference Level	Required Transmitter Power
150 dB	+30 dB above KTB +45 dB above KTB	5.5 K watts 174 K watts
165 dB	+30 dB above KTB +45 dB above KTB	174 K watts 5.5 M watts

Table 5.
Transmitter Power Requirements in Transauroral Paths for Binary Waveforms.

Not withstanding the large differences between Case 1 and Case 2, as far as the diversity condition is concerned, there is a close similarity between the two cases in terms of total bandwidth occupancy and total number of carriers used. These results give a clear indication how the adaptivity of the system could work. If we have available for instance 76 spectral lines in the HF band,

each with a bandwidth of 24 kHz, we could arrange 72 of them into four groups of nine carriers when the path is Case 1 type or we could use all 76 of them (Case 2) in 38 groups, each with a single carrier. This adaptivity adjustment to path conditions could be made based on the result of path sounding and channel probing.

4.2 The Use of M-Ary Codes

A substantial improvement in the link's performance can be obtained by using M-ary codes (or MFSK signalling). In our study we have considered M=8, which makes it possible to achieve a data rate of 24 kilobit/sec with a pulse length $T = 41.5 \log_2 8 = 125\mu$sec and bandwidth W = 8 kHz. Since

$$P_{bit,M} = 2^{k-1} P_{bit,\ binary} \tag{16}$$

then for $P_{bit,\ M=10^{-4}}$

$$P_{bit,binary} = \frac{10^{-4}}{4} = 2.5 \times 10^{-5} \tag{17}$$

Now following the stpes of (4 to 10), one can obtain $\bar{Z} = 1$ for Case 1 and $\bar{Z} = 12$ for Case 2. Since $Z_n = 20$ for both cases, we need 20 carriers, each with a bandwidth of 8 kHz and 2 such groups for Case 1 and 2 carriers with the same bandwidth and 13 groups in Case 2 to counteract fading and inter-symbol interference. To obtain an estimate of the total bandwidth occupancy required by the MFSK (with M=8) link, we fave for Case 1

$$2 \times 8 \times 20 \times 2 \times 8 \text{ kHz} = 5.12 \text{ MHz} \quad \text{and} \quad 2 \times 8 \times 2 \times 13 \times 8 \text{ kHz} = 3.3$$

MHz for Case 2. Similarily for simplex link, the total number of carriers used are 640 and 416 respectively for Case 1 and 2.

Concerning the chip energy, we point out that in the M-ary codes, the time duration of the chip is $\log_2 M$ longer than in the binary case. Therefore we need smaller (by the factor of $\log_2 M$) E/N_0 than computed as in (20). The required levels of transmitted power are given in Table 6. It can be seen that there is a factor of almost 10 improvement with respect to the binary-system approach, at the expense of bandwidth occupancy and equipment complexity.

Path Losses	Noise and Interference Level	Required Transmitter Power	Radiated Power Density	Equivalent Power in a 3-KHz Channel
Midlatitude Path				
130 dB	+30 dB above KTB	6.67 watts	$1.3\ 10^{-6}$ watts/Hz	4 milliwatts
	+45 dB above KTB	211 watts	$4.11\ 10^{-5}$ watts/Hz	123 milliwatts
Transauroral Path				
150 dB	+30 dB above KTB	667 watts	$2.4\ 10^{-4}$ watts/Hz	0.6 watts
	+45 dB above KTB	21.1 Kwatts	$6.34\ 10^{-3}$ watts/Hz	19 watts
165 dB	+30 dB above KTB	21.1 Kwatts	$6.34\ 10^{-3}$	19 watts
	+45 dB above KTB	667 Kwatts	$2\ 10^{-1}$	600 watts

Table 6. Transmitter Power Requirements in HF Paths for M-ary Transmission With M = 8.

*Note: Pulsewidth = 125 microsec

The power densities (watts/Hz) radiated by the proposed M-ary coded emissions are given in the fourth column of Table 6. The equipment power of a transmitter that would generate equal power density in a 2-kHz voice channel is given in the fifth column of the same table. These levels are relatively low, with the exception of the one on the last line. We have also to point out that our transmitter would be even less bothersome because the emissions would not stay consistently on the same frequencies, but would wander around to follow the adaptivity instructions generated by the microprocessors and control logies.

5. NON-ADAPTIVE, ONE-WAY LINK FOR INFORMATION TRANSFER WITHOUT FEEDBACK ACKNOWLEDGEMENT

There are cases in which it is neither operationally possible nor advisable to establish a two-way link with feedback acknowledgement between two terminals. This case is HF communication is normally handled by repeating the message on several frequencies.

Another alternative, that uses some of the concepts of M-ary transmission discussed in Section 4, would be to spectrally spread the information (in the limit, to the entire HF band) so that even if a portion of the HF band is lost, still the residual waveform arriving to the receiver could carry intelligible information. An example of this approach is contained in Section 2.7 of Gupta and Grossi, 1980 (7).

6. APPLICATION OF ADAPTIVITY AND M-ARY CODES TO OTHER DISPERSIVE CHANNELS, SUCH AS METEOR BURST CHANNEL

The adaptivity approach discussed in this paper, although general, is specifically referred to two cases of HF communciations: a mid-latitude link and a transauroral link. These links are characterized by different path losses, doppler and multipath spread. Application of this approach to a meteor burst channel, characterized by time-varying path losses, time-varying Doppler and multipath spread is particularly of interest. For a 4MHz system and assumed link and error rate ($\leq 15^{-3}$), Grossi and Javed (20) obtained the data rate which decreased from the time of trail's formulation to the end of the trail life time.

This limitation of decreasing data rate can be removed by the use of M-ary codes. Since the data rate for M-ary signalling is $\log_2 M$ times that of binary modulation, variable M-ary signalling can be adopted to counteract the increasing multipath spread. Larger the multipath spread, larger the value of M is required to maintain the constant data rate at a specified performance over the entire trail life. Thus higher data rate can be achieved for meteor burst link with the use of variable M-ary signalling.

7. CONCLUSIONS AND RECOMMENDATIONS

The conclusions of our study on adaptive utilization of HF propagation paths indicate that by spreading the radiated power across several MHz of bandwidth occupancy, it is possible to counteract effectively both waveform fading and inter-symbol interference. By this method, data rates of 24 Kbit/sec with 10^{-4} error rates are possible. If this is confirmed by the proposed experimentation, HF would acquire a reliability and channel capacity of the degree that is enjoyed by other communications media. The penalty that must be paid to achieve these results is equipment complexity and bandwidth occupancy. Concerning the first point, modern advances in microprocessor technology, high-density packaging, frequency agility, etc., offer concrete promises. As far as the second point is concerned, the power of the transmitter is so spread that the link would hardly interfere (and for not very long time intervals, because of frequency wandering) with a receiving site.

Our recommendation is that additional R&D activity be performed in three basic directions: a deeper understanding of the propagation properties of the two types of paths considered, a gathering of experimental data, and an engineering study of the availability and applicability of such modern technological breakthroughs as microprocessors, frequency-agile transmitters and receivers, broad-band HF antennas and matching units, switch circuits, decision logics, etc.

8. ACKNOWLEGEMENTS

The authors thank the Air Force Geophysics Laboratory for having sponsored this effort with contract F19628-79-C-0132 to the Smithsonian Astrophysical Observatory. In particular they wish to gratefully acknowledge the technical guidance provided by Dr. Jules Aarons who conceived the project and supervised the effort for AFGL.

9. REFERENCES

(1) Baghdady, E.J. (1969) Principles of simulation of randomly time-varying channels, IEEE International Conference Communications, Boulder, Colorado, June 9-11, pp. 40-3 to 40-18. (2) Price, R. and Green, P.E. (1958) A communication technique for multipath channels, Proc. IRE, pp. 555-570. (3) Kailath, T. (1959) Sampling Models for Linear Time-Variant Filters, MIT-RLE Report No. 352, Cambridge, Mass. (4) Esposito, R. and Grossi, M.D. (1971) Channel Characterization for Digital Communications in Ground-to-Space HF Paths, 11th Technical Meeting of the Joint Satellite Study Group (JSSG), Florence, Italy, October 4-9. (5) Gupta, A.K. (1979) Complex Envelope Simulation of Scatter Channels, presented at the IEEE National Telecommunications Conference, Washington, DC, November 27-29. (6) Lomax, J.B. (1970) HF Propagation Dispersion, in "Phase and Frequency Instabilities in Electromagnetic Propagation," K. Davies, Editor, pp. 497-510, AGARD Conference Proceedings, Number Thirty-Three, Techvision Services, Slough, England. (7) Gupta, A.K. and Grossi, M.D. (1980) Adaptive HF Propagation Path Utilization, AFGL-TR-79-9236, 30 June, Final Report on contract F19628-79-C-0137 to Smithsonian Astrophysical Observatory, Cambridge, Mass. (8) Bello, P.A. (1969) Measurement of random time-variant linear channels, IEEE Trans. on Info. Th., Vol. IT-25, No. 4, July. (9) Bello, P.A. and Esposito, R. (1970) Measurement techniques for time-varying dispersive channels, Alta Frequenza, No. II, Vol. XXXIX, pp. 980-996. (10) Gupta, A.K. (1979) Dissertation abstract, IEEE Trans. Info. Theory, p. 760, November. (11) Davies, K. (1965) Ionospheric Radio Propagation, NBS Monograph No. 80. (12) Bailey, D.K. (1959) The effect of multipath distortion on the choice of operating frequencies for HF communications circuits, IRE Trans. Antenna and Propagation, AP-7, 398. (13) Lomax, J.B. (1970) HF propagation dispersion, in Phase and Frequency Instabilities in Electromagnetic Wave Propagation, K. Davies, editor, pp. 497-510 AGARD Conference Proceedings No. 33, Techvision Services, Slough, England. (14) Shaver, H.N., Tupper, B.C., and Lomax, J.B. (1967)

Evaluation of a Gaussian HF channel model, IEEE Trans. on Communications Technology, Vol. 15, No. 1, pp. 79-88. (15) Shepherd, R.A. and Lomax, J.B. (1967) Frequency spread in ionospheric radio propagation, IEEE Trans. on Communications Technology, Vol. 15, No. 2, pp. 268-275. (16) Gupta, A.K. (1981) "On spectral moments of a narrow-and Gaussian process from its distorted and undistorted quadrature components," Proceedings of the 1981 Conference on Information Sciences and Systems, held at the John Hopkins University, Baltimore, Maryland, March 25-27, 1981. (17) Bello, P.A. (1965) On the rms bandwidth of non-linearly envelope detected narrowband Gaussian noise, IEEE Trans. on Info. Theory, pp. 236-239. (18) Bello, P.A. (1965) Some techniques for the instantaneous real-time measur measurement of multipath and Doppler spreads, IEEE Trans. on Comm. Tech., Vol. 13, No. 3, pp. 155-192. (19) Kennedy, R.S. and Lebow, I.L. (1964) Signal design for dispersive channels, IEEE Spectrum, Vol. 1, pp. 231-237. (20) Grossi, M.D. and Javed A. (1977) "Time and frequency spread in meteor burst propagation paths," AGARD paper No. 244.

MULTIPATH MEASUREMENTS IN THE ATHENS-SALISBURY T.E.P LINK

G.Stephanou, C.Caroubalos and N.Corallis
University of Athens, Department of Physics
Division of Electronics , ATHENS / GREECE

ABSTRACT

In this paper a systematic investigation of radio pulses propagation time, as well as ,of double-pulses delay time is reported, concerning the Athens-Salisbury (6250 Km) T.E.P link at 34.0 MHz.

A computer simulation, by a ray tracing program is employed, in order to elucidate the path type that could realize the link, (in the case of a simple or a double-pulse) and also to estimate the corresponding propagation times.

The results calculated in this way, are compared with the experimental ones and are used to interpret the seasonal variation of the relative delay times of the second pulse.

1. INTRODUCTION

Over the past thirty years reports have been accumulated in the literature dealing with unusual VHF conditions in and across the equatorial regions. In spite of this, the evidence is neither numerous nor concise enough in order to lead to a generally accepted theory of mode structure for Trans-Equatorial Propagation (T.E.P) . Besides the complexity of the ionosphere over the equatorial regions this has also been inhibited because of difficulties associated with coordinating evidence and theory, based upon test programs widely spaced in time, geographic location and primary objective.

One of the methods that can be used to deduce characteristics of the T.E.P path mode structure is measuring the absolute time of flight (propagation time) of radio pulses travelling between two experimental sites. When this is done, the existence of multipath is associated either with the broadening and / or the doubling of received pulses.

While double-pulse reception, in T.E.P, has been reported by several researchers, systematic investigation of this has not yet been undertaken. Moreover the existence of the second pulse has not yet been adequately explained.

The ^{n}F transmission model proposed by Villard et al. (1957) to explain T.E.P , is generally accepted , among the investigators, for the occurence of a single pulse or the first pulse of a douple pulse configuration. On the contrary a variety of suggestions have been proposed for the second pulse .

According to BOWEN et al. (1968) the second pulse can be attributed to field-guided paths proposed by OBAYASHI (1959).

ANASTASSIADIS and ANTONIADIS (1972) accept the possibility of a common 2F path to explain multipath propagation observed in the Athens-Roma T.E.P link. The second pulse is also attributed (VASSILARAS 1973) to the high angle (of arrival) rays of a ^{2}F path.

Magnetoionic splitting and off-great circle propagation have also been considered (CARMAN et al . 1973) responsible for the various signal fading-forms recorded in the Athens-Roma link.

In this paper a systematic investigation of radio pulse propagation times, as well as double-pulse

configuration delay times is reported. The computer simulation results obtained for the interpretation of our observations are also presented and discussed in comparison with them.

2. EXPERIMENTAL SET UP AND RESULTS

The experimental arrangement for the measurements reported here, has been briefly described by ANASTASSIADES and ANTONIADIS (1972) . Figure I shows the propagation path involved.

In this experiment, the frequencies employed for transmission from Athens to Salisbury and vice versa are 34.0 MHz and 33.8 MHz respectively. The low-power (I00 W) pulse transmitters at each end of the link were coupled to horizontal Yagi autennas. The outputs of the receivers at the Athens receiving centre were connected to a 556 Tektronix oscilloscope used on the delayed sweep mode (see Fig. 2)

With a pulse width of 200 μsec, this arrangement provides propagation time measurements accurate to $\pm$50 μsec for the double path propagation time, of the order of 43 msec. This accuracy is determined by the resolution that can be achieved with the receiver bandwidth of 8 KHz, and a sweep rate of 0.5 msec cm^{-1} . The relative delay time of the second pulse, if any, was measured directly on the oscilloscope.

The data collected cover four consecutive equinox periods from October I973 till March I975. Note that I975 was a year of low solar activity. T.E.P occurence and signal strength have a maximum at equinoctial periods and consequently any kind of interference is of least significance during these periods.

From 5652 sample measurements taken during the above periods, a total of 367 cases refer to double pulse reception. A third pulse was also observed in 9 cases.

Figure 3 shows the distributior of measured propagation times, for both the single and the second pulse. The main statistical parameters of these distributions are given in Table I.

TABLE I

	Single pulse	First pulse	Second pulse
Number N	5652	367	367
Mean Value $\bar{\tau}$ (msec)	2I.50	2I.50	2I.82
Stand.Deviat. s(msec)	0.I2	0.I2	0.I7

The second column of the table is refered to the first pulse of a double-pulse configuration,while the third column is refered to the second one. Comparison of columns I and 2 shows that the appearence of a second pulse does not influence the propagation times of the first one. Therefore we may accept that the second pulse is a circumstrantial consequence of the physical situation supporting this type of propagation.

The distribution of the second pulse delay times is shown in figure 4. The mean delay time (320 μsec) corresponds to a group path-difference of almost I00 KM.

The delay times of the second pulse did not show any time dependence during a day. On the contrary their distributions exhibited a significant seasonal variation consisting of a diminution of delay times as it is shown in figure 5.
To the best of our knowledge, this phenomenon, has been observed for the first time. It may be related to sunspot-cycle variations (see also discussion) .

3. COMPUTER SIMULATION

In order to better understand how the data presented could be interpreted, we proceeded with simulation tests using the Jones-Stephenson (I975) Ray Traycing Computer Program .

Since our measurements refered mainly, to presunset hours of the day (during which it is generally accepted that east-west gradients do not exist), we confined our study to north-south propagation across the geomagnetic equator. Moreover since it is not possible to define the ionosphere accurately enough from one day to the next, we studied order of magnitude effects, using the following smoothed models .

The electron density model is a Chapman-layer whose parameters have a latitude dependence given by the following analytical expressions (see also [7]):

$$f_c^2 = \left\{\left(\frac{f_{max}}{2}\right)^2 - \left(\frac{f_{co}}{2}\right)^2\right\}\left(\frac{4\lambda}{W}\right)^2 \exp\left\{2 - \left|\frac{4\lambda}{W}\right|\right\} + f_{co}^2$$

$$H_{max} = H_o + 80\left[1+\cos\left(\frac{180}{24}\lambda\right)\right] \quad \text{for} \quad |\lambda| < 24$$

$$H_{max} = H_o \text{, KM} \quad \text{for} \quad |\lambda| \geq 24^\circ$$

where : λ : geomagnetic latitude

f_{cmax} : critical frequency of the anomaly crest

f_{co} : critical frequency of the trough

W : distance between crests (in geomagn. latitude)

H_o : Hmax for $|\lambda| \geq 24^\circ$

The various parameters of the equatorial anomaly were changed between the following limits.

f_{cmax} = 9 to I6 MHz

f_{co} = 6 to 8 MHz in steps of 7 Mhz .

W = 45° to 75° dip angle in steps of I0°

H_o = 200 to 300 KM in steps of 50 KM

Finally the earth-centered dipole was used as an approximation to the real geomagnetic field.

A large number of rays have been calculated with these models, for transmitter located at 30°north latitude. Since this program does not include a homing feature, we interpolated linearly between adjacent landing points, in order to estimate group paths corresponding to a surface distrance of 6250 KM (distance between Athens-Salisbury).

4. DISCUSSION

The results of the calculations are sammarized in Table 2 .

TABLE 2

Path type	Calculated Times (msec)
^2F	2I.35 - 2I.64
2F	2I.58 - 22.20

Comparison between calculated and observe propagation times (see fig. 3) shows that:

a. The calculated values of the propagation times are in good agreement with the measured ones , for the case of a single pulse or the first pulse of a double - pulse configuration .

For these cases, the existence of ^2F paths proposed by VILLARD et al. is the most prevalent mode of

propagation.

b) The second pulse of the double-pulse cannot be explained by the ^2F paths. However this can be done by considering paths with an intermediate ground reflexion(2F)

In several cases, in the same simulation test, rays transmitted at different directions were landed at the same distance from the transmitter (6250 KM).The calculated delay times (differences in propagation times of any two such rays, see fig. 6) are summarized in Table 3, for the various modes ("φ" : low angle , "Φ" : high angle of arrival) of both wave components ("O" and "X"). Comparison of the later results with experimental data shows that:

a) The calculated delay - times cover the range of the experimental values. However, none of the various classes of this table can, by itself, explain the whole set of observations. It is therefore suggested that the second pulse may be explained as a combination of the various classes, in Table 3.

b) The upper half of the Table is refered to rays of the same path-type (^2F or 2F). Differences larger than 240 μsec, cannot be explained by these rays, that however, can account for the broadening of the received pulses.

c) Rays of different path types, (2F and ^2F) gave delay times in the range from 240 to I040 μsec. The majority of the observed values (73 %) lies in this range. In all cases, the second pulse corresponded to a 2F path.

Note that in the simulation tests, the 2F type was occuring less frequently as the critical frequencies of the F-layer were decreased. Since the critical frequencies of the F-layer equatorial anomaly are known to be positively correlated with the solar cycle, it may be suggested that, the progressive reduction of delay times observed in our experiment from I973 to I975 (see fig.5), is due to the effect of a corresponding reduction of solar activity.

5. CONCLUSIONS

The propagation times of radio pulses in a TransEquatorial Porpagation Link between Athens and Salisbury, have been studied, both experimentaly (using the forward propagation pulse technique 4) and by computer simulation (using a Ray Tracing program 7 and plausible models for the propagation medium).

This comparative study is proved to be efficient in determining the dominant mode in the type of propagation. With this method we have obtained satisfactory explanation for the doubling observed in the received pulses, in T.E.P , during multipath conditions.

The above results contribute to a better understanding of the propagation mechanism in T.E.P links Besides, they may be usefull in any attempt to apply pulse code modulation techniques in communication via TransEquatorial Propagation.

AKNOWLEDGEMENTS

The authors express their many thanks to professor Ray Stevenson (University College of Rhodesia), for his assistance and all the offered fascilities for the installation of measuring equipment in Salisbury.

We also thank the collegnes of the Ionospheric Institute of the Athens Observatory and the group of Ionospheric Research of the Electronics Laboratory (University of Athens) for many usefull discussions.

REFERENCES

I.	Villard O.G, Stein S. , Yeh K. G	I957	J.G.R , 62 , p 399
2.	Bowen E. D, Fay W.J, Heritage J. L	I968	J.G.R 73, 7, p 2469
3.	Obayashi T.	I959	J. Radio Res. Lab. (Japan), 6, p 603
4.	Anastassiadis M, Antoniadis D.	I972	J.A.T.P,34, 7 , p I2I5
5.	Vassilaras J.	I973	Thesis, Athens University
6.	Carman E. W , Heeran M. P, Anastassiadis M.	I973	J.A.T.P,35, p I23I
7.	Jones M, Stephenson J.	I975	O.T Report

TABLE 3

				SECOND PULSE DELAY TIME (msec)
RAYS OF THE SAME TYPE	^2F	O-X		*0.01-0.24*
		φ-Φ	O	*0.01-0.13*
			X	*0.05-0.08*
	2F	O-X		*0.02-0.16*
		φ-Φ	O	*0.21*
			X	—
RAYS OF DIFFERENT TYPE		O-X		*0.40-1.04*
		φ-Φ	O	*0.27-0.99*
			X	*0.24*

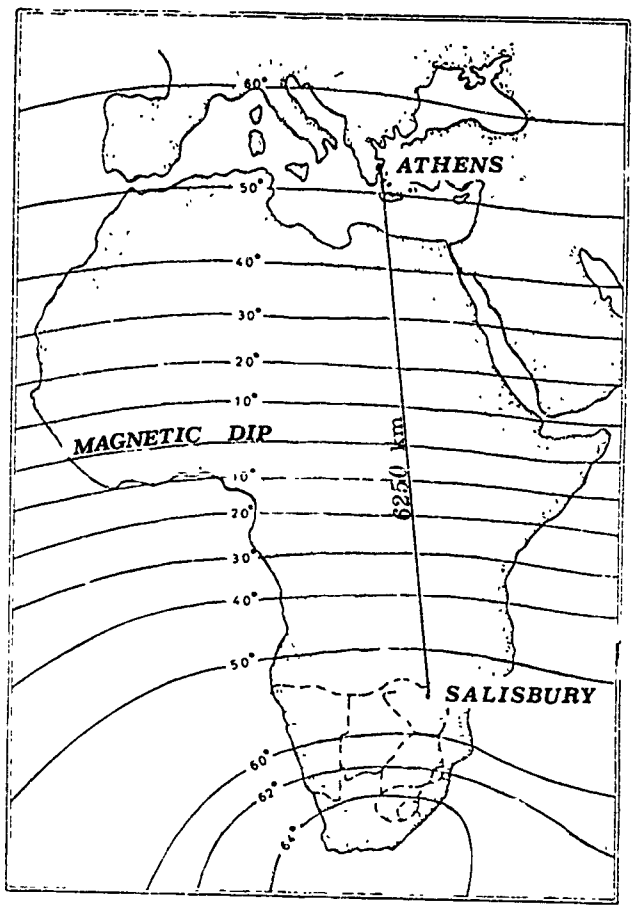


Fig. I. Propagation path geography

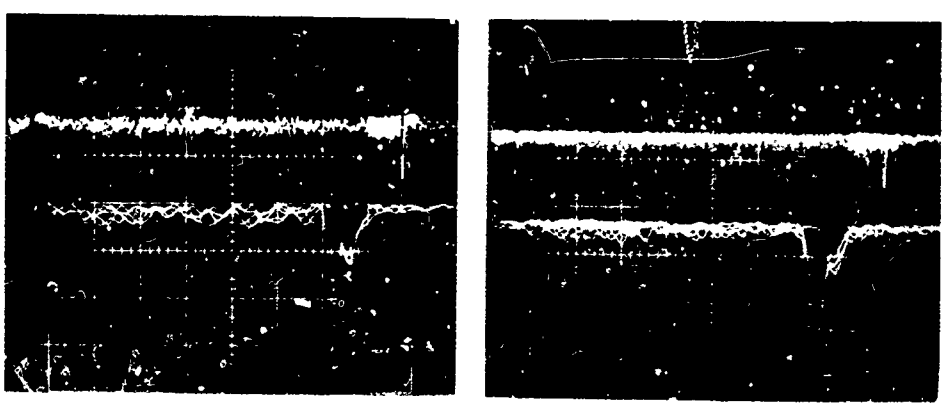


Fig. 2. Example of double-pulses received at Athens, in March 1975

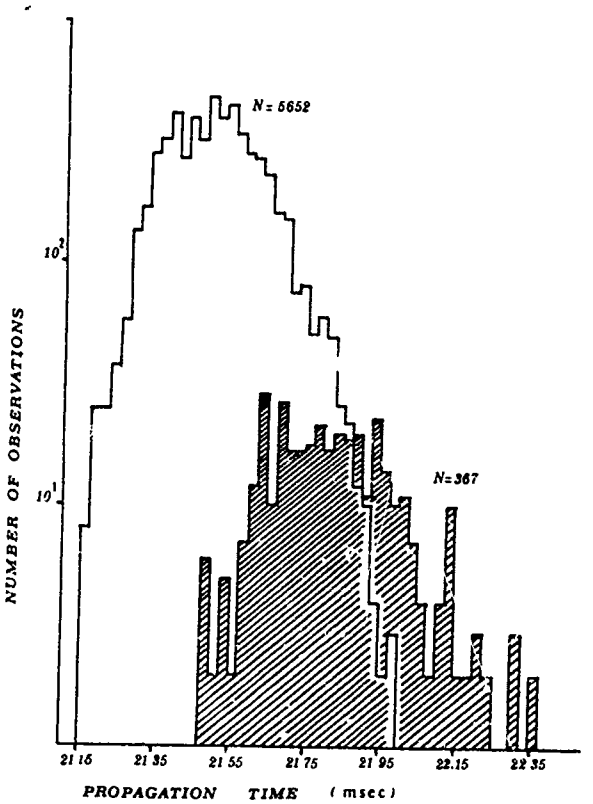


Fig. 3. Distributions of measured propagation times. The shaded area is refered to the second pulse.

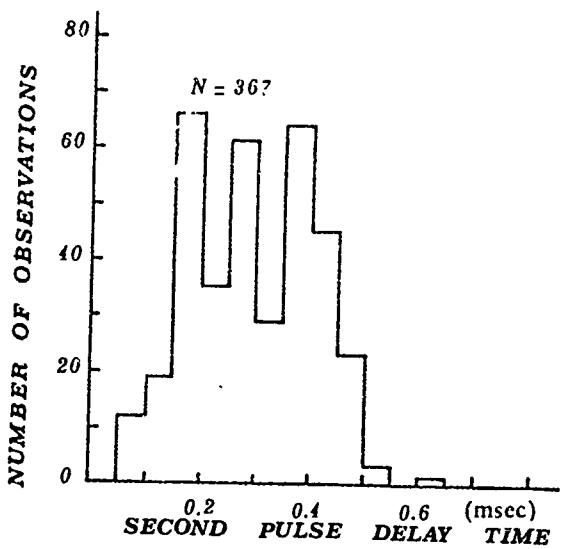


Fig. 4. Distribution of the second-pulse delay times.

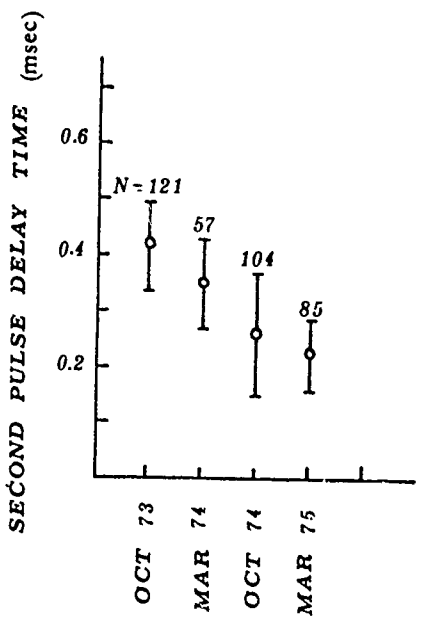


Fig. 5. Seasonal variation of second-pulse delay times.

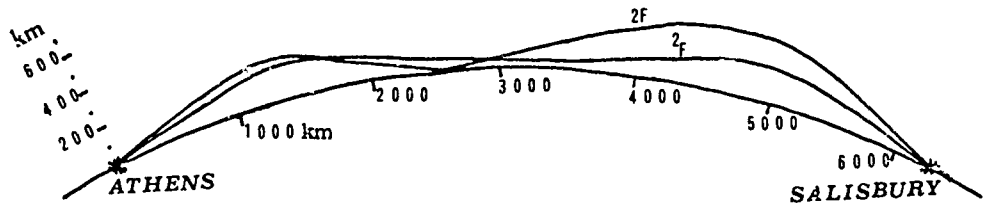


Fig. 6. Example of calculated $2F-^{2}F$ paths.

AN EMPIRICAL MODEL FOR THE PROBABILITY DISTRIBUTION OF THE LOWEST OBSERVED FREQUENCY

David B. Sailors
Naval Ocean Systems Center
San Diego, CA 92152

ABSTRACT

A model of the distribution representing the lowest observed frequency (LOF) is presented. The frequency spectrum of recorded minute-by-minute LOF data was bounded first on its lower end by radio noise and ionospheric absorption during the daytime and by lowest sounded frequency at night, and second at its upper end by the highest observed frequency path (MOF). The Johnson S_B system of frequency curves was fitted to LOF data using the method of moments (i.e., mean, standard deviation, etc.). The Johnson S_B system was chosen because (1) it represents all bounded distribution systems; (2) its simplicity of calculation makes it adaptable to minicomputer type applications; and (3) the transformation of the variables to the normal system allows use of normal probability algorithms. The four parameters of the distribution were determined as a function of path local time. The results for one path, France to Iceland, for October 1975 are presented. It was found for this data that the model was most accurate during the daytime, where the lowest observed LOF during an hour was above the lowest sounded frequency.

INTRODUCTION

This paper presents a model for determining whether an operational frequency is above the path lowest observed frequency (LOF). The method presented is the result of a study to develop a minicomputer algorithm for assigning frequencies at hf based on the probability that an operational frequency f_0 is between the two random variables: lowest usable frequency (LUF) and maximum usable frequency (MUF). The model was developed using digitally recorded, minute-by-minute, oblique incidence LOF data. The results presented are for one path, France to Iceland, for October 1975.

Many reasonably good computer models are designed to predict the long term properties of radio signals received via sky-wave modes of ionospheric propagation (Lucas and Haydon, 1966; Barghausen et al, 1969; Haydon et al, 1976). These programs calculate a parameter called circuit reliability. Circuit reliability is the likelihood that the signal level at the receiving terminals will sufficiently exceed the summation of the expected noise levels at these terminals so as to provide the type and quality of service desired.

For minicomputer applications it was assumed that a suitable approach would be to represent circuit reliability by

$$P\{\underline{MUF} > f_0\} \cdot P\{\underline{LUF} < f_0 \mid \underline{MUF} > f_0\} \qquad (1)$$

Where: 1) the probability $P\{\underline{MUF} > f_0\}$ gives the precentage of days within a month at a given hour that a sky-wave path will exist (probability of ionospheric support); and 2) the conditional probability $P\{\underline{LUF} < f_0 \mid \underline{MUF} > f_0\}$ gives the likelihood that, if the signal is reflected, the path LUF will be below the operational frequency f_0. The probability of ionospheric support can be evaluated using a technique developed by Zacharisen and Crow (1970) to fit the median and upper and lower deciles of the MUF distribution to a chi-square distribution function. A semi-empirical model, called MINIMUF-3, suitable for predicting the median MUF on a minicomputer is available (Rose et al, 1978; Rose and Martin, 1978). The upper and lower deciles of the MUF distribution can be obtained by multiplying the predicted median MUF by the factors given in Table 1 in CCIR Report 252-2 (CCIR, 1970). The chi-square distribution can be evaluated on a

minicomputer using a normal approximation (Peizer and Pratt, 1968). The predicted LUF can be calculated on a minicomputer using an algorithm called QLOF (Argo and Sailors, 1979).

The body of the paper describes how oblique incidence sounder data, recorded by the Naval Electronics Laboratory Center in the 1970's, was used to determine the nature of the conditional probability representing the path LUF in Eq. 1. Because the LOF data was bounded first on its lower end by radio noise and ionospheric absorption during the daytime and by lowest sounded frequency at night and second at its upper end by the maximum observed Frequency (MOF), a statistical model was sought that could represent bounded data. The Johnson S_B system was chosen because (1) it represents all bounded distribution systems; (2) its simplicity of calculation makes it adaptable to minicomputer applications; and (3) the transformation of the variables to the normal system allows use of normal probability algorithms. The parameters of the distribution were determined from the moments of the LOF data for each hour local time (i.e., mean, standard deviation, skewness, and kurtosis). It was found for the data sampled that the model was most accurate during the daytime, where the lowest observed LOF during an hour was above the lowest sounded frequency.

DIGITIZED NAVY TACTICAL SOUNDER DATA

In order to determine the nature of the LUF distribution, it was necessary to examine oblique incidence sounder data. The most suitable data for that purpose was that recorded by NELC at Navy communication stations in the 1970's. Hf digitizer/recorder (HFDR) systems were installed at communications stations in Greece, Honolulu, Iceland, and Japan. In addition, a reference system was installed at NELC's remote observatory at La Posta in Campo, CA. The paths being recorded were: Guam to Japan; Honolulu to Japan; Toulouse, France to Greece; Toulouse, France to Iceland; and Honolulu to La Posta.

The oblique sounder data collection system was comprised of the Navy Tactical Sounder System sounder transmitter, AN/FPT-11, and receiver, AN/UPR-2, and of the hf sounder digital data recording system (HFDR) Once each minute the FPT-11 transmitter sequentially transmitted a double, biphase, Barker-coded pulse on each of 80 discrete frequencies between 2 and 32 MHz; the total scan, consisting of 160 pulses, lasted 16 seconds. The 80 frequencies were spaced in 100 kHz increments in the 2 to 4 MHz range (Band A), 200 kHz increments from 4 to 8 MHz (Band B) and 400 kHz increments from 8 to 16 MHz (Band C) and 800 kHz increments from 16 to 32 MHz (Band D). The UPR-2 receiver sequentially processed the pulse-train input by starting the gated receiver scan at the same time as the transmission. This was accomplished by synchronizing to a common timing source (i.e., WWV) and maintaining an accurate time base generator in the receiver. The resultant pulses were then digitized and recorded on magnetic tape. When processed on digital computers, a minute-by-minute picture of mode of propagation, frequency and signal strength for the entire 2-32 MHz spectrum could be constructed.

Since the development of the first HFDR unit in 1968, concurrent efforts were conducted in developing data processing techniques to facilitate hf propagation analysis. This included a technique for the determination of the MOF and LOF for each sounder scan. These minute-by-minute processed MOF/LOF data were available on magnetic tape for several paths and months. Data for the France to Iceland path recorded during October 1975 was used to develop the model presented here.

DATA PREPARATION

The first step in the development of the model was to improve the quality of the sounder data. This was accomplished by both smoothing of the data and the elimination of outliers.

MOF/LOF Smoothing

To eliminate some of the high-frequency oscillations occurring on the minute-by-minute variations of MOF/LOF, the estimated MOF and LOF values at each scan were digitally filtered by a three-term linear formula of the form

$$g_k = {}^1/_4 \, (f_{k-1} + 2f_k + f_{k+1}) \qquad (2)$$

Elimination of Outliers

Figure 1 is a histogram of the LOF distribution for the France to Iceland path for the 10:00 hour of October 1975. One might suspect that outlying data points beyond 16.75 MHz are spurious observations. Such a spurious observation might arise from the inability to process out recorded noise pulses or a particular day might not be representative of the month as a whole (i.e., there might have been a solar disturbance). A procedure is then needed to eliminate such data points when they arise.

An approach due to Anscombe (1960) somewhat modified by Tiao and Gutmann (1967) was adopted for rejecting the suspected spurious observations. They discuss the situation of sampling from a normal population $N(\mu,\sigma)$ with variance σ^2 known. This leads to a set of estimation procedures by which the observations associated with the largest adjusted residuals are excluded. They then construct the adjusted residuals

$$z_i = y_i - \bar{y} + \frac{v}{\sqrt{n}} , \quad i = 1,\ldots,n \tag{3}$$

Where $y_1,\ldots, y_n$ are the observations, $\bar{y}$ is the mean observation, and v is an independent observation from $N(0,\sigma^2)$. For example, v might be obtained from a table of random normal deviates.

In the case of the LOF/MOF sounder data, the estimate of σ was found by estimating the standard deviation for the month at each hour. Then the adjusted residuals for each day of the month at each hour were determined. Figure 2 is an example of the results of the application of these procedures to the same data as in figure 1. Note a reduction in the frequency range of this distribution of 6 MHz. The application of this procedure to other hours consistently improved the quality of histograms (no spurious data points at the tails).

DETERMINATION OF A STATISTICAL MODEL

The second step was to hypothesize a distribution function and determine its parameters. In the case of the application at hand, the determination of the distribution function was motivated by (1) a desire for ease of calculation so that the results could be used in minicomputer type applications, and (2) a need to automate the data analysis. In addition, it was expected that the measured LOF values would be bound on the upper end by the MOF and on the low end by radio noise and ionospheric absorption. In particular, the distribution function chosen for the LOF data was one known as the Johnson curves (Johnson, 1949).

Johnson Curves

The Johnson curves are an empirical family of curves chosen because (1) they are easy to evaluate once their parameters are determined; (2) their parameters can be determined either by using moments or percentile points; and (3) the system of curves can be used to represent all existing distribution functions. Johnson proposed basing empirical distributions on the transformation of a standard normal variate. The Johnson system of frequency curves consist of:

the lognormal system (or S_L): $z = = \gamma + \delta \ln (x-\xi) \quad \xi < x,$ (4)

the unbounded system (or S_u): $z = \gamma + \delta \sinh^{-1} [(x-\xi)/\lambda],$ (5)

the bounded system (or S_B): $z = \gamma + \delta \ln [(x-\xi)/(\xi+\lambda-x)] \quad \xi < x < \xi + \lambda$ (6)

where z is the standardized normal variate in each case. The parameters γ and δ determine the shape of the distribution of x; λ is a scale factor; and ξ is a location factor.

To decide which of the three Johnson families should be used for a given set of data, the usual procedure is to obtain the data estimates of the skewness $\sqrt{\beta_1}$, and the kurtosis, β_2. These then are plotted on a figure such as figure 3 which shows the region in the (β_1,β_2) plane for the three Johnson families. Also shown are other common sampling distributions. If the (β_1,β_2) point is close to the S_L curve, S_L family is chosen. If it is in the region above the S_L curve, the S_B family is chosen; and if it is below the curve, the S_u family is used. The S_L curve can be extended by use of the parametric equations

$$\beta_1 = (\omega - 1)(\omega + 2)^2 \tag{7}$$

$$\beta_2 = \omega^4 + 2\omega^3 + 3\omega^2 - 3. \tag{8}$$

The impossible region in the figure is bounded by the line $\beta_2 - \beta_1 - 1 = 0$.

Fitting Johnson Curves by Moments

To determine the parameters for the Johnson curves for the LOF data an algorithm, called JNSN, due to Hill et al (1976) was used. This altorithm uses the sample moments (i.e., the mean, standard deviation, skewness ($\sqrt{\beta_1}$), and the kurtosis (β_2)) to determine the type of Johnson curve and its parameters, Eqs. 7 and 8 are solved in JNSN to determine the type of Johnson distribution.

To determine the parameters for the Johnson S_B system, the expected LOF distribution, JNSN approximates initial values for δ and γ. These initial values are determined from β_1 and β_2. Evaluation of the first six moments at the given δ and γ values, using Draper's (1952) form of Goodwin's (1949) integral, then enables a two-dimensional Newton-Raphson process to converge on δ and γ values which yield the required input β_1 and β_2 values.

For the Johnson S_B system r^{th} moment of $y = (x-\xi)/\lambda$ is

$$\mu_r'(y) = \frac{1}{\sqrt{2\pi}} \int_{-\infty}^{\infty} e^{-z^2/2} \left(1 + e^{-(z-\gamma)/\delta}\right)^{-r} dz. \tag{9}$$

Applying Goodwin's quadrature formula for evaluation of integrals of the form $\int_{-\infty}^{\infty} f(x)e^{-x^2}\,dx$, results in

$$\mu_r'(y) = \frac{h}{\sqrt{\pi}} \sum_{n=-\infty}^{\infty} \left(1 + e^{(\gamma - \sqrt{2}\,nh)/\delta}\right)^{-r} e^{-n^2h^2} - E_r(h) \tag{10}$$

The error term $E_r(h)$ tends rapidly to zero with the decrease in interval size h. The initial value of h is 0.75. Beginning with n = 0 term, successive terms are added in Eq. 10, until $[\mu_r'(n) - \mu_r'(n-1)]/\mu_r'(n-1)$ is less than 1.0×10^{-5}. JNSN allows up to 500 terms in the series evaluation. The derivatives required by the Newton-Raphson process are obtained from the relationship between higher order moments of the Johnson S_B system

$$\mu_{r+1}' = \mu_r' + \frac{\delta}{r} \frac{\partial \mu_r'}{\partial \gamma} \tag{11}$$

When the required δ and γ have been found, ξ and λ can be found from

$$\sigma(x) = \lambda\left(\sigma(y) - \mu_1'(y)^2\right)^{1/2} \tag{12}$$

$$\mu_1'(x) = \xi + \lambda\, \mu_1'(y) \tag{13}$$

$\sigma(y)$ and $\mu_1'(y)$ are available at each stage of the evaluation for δ and γ, and $\sigma(x)$ and $\mu_1(x)$ are the input standard deviation and mean, respectively.

RESULTS

A detailed statistical analysis was conducted of the minute-by-minute smoothed LOF/MOF data from which spurious observations had been removed as previously described.

Propagation Properties France-to-Iceland Path

The propagation properties for the France to Iceland Path for October 1975 are presented in Table 1 and figure 4. At each hour Table 1 gives the predicted LUF, the four sample moments of the LOF data, the lower decile of the MOF data, the predicted FOT (the predicted lower decile of the MUF distribution) and the zenith angles at the transmitter and receiver locations. The table shows that during the nighttime that the median LOF is at the sounder minimum frequency. During the nighttime, the larger values of β_1 indicate that at night the distributions are seriously skewed. Also at night the high values of β_2 indicate a distribution with long tails and high peaks. Figure 4 shows select percentile points and the minimum and the maximum values for the LOF data. Included are only those values greater than the sounder minimum. The minimum value is above the sounder minimum only during the hours 8 to 16 UT. Even the 25% percentile point

occurs above the minimum only during the hours 6 to 18 UT. The first percentile curve illustrated entirely above the minimum is the 75% percentile point.

Histograms were plotted of the LOF distribution from 0 hours UT through 23 UT. The first five distributions for the night hours 0 UT through 4 UT were highly skewed with a single mode at the sounder minimum. Figure 5 is a sample for the hour 0 UT. In Figure 6 for the hour 5 UT, a second mode can be seen to begin at about 6.25 MHz. For the hours 6UT through 10 UT, this mode became more pronounced. This secondary mode disappeared at 1100 UT only to reappear at 1200 UT. In this instance the secondary mode is very strong. By 1400 UT it disappeared. The daytime main mode is characterized by a somewhat skewed distribution. By 1800 UT the highly skewed nightime distribution returned.

A partial explan tion for the appearance of the secondary mode on the morning histograms can be obtained by examining Table 1. In addition to QLOF predicted LUF and MINIMUF predicted FOT, Table 1 gives the sun's zenith angle at the transmitter and the receiver. As shown in Table 1 the sun's zenith angle at the transmitter is decreasing with increasing hour. During the 5th hour it reaches a value less than 108.9° and greater than 98.3°. If the zenith angle for the starting or stopping of daytime absorption is 102.2° (Laitinen and Haydon, 1962), during this hour the path is partially exposed to sunlight, causing on some of the days higher absorption. By 9 UT the path is entirely in daylight, and the effect can be seen to begin to decrease. But at 12 UT it reappears and is at its strongest. In this case, the increased LOF is due possibly to: (1) automobile emission noise on a highway near the receiver site or (2) increased noontime automobile emission noise at the communication station where the receiver was located.

Table 2 presents the Johnson curve parameters for the France to Iceland path for October 1975. It was found, as was expected, that at every hour the LOF data followed a Johnson S_B curve. The correlation coefficient between the mean LOF in Table 1 and ξ in Table 2 is 0.958. The correlation coefficient between the MOF lower deciles values in Table 1 and the parameters $(\xi + \lambda)$ in Table 2 is not quite as high but is still 0.906.

Verification of the Model

The final step in the determination of a distribution representing the LOF data was the verification that the fit was adequate. The procedure followed for the test of fit was that for the chi-square test of fit (Williams, 1950; Hahn and Shapiro, 1967). The number of classes k is selected by means of a formula depending upon the sample size N and the level of significance c, and the class limits are chosen such that each class contains the same number of items under the null hypothesis (i.e., a Johnson distribution). For this distribution, the class boundaries were found by inputting 1/k, 2/k, ..., (k-1)/k into an algorithm for finding the normal deviates corresponding to the lower tail area (Beasley and Springer, 1970). Then an algorithm was used to find the Johnson deviates corresponding to the normal deviates (Hill, 1976).

When the chi-square test-of-fit was was applied for a 5% level of significance to the LOF data, it failed to provide a useful conclusion. This occurred for two reasons. The first is that even though the sounder data was smoothed, it tended to cluster about the discrete sounder frequencies. The chi-square test depends on the data being ungrouped. Second, at night the bulk of the data is in the cell at one endpoint causing the technique to fail.

Even so, the procedure provided cumulative distributions to graphically compare to the assumed null hypothesis. Figs. 7 through 11 compare the assumed Johnson distribution against the observed data for the hours 0, 5,7, 10 and 16 UT, respectively. Examination of Fig. 7 for 0 UT shows an error in probability of more than .1 below 3 MHz. As 75% of the data is below 3 MHz, the fit is obviously very poor. Even in Fig. 8 for 5 UT, the improvement is not too encouraging. As the morning hours arrive in Fig. 9 for 7 UT the fit begins to improve. Figs. 10 and 11 for daytime conditions show quite good fits.

DISCUSSION

In this paper a model for determining the probability whether an operational frequency is above the path lowest observed frequency (LOF) was presented. The Johnson S_B system of frequency curves was fitted to minute-by-minute oblique inci-

dence LOF data using the method of moments (i.e., mean, standard deviation, etc.) It was found for the France-to-Iceland, October 1975, data that the model was most accurate during the daytime, where the lowest observed LOF was well above the lowest sounder frequency. The poor results at night occur in part due to the large number of measurements at sounder minimum. Because the real data points are at or below the minimum, the sample estimates of the moments used to determine the Johnson S_B parameters are not very accurate. The method of moments itself may not be accurate for the large values of skewness and kurtosis existing in the LOF data at night.

A better approach would be to separate out the data points at sounder minimum and find the Johnson S_B parameters for the remaining data points using either maximum likelihood estimates (MLE) or the percentile points assuming ξ equal the sounder minimum. Then the probability of a frequency being above the LOF would be $100 \cdot [P_0 + P_1(1 - P_0)]$ where P_0 is the ratio of LOF data points at sounder minimum to the total data points in the sample and P_1 is the probability given by the model determined from the points above the sounder minimum.

There is still the need to extend this model to other regions and seasons. The correlation between ξ and the mean LOF and between $(\xi + \lambda)$ and the lower decile of the MOFs might imply that predicted LUFs and FOTs might be used for this purpose within the latitude region for which the parameters are obtained.

Finally, once fully developed, the ability to predict the reliability of a frequency between the MUF and LUF on a minicomputer will present numerous applications with growing minicomputer technology.

ACKNOWLEDGEMENTS

The author is deeply indebted to the numerous individuals at the Naval Electronics Laboratory Center in the early 1970's who developed the oblique sounder digitizing system, and who developed the software to process the magnetic tapes to get the minute-by-minute LOF/MOF data used in the analysis.

REFERENCES

Anscombe, F. J., (1960), "Rejection of outliers," Technometrics, 2(5), 123-147.

Argo, P. E. and D. B. Sailors (1979), "Quiet time lowest observable frequency (QLOF) calculation program," Tech. Doc. 231, Naval Ocean Systems Center.

Barghausen, A. F., J. W. Finney, L. L. Proctor, and L. D. Schaultz (1969), "Predicting long-term operational parameters of high frequency sky-wave systems," Tech. Rep. ERL 110-ITS 78, Environmental Sciences Services Administration.

Beasley, J. D. and S. G. Springer (1977), " Algorithm AS 111. The percentage points of the normal distribution," Appl. Statist., 26, 118-121.

Consultative Committee International Radio, New Delni (1970, Rept. 252-2; CCIR interim method for estimating sky-wave field strength and transmission loss at frequencies between the approximate limits of 2 and 30 MHz, pp 84-88, International Telecommunications Union, Geneva.

Draper, J. (1952), "Properties of distributions resulting from certain simple transformations of the normal distribution," Biometrika. 39, 290-301.

Goodwin, E. T. (1949), "The evaluation of integrals of the form $\int_{-\infty}^{\infty} f(x)\, e^{-x^2} dx$," Cambridge Phil. Soc., 45, 241-245.

Hahn, G. J. and S. S. Shapiro, (1967), Statistical Models in Engineering, pp. 302-308, John Wiley and Sons, New York.

Haydon, G. W., M. Leftin, and R. Rosich (1976), "Predicting the performance of high frequency sky-wave telecommunications systems (the use of the HFMUFES 4 program)," Rep. 76-102, Office of Telecommunications.

Hill, I. D. (1976), "Algorithm AS 100. Normal-Johnson and Johnson - normal transformations," Appl. Statist., 25, pp 190-192.

Hill, I. D., R. Hill, and R. L. Holder (1976), "Algorithm AS 99. Fitting Johnson curves by moments," Appl Statist., 25, 180-189.

Johnson, N. L. (1949) "Systems of frequency curves generated by methods of translation," Biometrika, 36, 149-176.

Laitinen, P. D. and G. W. Haydon (1962), "Analysis and prediction of sky-wave field intensities in the high frequency band," Tech. Rept. 9, Army Signal Radio Propagation.

Lucas, D. L. and G. W. Haydon (1966), "Predicting statistical performance indexes for high frequency ionospheric telecommunication systems," Tech. Rep. IER-ITSA 1. Environmental Sciences Services Adinistration (AD644827).

Peizer, D. B. and J. W. Pratt (1968), "A normal approximation for the binnomial, F, beta and other common, related tail probabilities, I.," J. Amer. Stat. Assn., 63, 1416-1456.
Rose, R. B. and J. N. Martin (1978), "MINIMUF-3.5," Tech. Doc. 201, Naval Ocean Systems Center.
Rose, R. B., J. N. Martin, and P. H. Levine (1978), "MINIMUF-3: a simplified hf MUF prediction algorithm," Tech. Rep. 186, Naval Ocean Systems Center.
Tiao, G. L. and I. Guttman (1967), "Analysis of outliers with adjusted residuals," Technometrics, 9(11), 541-559.
Williams, Jr., C. A. (1950), "On the choice of the number and width of classes for the chi-square test of goodness of fit," Jour. Am. Sta. Assn., 45, 77-86.
Zacharisen, D. H. and E. L. Crow (1970), "Fitting distributions of telecommunication variables with chi-square distributions," Radio Science, 5(11), 1307-1315.

Table 1 Propagation properties for the France to Iceland path for October 1975

Hour UT	LUF	LOF Mean	LOF c	LOF β_1	LOF β_2	MOF (10%)	FOT	Zenith Angle Transmitter	Zenith Angle Receiver
0	2.00	2.91	1.60	4.80	7.33	8.29	5.48	154.8	132.6
1	2.00	2.88	1.44	3.68	5.78	8.39	5.33	149.6	134.3
2	2.00	2.70	1.12	5.39	8.29	8.29	5.25	141.0	133.8
3	2.00	2.92	1.40	4.05	6.42	9.49	5.52	130.7	131.0
4	2.00	2.74	1.32	5.04	7.24	8.81	5.44	119.9	126.6
5	2.00	2.64	1.29	7.43	9.51	8.13	5.36	108.9	121.0
6	2.00	2.93	1.26	2 69	4.67	7.94	9.12	98.3	114.6
7	2.00	5.32	3.49	1.59	3.40	10.49	11.34	88.1	108.1
8	2.60	5.94	3.21	1.79	3.84	13.89	12.47	78.9	101.6
9	5.61	6.61	2.69	1.49	3.83	13.49	13.28	71.1	95.6
10	7.16	7.17	1.90	1.62	4.26	14.99	13.77	65.3	90.5
11	8.30	7.79	2.10	1.14	3.30	15.77	13.22	60.0	86.3
12	8.34	8.26	2.41	1.26	3.11	15.89	13.38	61.6	83.6
13	8.21	8.23	2.55	1.92	4.61	15.89	13.30	64.3	82.4
14	7.60	7.20	1.94	1.59	4.33	15.89	13.15	69.7	82.8
15	6.40	6.29	2.05	1.97	5.22	15.89	13.61	77.2	84.8
16	4.21	5.01	1.98	3.00	6.10	15.29	13.04	86.1	88.3
17	2.00	3.75	1.88	3.90	6.82	12.79	12.15	96.1	93.0
18	2.00	2.92	1.54	9.49	14.07	13.89	10.85	106.7	98.6
19	2.00	2.79	1.72	11.22	14.57	12.69	7.37	117.6	104.9
20	2.00	3.03	1.94	7.56	10.78	10.69	6.54	128.5	111.5
21	2.00	2.70	1.49	8.78	11.76	9.49	6.23	139.0	118.0
22	2.00	3.10	1.75	3.77	6.37	9.09	5.93	148.1	124.1
23	2.00	2.90	1.52	4.56	7.10	8.29	5.63	154.3	129.2

Table 2. Johnson S_B Distribution Parameters for LOF Data for the France to Iceland Path October 1975

Hour UT	γ	δ	λ	ξ	ξ + λ
0	1.421	0.429	7.894	1.894	9.788
1	1.229	0.374	6.058	1.959	8.017
2	1.553	0.479	6.179	1.976	8.155
3	1.340	0.431	6.543	1.983	8.526
4	1.332	0.335	5.843	2.003	7.846
5	1.411	0.393	5.765	2.080	7.845
6	0.906	0.411	5.014	2.019	7.033
7	0.906	0.411	12.626	2.409	15.035
8	1.042	0.507	13.214	3.077	16.291
9	1.167	0.693	13.087	3.658	16.745
10	1.399	0.828	10.912	4.895	15.807
11	0.998	0.659	9.348	5.420	14.768
12	0.860	0.463	8.886	6.030	14.916
13	1.444	0.789	14.745	5.361	16.738
14	1.497	0.898	11.995	4.743	20.240
15	2.015	1.081	16.743	3.497	15.920
16	1.690	0.762	12.873	3.047	12.999
17	1.555	0.592	10.742	2.256	14.068
18	1.986	0.539	12.056	2.012	13.441
19	1.719	0.339	10.363	2.058	14.421
20	1.660	0.432	11.487	1.954	13.441
21	1.619	0.352	8.342	1.991	10.333
22	1.416	0.515	8.922	1.792	10.714
23	1.428	0.455	7.595	1.893	9.487

CUM FREQ	FREQ	NUMBER
0.1	3.75 / **	1
1.5	4.25 / ***	16
3.0	4.75 / ***	17
13.6	5.25 / **********	118
27.8	5.75 / **********	160
43.4	6.25 / **********	175
56.6	6.75 / **********	148
67 5	7.25 / **********	122
77.6	7.75 / **********	113
80.8	8 25 / ******	36
83 5	8.75 / *****	30
85 2	9.25 / ****	19
89.	9.75 / ********	47
90 4	10.25 / **	11
91.8	10.75 / ***	16
92.2	11.25 / *	4
94.7	11.75 / *****	29
96 8	12.25 / ****	23
98.0	12 75 / ***	14
98.5	13.25 / *	5
98.7	13.75 / *	2
98.8	14.25 / *	2
99.0	14.75 / *	2
99.5	15.25 / *	5
99.6	15.75 / *	2
99 6	16.25 /	0
99.6	16.75 /	0
99.8	17.25 / *	2
9	5 /	0
9	18 /	0
9	8,75 /	0
9	19.25 / *	1
100.0	19.75 / *	1

Figure 1 LOF distribution for France to Iceland for the 10.00 UT hour during October 1975 (each * represents 6 values in the frequency band).

CUM FREQ	FREQ	NUMBER
0.1	3.75 / *	1
1.5	4.25 / ***	16
3.1	4.75 / ***	17
13.8	5.25 / **********	118
28.3	5.75 / **********	160
44.2	6.25 / **********	175
57.7	6.75 / **********	148
68 8	7.25 / **********	122
79.0	7.75 / **********	113
82.3	8.25 / ******	36
85.0	8.75 / *****	30
86.7	9.25 / ****	19
91.0	9.75 / ********	47
92.0	10.25 / **	11
93.5	10.75 / ***	16
93.8	11.25 / *	4
96.5	11.75 / *****	29
98.5	12.25 / ****	23
99.7	12.75 / ***	13
99.9	13.25 / *	2
100.0	13.75 / *	1

Figure 2. LOF distribution for France to Iceland for the 10·00 UT hour during October 1975. Elimination of outlier procedure applied. (Each * represents 6 values in the frequency band.)

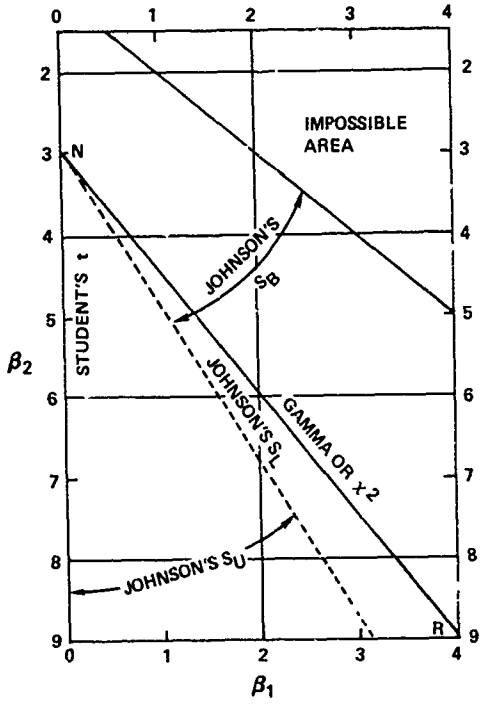


Figure 3. Chart for determining Johnson distribution approximation, N: normal point (0, 3), R: Raileigh power distribution (4, 9).

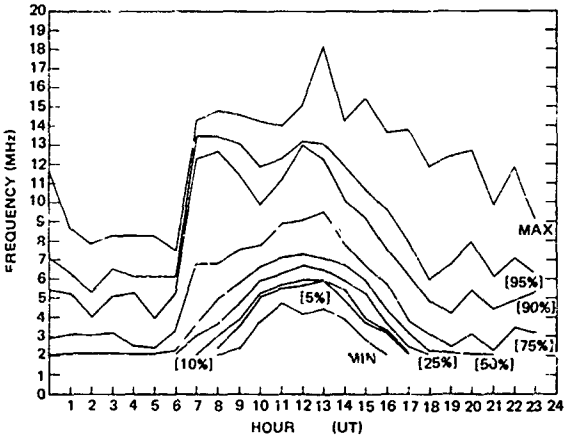


Figure 4. Select percentile points and minimum and maximum values of the LOF data for the France to Iceland path for October 1975.

CUM FREQ	FREQ	NUMBER
73.5	2.25 /**********	919
75.1	2.75 /****	21
80.0	3.25 /**********	61
82.7	3.75 /******	33
85.7	4.25 /********	38
88.4	4.75 /*******	34
90.0	5.25 /****	20
91.7	5.75 /****	21
93.8	6.25 /******	27
94.8	6.75 /**	12
97.4	7.25 /*******	33
98.6	7.75 /***	16
99.4	8.25 /**	9
99.4	8.75 /*	1
99.7	9.25 /*	3
99.7	9.75 /	0
99.8	10.25 /*	1
99.8	10.75 /*	1
99.9	11.25 /*	1
100.0	11.75 /*	1

Figure 5 LOF distribution for France to Iceland for the 0 UT hour during October 1975. (Each * represents 6 values in the frequency band.)

CUM FREQ	FREQ	NUMBER
76.2	2.25 /**********	650
85.1	2.75 /**********	76
88.7	3.25 /******	31
90.3	3.75 /***	12
90.9	4.25 /*	5
91.2	4.75 /*	3
92.4	5.25 /**	10
94.0	5.75 /***	14
95.9	6.25 /***	16
97.5	6.75 /***	14
98.4	7.25 /**	7
99.4	7.75 /**	9
100.0	8.25 /*	5

Figure 6 LOF distribution for France to Iceland for the hour 5 UT during October 1975. (Each * represents 6 values in the frequency band.)

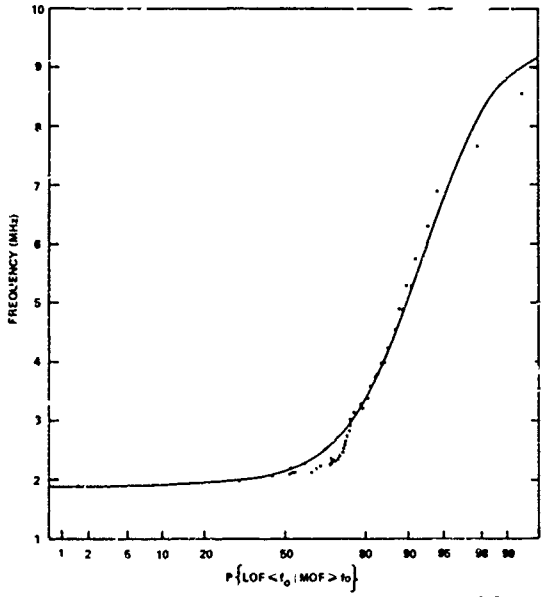


Figure 7. Comparison of measured and theoretical LOF distribution for France to Iceland path October 1975 at 0 UT.

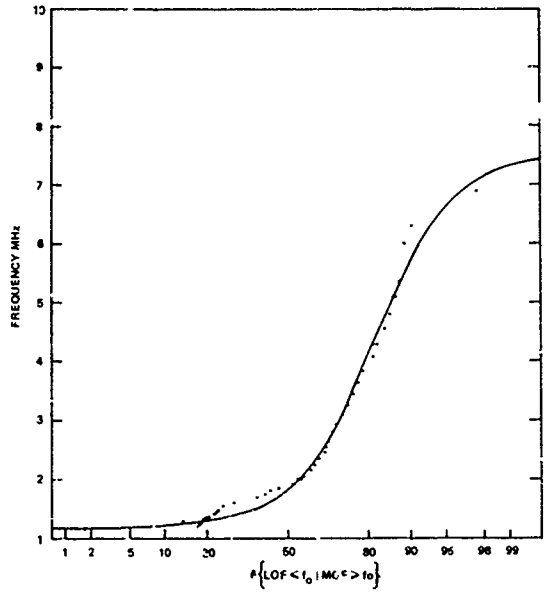


Figure 9. Comparison of measured and theoretical LOF distribution for France to Iceland paths for October 1975 at 07 UT.

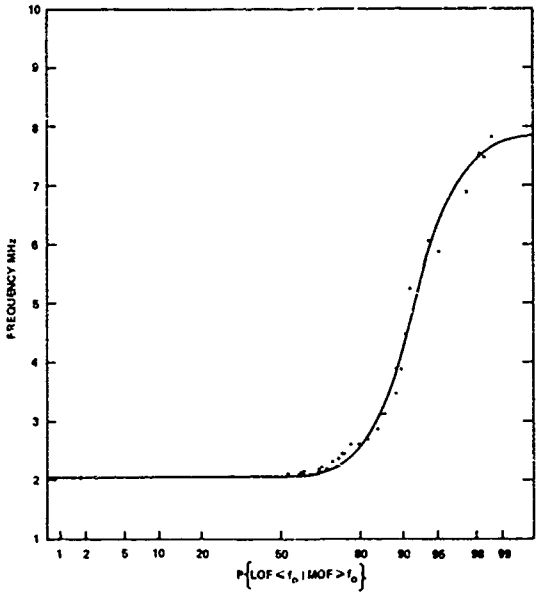


Figure 8. Comparison of measured and theoretical LOF distribution for France to Iceland path for October 1975 at 05 UT.

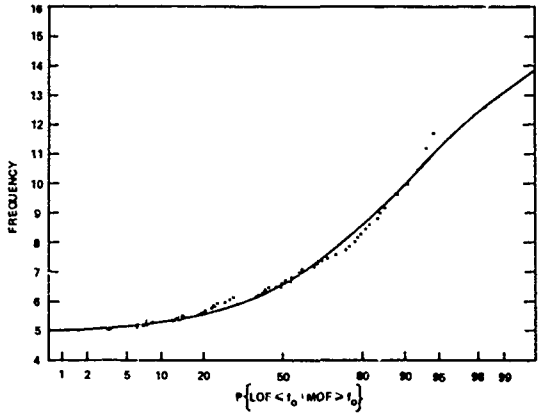


Figure 10. Comparison of measured and theoretical LOF distribution for France to Iceland path for October 1975 at 10 UT.

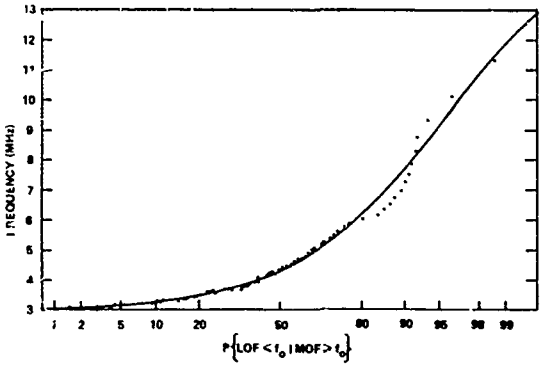


Figure 11. Comparison of measured and theoretical LOF distribution for France to Iceland path for October 1975 at 16 UT.

ACCOUNTING FOR IONOSPHERIC VARIABILITY AND IRREGULARITY IN HIGH FREQUENCY DIRECTION FINDING

Melvin G. Heaps

USA Atmospheric Sciences Laboratory
Electronics Research and Development Command
White Sands Missile Range, New Mexico 88002

I. INTRODUCTION

High frequency (HF) radio wave propagation has long been the backbone of many short and long range communication networks. The reasons are straightforward: the technology is well established, the systems are relatively inexpensive and easily fieldable, and the reliability (while not perfect) is high enough to meet a large body of communication needs. In a similar vein, the interception of HF communications is one means of gaining additional intelligence in a military engagement. Taken an additional step, the interception of HF communications can also be used to locate the position of the transmitter, which knowledge may be of strategic or tactical value. HF radio waves propagate in two modes: the ground wave, which can be detected out to distances on the order of 50 km, and the direct wave, which when reflected from the ionosphere, is known as the skywave and can be detected out to very long distances. The radio source location technique for a single station receiver consists of measuring the angle of arrival of the signal of interest in three dimensions. The problem then consists of tracing the signal's path back through the ionosphere to the source, the successful solution of which depends on a knowledge of the state of the ionosphere.

Currently attainable accuracies in HF position location using a single station locator still carry inherent errors of tens of kilometers or worse. These errors arise from three main areas: ionospheric variability and irregularity, locator system size limitations, and problems with data acquisition, processing and interpretation. Of these areas the ionosphere is the single largest source of error and is the principle concern of this paper.

It is now generally agreed that real time information on the state of the ionosphere is required for optimum performance of a single station locator. Most often this is accomplished via a vertical ionospheric sounding made at that station. It then becomes important to quantify the spatial and temporal irregularities of the ionosphere and estimate the spatial and temporal ranges within which ionospheric sounding information gathered at one point can be extrapolated to another point with minimal loss of position location accuracy.

The next section will then outline the general properties of ionospheric irregularities and the order of magnitude of the errors which are introduced into position location accuracies. The spatial and temporal coherence of the ionospheric irregularities will then be estimated. The final section will address the usefulness of a single ionospheric sounding as opposed to multiple and spatially separated soundings.

II. IONOSPHERIC IRREGULARITIES MOST IMPORTANT FOR RADIO SOURCE LOCATION

In principle the location of an HF transmitter can be found by simply measuring the azimuthal and elevation angles of the incoming signal and determining the height of the ionospheric reflecting layer. In practice there are several types of ionospheric irregularities which distort the otherwise straight forward picture of a uniform, concentric, smoothly reflecting ionosphere.

Irregularities can arise from a multitude of causes ranging from plasma instabilities and non-uniform ionization sources, through atmospheric winds and traveling waves, to mass motions of the atmosphere due to tidal and heating effects (see, for example, Kent, 1970; Yeh and Liu, 1974; Fejer and Kelley, 1980; and the references therein). Three phenomena shall be singled out because of their effects on radio source location: sporadic E, ionospheric tilts, and traveling ionospheric disturbances (TID's). Of the above, TID's are the more important.

Sporadic E is a thin layer of ennanced ionization confined to the E region (principally 100-110 km) and is felt to be a result of an enhanced concentration of ions (principally metallic ions) caused by wind shears in that region. These wind shears may in turn be a result of the propagation of acoustic-gravity waves through the E region, and thus sporadic E may be an E region manifestation of one type of traveling waves which also produce F region irregularities in the form of one class of TID's. The properties of sporadic E are summarized in Table 1 for the midlatitudes (after Smith, 1957; Peterson, 1980). Sporadic E-like occurrances are also found at low and high latitudes and are associated with the equatorial and polar electrojets; these are not considered further here.

The patchiness of sporadic E is somewhat a function of the radio frequency employed. Higher density patches of a few hundred kilometers extent or less, as discerned by radio frequencies above 7 MHz, may be found embedded in larger, lower density patches as defined by radio frequencies below 3 MHz (Peterson, 1980). Both scales of patches apparently move with similar velocities which suggest similar orginating mechanisms.

Sporadic E, because of the thiness of the layers and the sharpness of the vertical gradients, is often an aid in communications and often provides a better reflecting surface than the F region. The principal deleterious effect of sporadic E for radio source location is the creation of multi-mode interference. Signals from the source may suffer reflections from both E and F layers, thus causing additional problems with resolution of modes.

Ionospheric tilts refer to generally large scale horizontal gradients in the electron density, such that contours of constant electron density are no longer parallel to the earth's surface. The effects are to cause an error in estimating the angle of arrival, principally the elevation angle, and to misjudge the virtual height of the reflection point. The most regular and predictable ionospheric tilt is that caused by the diurnally varying solar ionization rate in the F region. The effect is most noticeable at sunrise and sunset, and can thus be anticipated and accounted for. In a generic sense, ionospheric tilts can refer to any deviation from the horizontal plane of the contours of constant electron density, whether caused by large-scale phenomena such as solar ionization mentioned above, or due to more transient, localized disturbances discussed next.

Traveling ionospheric disturbances have been noticed since the earliest days of radio wave propagation and were first studied extensively by Munro (1950, 1958). TID' are essentially an ionospheric manifestation of an entire spectrum (not necessarily continuous) of waves propagating through the atmosphere. The spectrum of TID's can be placed in at least two distinct categories: large-scale and medium-scale. According to acoustic-gravity wave theory, large-scale TID's are associated with a discrete spectrum of guided waves whose modes are excited only by upper atmospheric sources and whose horizontal speeds are substantially greater than the (lower atmospheric) speed of sound. Medium-scale TID's are associated with a spectrum of freely propagating internal waves which can be excited by sources at any altitude and whose horizontal speeds are less than the speed of sound. As one might suspect, medium-scale TID's are much more common. (For more information on waves in the atmosphere see Georges, 1967; Yeh and Liu, 1974, Hines, et. al., 1974; and the references therein.) A third category of small-scale TID's exists which is most likely the extension to higher frequencies and smaller size of the medium-scale TID's is generally below the Fresnal-zone size of ionospheric sounders and thus has not been as well documented. Table II summarizes the properties of these different categories (after Georges, 1967; Rao, 1981).

Again the main effects of TID's are to cause errors in the angle of arrival, measured as the azimuthal and elevation angles, and the virtual height of reflection. Table III gives the magnitude of the errors in position location which the ionosphere can cause for selected ranges. A quick "rule of thumb" seems to be 10 km or 10% of range, whichever is worse.

III. SPATIAL AND TEMPORAL COHERENCE OF IONOSPHERIC IRREGULARITIES

The basic question which needs to be answered can be stated as follows, "If the state of the ionosphere can be determined at one point, over what spatial ranges can that information be transferred, and for what time period is it valid?" For the simple case of a single vertically incident ionosonde, the pertinent information would be the height of the reflecting layer and the tilt of the ionosphere. The problems encountered are shown in Figures 1 and 2.

Figure 1 is a plot of the incident angle (plotted as radial distance from the origin) versus the azimuthal angle of arrival (plotted as polar angle) for the return signal of a vertically incident ionosonde. The numbers represent one sounding each minute from 11:49 to 12:39 local standard time (Ernst, et. al., 1974; Rao, 1981). The general pattern of a NW-SE propagation wave is apparent, but so are the patterns of other smaller and differently oriented waves. This is often typical for medium-scale TID's, which are superpositions of several frequency components. Figure 2 shows the constant plasma frequency contours (i.e., variation of reflection heights) as a function of time

TABLE I

IONOSPHERIC IRREGULARITIES

TYPE OF IRREGULARITY	SPORADIC E	IONOSPHERIC TILT
Structure	Patches of enhanced electron and ion density often hundreds of kilometers in horizontal extent.	Horizontal gradient in electron density over distances on the order of 1000 km.
	Vertical thickness generally 1-2 km at an altitude of 100-110 km.	Found at F region altitudes (>140 km).
	Patches of higher density often embedded in larger, lower density patches.	
Origin	Wind shears, probably from propagating acoustic-gravity waves.	Daily variation of solar ionizing radiation.
Motion	Patches (not plasma) generally move 50-100 m/s, no preferred direction.	Pattern shifts with sun's diurnal motion.
Duration	Several minutes to several hours.	Over the order of 1-2 hours.
Occurrence	More frequent during day, with peak occurrence before noon, and in some locations a secondary peak near sunset likely in summer.	Daily near sunrise and sunset.
	More frequent off the east coast of Asia and Indonesia.	
	Frequency of occurrence (i.e., detection) more likely for lower radio frequencies.	

for the passage of a large-scale TID. This large-scale wave can cause ionospheric tilts of up to 3^0 - 4^0 for short periods of time, with tilts on the order of 1^0 being common for an hour or more (after Rao, 1975).

The relevance of ionospheric data taken at one point when extrapolated out to successsively larger distances may be estimated from the following example. The positions of known transmitters are estimated from received signals, and the fixing errors between calculated and known ranges are determined. This is done two ways: first, by assuming the ionosphere is uniformly flat, and second by assuming the ionosphere is tilted, based on the ionosonde data at the receiver. The scatter plots of the fixing errors are shown in Figure 3 for three transmitters at successively greater distances (Rao, 1981). The 45^0 line in each plot corresponds to the condition of equal errors from the two models. Thus, points above the line correspond to smaller "tilted ionosphere" fixing errors, while points below the 45^0 line correspond to smaller "untilted" fixing errors. For the case where the actual range is 30 km, using the tilted ionosphere model (i.e., "extrapolating" the ionosonde data out to a point 15 km away) produces noticably smaller fixing errors. The same is true when the range is extended to 70 km (i.e., ionosonde data are extrapolated to a point 35 km

TABLE II

TRAVELING IONOSPHERIC DISTURBANCES

TYPE OF DISTURBANCE	WAVELENGTH AND STRUCTURE	MOTION	PERIOD	FREQUENCY OF OCCURRENCE	SOURCE
Large-scale	>1000 km horizontal wavelength. Wave front width on order of 1000 km. Phase fronts tilted nearly horizontal. Retains shapes over thousands of kilometers.	>300 m/s north to south.	30 min - 3 hr usually 1-3 cycles.	Infrequent, less than daily.	Events in the auroral zone. Strong correlation with magnetic activity.
Medium-scale	10's-100's km horizontal wavelength. Wavefront width 100's to over 1000 km. Phase fronts tilted 30^{o}-60^{o} from vertical. Do not retain shapes well over distances >100 km; energy does propagate globally.	100-250 m/s variable directions, with seasonal trends.	10-100 min several cycles to trains.	Daily, more common in daytime.	Tropospheric phenomena. Upper atmospheric and polar winter sources.
Small-scale	$\leq$10 km horizontal wavelength. Structure not well resolved.	100-250 m/s (est) variable directions.	<10 min long trains to families as wavelength decreases.	Daily	Probably tropospheric; not well established.

TABLE III

MAGNITUDE OF ERRORS IN POSITION LOCATION
ACCURACY DUE TO ERRORS IN ANGLE OF ARRIVAL
FOR E AND F REGION LAYERS

RANGE (ACTUAL)	200 km		300 km	
HEIGHT OF REFLECTING LAYER	105 km	250 km	105 km	250 km
Range Error (km) for:				
1^0 elevation angle uncertainty	6.9	9.8	10.9	11.5
3^0 elevation angle uncertainty	20.6	29.3	32.7	34.4
Cross-Range Error (km) for:				
1^0 azimuthal angle uncertainty	3.5	3.5	5.2	5.2
3^0 azimuthal angle uncertainty	10.5	10.5	15.6	15.6
Range Error (km) for:				
10 km height uncertainty	19.4	7.7	27.6	11.5

away), although to a somewhat lesser extent. However, when the range is extended to 170 km (ionosonde data must be extrapolated 85 km), there is no advantage to using the tilted ionosphere model over the untilted one. Thus, in this example, extending ionospheric data from one point to another for distances of more than 50-100 km does not seem to be of any advantage.

Similar conclusions have been reached by measuring the angles of arrival of HF signals from a series of geographically spaced transmitters. Assuming a one-hop propagation path some useful results were obtained by cross-correlating the angle of arrival deviations of the signals from pairs of transmitters (Ernst, et. al., 1975; Hoover, 1976; Rao, 1981). A maximum in the cross-correlation function means that the variations in the angle of arrival at one location are reproduced at the second location some time T later. The results indicate significant decorrelation of a persistent ionospheric pattern over distances of 50 to 100 km. This does not mean that a single frequency component of the composite disturbance necessarily decorrelates over distance on the order of 100 km, but rather that interference between waves from different sources, or from the same source traveling different paths, can result in the observed decorrelation.

The approximate ranges of the quasi-periodic variations of several ionospheric irregularities have been listed in Table II. While there have been numerous studies pertaining to the statistics of occurrence of characteristic periods or frequencies, the subject of temporal coherence seems to have received less attention. The temporal coherence of the ionospheric waves (particularly the medium-scale TID's) observed at a given location depends on the sources of the waves and the sources' duration. In practice, many waves due to several sources or multi-path propagation from a single source are probably present at any given instant of time. One can assign a decorrelation time to a group of waves which would essentially represent the time it takes the group to change form due to interference of the several components. This approach was taken by Walton (1971) who found a predominant decorrelation time of approximately 5 minutes. The approximate range in speeds for medium-scale TID's is 100-250 m/s. Using the decorrelation time of 5 minutes, this would yield a "decorrelation distance" in the range of 30-75 km, in good agreement with the previous estimates of spatial coherence.

Therefore, the spatial and temporal coherences of ionospheric sounding information appear to be on the order of 50-100 km and 5 minutes, unless sophisticated techniques of spectral analysis are employed to extract individual waves which remain coherent over much longer distances and time periods.

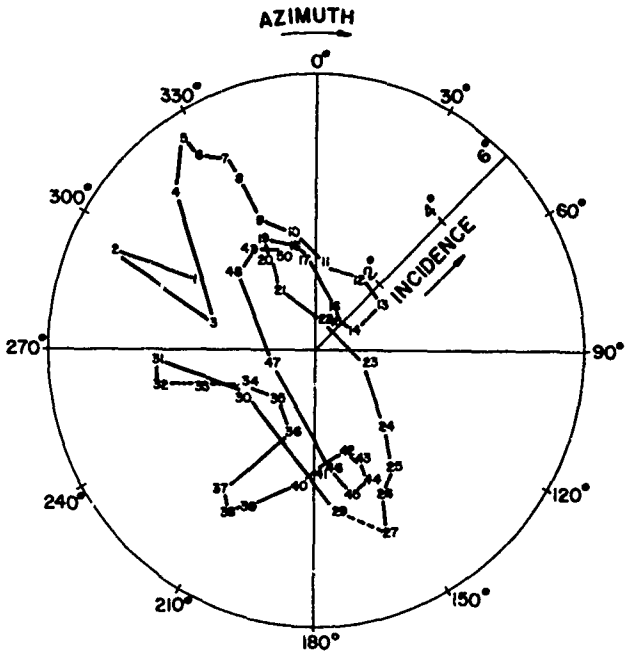


Figure 1: Incident angle, plotted as radial distance from the origin, versus azimuthal angle, plotted as the polar angle, for the return signal of a vertically incident ionosonde over a 50 minute interval.

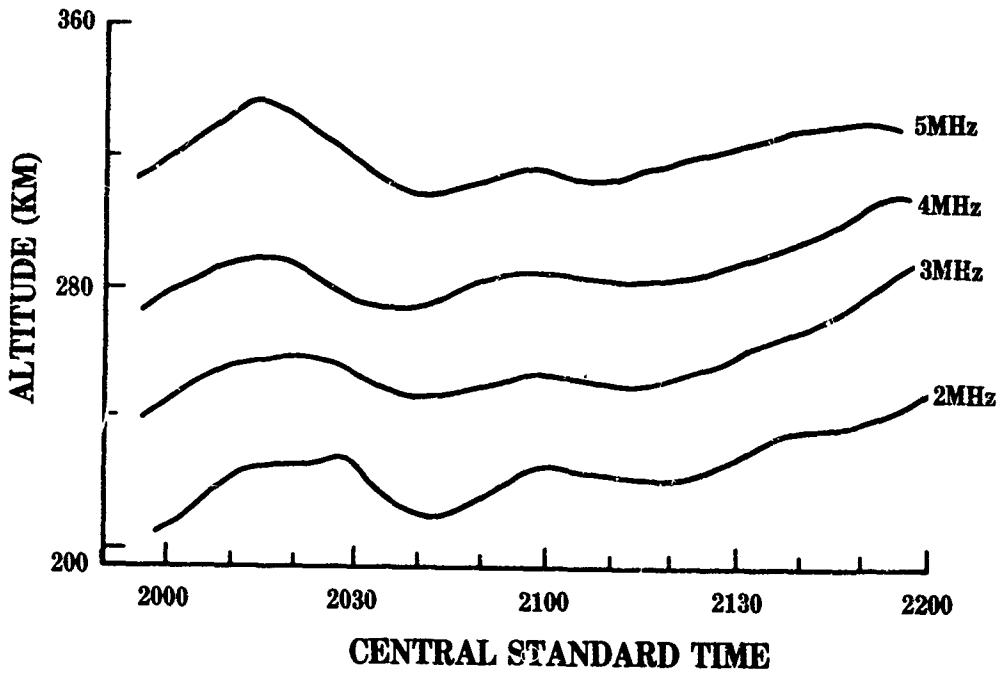


Figure 2: The altitude of the constant plasma frequency contour (i.e., signal reflection heights) versus time for a large-scale TID.

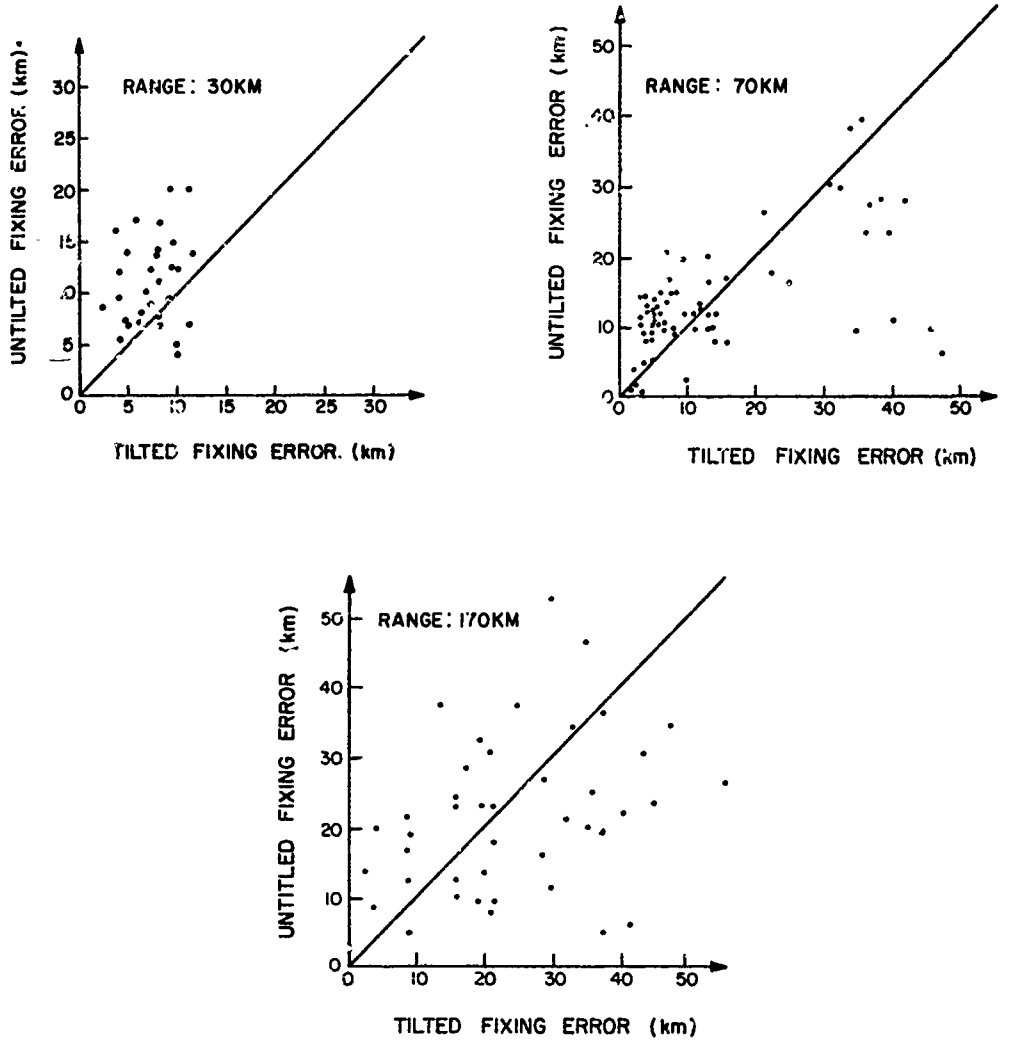


Figure 3: The scatter plots of the radio source location errors using two ionospheric models for three t mitters at successively greater distances.

IV. DISCUSSION AND CONSLUSION

The previous sections have outlined the types of ionospheric phenomena which are most likely to affect radio source location. Estimates of the magnitude of some of the induced errors in position location accuracies have been given, and estimates of the spatial and temporal coherence of medium-scale TID's have been made.

Most of the ionospheric irregularities considered here can be thought of as wave-like phenomena which propagate through the neutral atmosphere, with corresponding effects on the ionized component of the atmosphere. The possible exception, at least in behavior, is sporadic E, although the wind shears thought to be responsible for sporadic E may be due to gravity waves. The main effect of sporadic E is to introduce uncertainity as to which ionospheric layer, E region or F region, the signal is returning from and to provide additional opportunities for multimode reflection of signals. The concepts of spatial and temporal coherence, or decorrelation, are applicable to sporadic E only in the sense that the physical size of the patch and its motion will give some estimate of how long the phenomenon is expected to persist at any one given point.

Ionospheric tilts and traveling ionospheric disturbances produce the major problems in radio source location by introducing error in the angle of arrival and uncertainties in the height of the reflecting layer. Multiple reflections from wave-like or corrugated layers also producing multimode interference. The ionosphere tilts due to solar influence and the large-scale TID's show good spatial and temporal coherence, maintaining their shapes over long distances and for times on the order of an hour or more. Superimposed on these more regular waves are the spectrum of medium-scale TID's. The medium-scale TID's are the most frequently occurring ionospheric irregularities and the ones showing the least spatial and temporal coherence. While single frequency components of medium-scale TID's may persist for longer distances and periods of time, the composite TID seems to decorrelate over distances of 50-100 km and times of 5 minutes. Table IV provides a summary of this group of properties for ionospheric irregularities.

The emphasis of this paper has been implicitly directed toward the concept of a single station locator using a single, vertically incident ionospheric sounder. One basic constraint in radio source location is that the reflection point of the intercepted HF signal is some distance from the receiving site. For a spatially and temporally uniform, or at least slowly varying, ionosphere this would present no problem. In reality, however, ionospheric irregularities reduce the usefulness of information gathered at one point when transferred to another point. Medium-scale TID's seem to place the severest limitations on extrapolation of state-of-the-ionosphere information. The basic space and time decorrelation parameters have been listed above.

If the concept of the single station locator is retained, then the inherent problem of errors due to the ionosoheric propagation path may be approached by either living within the constraints of a single point sounding, or expanding the ionosonde network. Four options will be explored: (1) Limit the use of the system to live within the current constraint of a single, overhead sounding; (2) Place the ionosonde at the anticipated mid-point of the propagation path; (3) Resolve the various frequency components of the ionospheric disturbance, and (4) Employ an integrated network of ionosondes.

The first option recognizes the basic constraints of the available ionospheric data and limits the use of a single station locator to within these constraints. This implies that the system would be of essentially strategic use, but has the advantage that it is essentially self-contained and could be fielded well behind the forward battle area.

The second option attempts to gather ionospheric data where it would be most useful, near the anticipated ionospheric reflection point. This implies a preselection of range and direction over which radio source location will be attempted so that the system and the sounder can be optimally positioned. Thus additional constraints on system use have been imposed, not the least of which is the transfer of data from the ionosonde to the receiving station. Once the step of moving the ionosonde has been taken, it is a natural extension to consider using several ionosondes.

A basically analytic approach to the problem of decorrelation of ionospheric data would be to resolve the various frequency components of the medium-scale TID. It is felt that the individual components maintain their coherence and propagate over distances and times longer than 50-100 km and 5 minutes. While this approach is conceptually straightforward, it is not clear how much ionospheric data would be needed as input. It would appear, however, that data from several ionsondes would be needed. The additional data correlation and analysis effort would place very large requirements on any fielded computer system.

TABLE IV

TYPE	SPATIAL DECORRELATION	TEMPORAL DECORRELATION	MAJOR EFFECTS
Sporadic E	~100s km, depending on the size of the patch.	Minutes to hours, depending on relative location of patch and its drift velocity.	Uncertainty in height of reflecting layer. Multimode propagation.
Ionospheric Tilt (solar effect)	100's-1000's km	Hour or longer	Uncertainty in angle of arrival and height of reflecting layer.
Large-scale TID	1000's km	30 minutes to several hours.	Uncertainty in angle of arrival and height of reflecting layer. Multimode propagation.
Medium-scale	50-100 km	~5 minutes	Uncertainty in angle of arrival and height of reflecting layer. Multimode propagation.

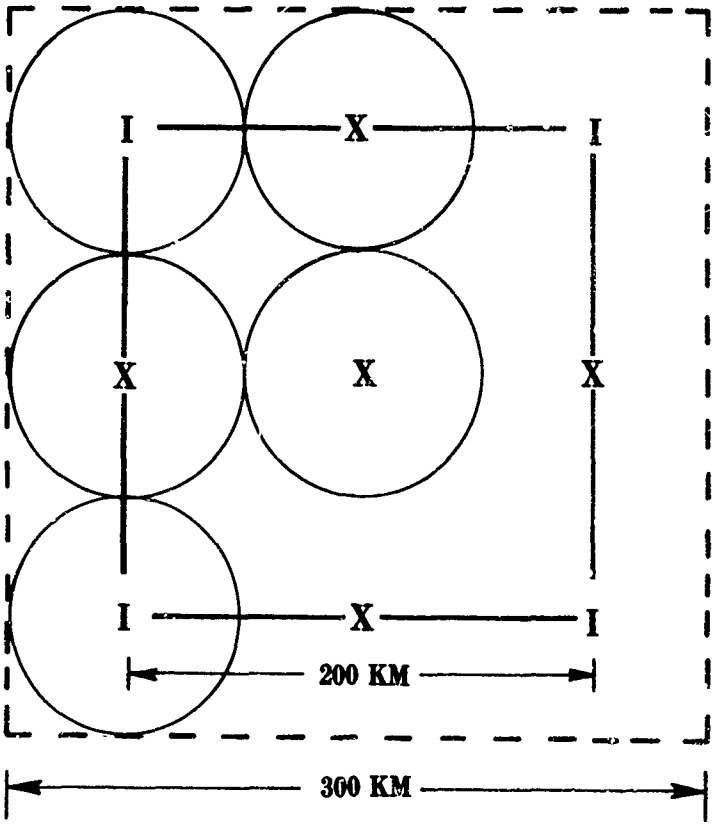


Figure 4: Schematic of how to locate four ionosondes such that maximum ionospheric data over a large area can be gathered from both vertical and oblique soundings.

The tendency toward the use of more than one ionosonde is apparent. An integrated network of ionosondes could be used effectively to gather sufficient ionospheric data over a large area. As an illustration, place four ionosondes on the corners of a square 200 km on a side. By using vertical soundings at each ionosonde, and oblique soundings between ionosondes, the state of the ionosphere could be determined at nine points along the perimeter and at the center of this square. Thus 78.5% of all the points within a slightly larger, superimposed square, 300 km on a side, would be within 50 km of a sounding point and nominally within the "decorrelation distance". No point in the square would be more than 71 km from a sounding point; see Figure 4. Therefore a relatively large area can be covered by as few as four ionsondes, provided they are integrated into a network using both vertical and oblique soundings. Nominally enough data can be gathered to adequately define the state of the ionosphere for any reflection point within the larger area, and potentially enough data are available for more complicated analysis approaches. The trade-off is that the complexity of a fieldable system has been greatly increased.

In summary, the effects of ionospheric irregularities on radio source locations have been investigated. It was found that ionospheric data (tilt and virtual height of reflection) taken at one point lose their validity when extrapolated over distances of 50-100 km or times of more than 5 minutes. Thus ionospheric soundings should be made more frequently than 5 minute intervals. A single ionosonde is usually not sufficient to adequately represent a large enough area of the ionosphere. An integrated network of ionosondes, using both vertical and oblique soundings is recommended.

REFERENCES

Ernst, E. W., J. D. Dyson and N. N. Rao, HF/DF Techniques Investigation, Final Report, Vol I, ECOM-0125-F, May 1974.

Ernst, E. W., J. D. Dyson and N. N. Rao, HF/DF Techniques Investigation, Second Quarterly Report, RRL Report No. 456, Radio Research Laboratory, University of Illinois, March 1975.

Fejer, B. G., and M. C. Kelley, Ionospheric Irregularities, Rev. Geophys. Space Phys. 18, 401-454, 1980.

Georges, T. M., Ionospheric Effects of Atmospheric Waves, ESSA TR-IER57-ISTA 54, US Govt Printing Office, Washington, D.C., October 1967.

Hines, C. O., and Colleagues, The Upper Atmosphere in Motion, Geophys. Mono. Series 18, Amer. Geophys. U., Washington, D.C., 1974.

Hoover, K. E., Ionospheric Modeling for HF Radio Source Location from a Single Site, PhD Thesis, University of Illinois, 1976.

Kent, G. S., Measurement of Ionospheric Movements, Rev. Geophys. Space Phys 8, 229-288, 1970.

Munro, G. H., Traveling Disturbances in the Ionosphere, Proc. Royal Soc., 202, 208-223, 1950.

Munro, G. H., Traveling Disturbances in the F Region, Aust. J. Phys. 11, 91-112, 1958.

Peterson, V. L., Ionospheric Irregularities, TR-80-005, Centenial Sciences, Inc., Colorado Springs, CO, September 1980.

Rao, N. N., A Large Scale Traveling Disturbance of Polar Origin, Planet. Space Sci. 23, 381-384,1975.

Rao, N. N., Ionospheric Irregularities in HF Radio Source Location, Technical Report, D.O. 1684, TCN-80-304, Battlele Columbus Laboraties, Research Triangle Park, NC, January 1981.

Walton, E. K., An Investigation of Directional Deviations of Steeply Downcoming HF Radio Waves Due to Traveling Ionospheric Disturbances, PhD Thesis, University of Illinois, 1971.

Yeh, K. C., and C. H. Liu, Acoustic-Gravity Waves in the Upper Atmospher, Rev. Geophys. Space Phys. 12, 193-216, 1974.

COMMUNICATIONS IN THE "SKIP-ZONE" USING BACKSCATTERING FROM IONIZED METEOR TRAILS AND E-LAYER TURBULANCES

H. E. R. Jones, Scientist
Litton Data Command Systems
Agoura, California 91301

I. Summary

The feasibility of using ionized meteor trails and E-layer type anomalies (Sporadic E-layer reflections) occurring in the lower ionosphere as passive reflectors for VHF transmissions is discussed in terms of their use and application to low rate data transfers and communications from remote transmitter and receiving sites to a central base station. Such networks can be used to provide weather reporting, oceanographic data, terrain intelligence data and for emergency and back-up communications to augment Command Control, Communications and Intelligence (C^3I) Systems.

Information on meteor-burst communications systems is well documented. One of the first systems to employ this technique was the "JANET" system (1) in Canada and the "COMET" system (2) in the United States and in Europe. These systems were originally designed to take advantage of oblique scattering or to use the communications engineering terminology forward scattering via ionized meteor trails. Using this propagation mode, the range of a communications system can be extended around the earth's bulge up to distances of 1200 miles. These systems employed high gain antennas at each end of the link with their main lobes directed slightly to the left or right of the path between the stations to take advantage of certain characteristics of the meteor trails (7). Typical transmitter powers range from 1KW to 10KW and are tailored to the information rates and service being offered.

The system to be described here represents a different approach to the meteor-burst communications technique. Quite early in the investigations of the meteor-burst phenomenon it was recognized that certain types of ionized meteor trails classified as overdense are near perfect reflectors of radio waves. The formation of these trails occurs normally from 75 to 120 kilometers above the surface of the earth. If a transmitting antenna is pointed upwards towards these trails, then some of the energy radiated will be scattered back to earth to be detected by a receiver with its antenna pointing upward in the same manner. Using this near-vertical incidence backscattering technique it is possible to communicate within the "skip-zone" the (region between the limit of a transmitter ground wave and the point at which the first sky wave is encountered). At frequencies of 30 to 80 MHZ which are above the critical frequency of the E-layer, the skip-zone is considered to be a blind radio zone. In addition, these passive reflectors will enable local communication obstacles such as mountains and deep valleys to be surmounted easily. While it is recognized that the intermittent nature of the communications channel will restrict its use, nevertheless there are enough potential applications of this type system to encourage further study.

Research has revealed that an enormous amount of data has been collected on meteor occurrences and ionized trails since 1945 by radio astronomers using VHF backscatter radars [3] [4]. This data is related to such properties as the electron density of the trails, the frequency of occurrence, radar cross sections, direction and angle of arrival and the duration of the echoes encountered.

Much of the data available from these radar measurements can be used in the design of a communications system, although most of it has been obtained using high power pulsed radars of from 10 to 300 KW peak power. The investigation began with a study of what could be achieved with small transmitter powers. In one of the early tests of this system in July of 1961 using transmitter powers of 100 milliwatts and narrow-band frequency shift modulation, signals were transmitted over a 7000 foot mountain range over a thirty mile circuit. The receiving station was located in a noisy electrical environment so that it had to be operated from 6 p.m. in the evening to 6 a.m. in the morning. Typically, during this time and on a number of occasions later, bursts of signal of two to eight seconds were recorded three to four times during the night and early morning. Admittedly, this is representative of a very low duty cycle for a communications link and yet for some types of data,

for example, snow water content data which is collected by men on foot in the mountains of the Western United States, a once-a-day reading or for that matter once-a-week reading is more than adequate.

During the winter of 1962 and 1963 a demonstration meteor-burst telemetry link was established between Mt. Hood, Oregon and Seattle, Washington to transmit snow water content dat. from one of the Soil Conservation Service's snow pillow sensors. The link employed a broad beamwidth low gain antenna at the receiving end and a 3-element yagi antenna at the transmitting end of the link oriented to produce a common area of illumination almost directly over the receiving station. Results of these tests were successful and lead to some of the conclusions presented later concerning antenna optimization. This encouraged the Department of Agriculture to proceed with the development of its Snonet Meteor Burst-Telemetry System.

One of the serious problems slowing further development of the system in the late 1960's was the inability to obtain frequency allocations for meteor-burst systems. In addition, satellite and tropospheric systems with their wider bandwidth, high data rate capabilities became much more attractive to military and commercial system users.

Recently, however, there has been renewed interest in meteor-burst systems. The nuclear survivability of the meteor-burst medium is known to be superior to tropospheric, satellite and HF skywave systems. The developement of small hand-held digital terminal equipments, inexpensive microprocessors and small solid state memories have all stimulated this revival.

Other indicators of this renewed interest are; The Defense Communications Agency (DCA) is studying the feasibility of developing meteor burst for the Minimum Essential Emergency Communications Network (MEECN). The Department of Agriculture is now implementing an extensive system of meteor-burst telemetry for the collection of hydrologic data. The Department of Energy (DOE) has initiated the development of a system for transmitting emergency messages between 14 DOE offices nationwide.

Topics to be discussed in this paper are:

1. The Nature of the Meteor Scatter Phenomena
2. The Radio Propagation Model
3. Some System Design Considerations
4. Some Examples of Baseline Systems Design.

The results presented here are a combination of field evaluations and analytical studies carried out over a number of years. Test links have been set-up at various locations in the western states. In one such test the remoted terminal was located in Bakersfield California with the Receiving station in the heart of Los angeles in an area of high man-made interference. Some of the methods used to suppress this interference are described here.

Generally it can be concluded that the meteor-burst medium is well understood for both the forward-scatter mode and the backscatter mode described here. This conclusion is based upon the results reported here and the results of others in the field whose efforts have been documented from time to time during the last forty years. Enough data has been collected to enable the design of data collection and emergency communications systems with the now predictable performance within the limitations of the meteor-burst propagation mode.

In the area of C3I Systems where secure communications and emergency communications are required to operate in a nuclear environment, more work is required in developing survivable antenna systems. As indicated, in these type systems, the transmitted power is directed upward. It is possible then to conceive of antennas flush with the ground and hardened in place with the operating personnel and equipment in protected bunkers below the surface.

From the equipment point of view, the antenna designs described here are simple, easily transportable and erectable. Transmitters, receivers and signal processing equipments adaptable to this type of system have now been reduced to back-pack dimensions and when not being used in the scatter mode, can readily be reoriented to the line-of-site (LOS) mode. The optimum frequency range still appears to be in the 80 to 100 MHz band and still produces the least costly designs.

Using this near vertical incidence scattering mode described here, broad-beamwidth antennas are desirable so as to provide broad coverage of the E-region above and around a transmitting and receiving station. Using this technique, sufficient overlap of patterns exists between the stations up to distances of 200 miles in any direction around the transmitting station. In addition, the diurnal and seasonal variations in the angle of arrival of meteor radiants can readily be accommodated.

II. Nature of the Meteor Scatter Phenomenon

It is well established both by visual and radar type measurements that as the earth moves through space in its orbit around the sun it is continuously bombarded with meteoric debris. As these meteors enter the atmosphere of the earth, they burn up leaving in their wake long columns of ionized particles at altitudes ranging from 80 to 120 kilometers. In some instances the electron density of these trails causes them to act as near perfect relectors of radio waves. Almost as soon as they form they begin the process of diffusion which is brought about in some instances by the action of high altitude winds. Based upon radar and radio measurements, these trails have been found to last from fractions of a second to several minutes and to be intense enough to support intermittant radio communications over distances up to 1200 miles. The intermittent nature of these trails has led to the term "Meteor-burst" radio propagation. Typcial frequency assignments for this type of propagation are in the range of 30 to 180 MHz. The lower frequency is bounded by critical ionospheric frequencies to avoid ionospheric refractions. The upper frequency limit is due to the characteristics of the receiving equipments and the meteor trails themselves.

Early observations of the effect of meteor ionization resulted from ionospheric research. Appleton and Barnett (9) and Breit and Touve (10) noticed sudden increases in the electron density of the ionosphere as far back as 1925. The first suggestion that meteors contributed to periodic increases in ionization came from Skellet (10) in 1931.

During World War II the presence of meteors was detected on the early search radars operating at 20 to 40 MHz. Later during the V2 missile raids on London the presence was detected on Army Gun Laying radars operating at 60 MHz.

Since that time radar backscatter measurements have continued to be recorded throughout the world at various locations (3) (4). As a result of these efforts, the distribution of meteoric particles in well defined streams of particles (meteor showers) is a matter of record (5). These streams represent only a small percentage of the total number of meteors entering the atmosphere of the earth daily. The figures are staggering and in the order of billions. In order to avoid a mass panic situation it should be noted that these particles range in size from 0.8 micron diameter all the way up to 8cm diameter and most totally disintegrate in the atmosphere (see TABLE I).

Meteor distributions. Backscattering and forward scattering from meteor trails are subject to diurnal and seasonal variations in the occurrance of useable trails. The rotation of the earth produces a diurnal variation in the rate of arrival of meteor radiants, for a particular location on the earth's surface. Meteors along the apex of the earths way are swept up by atmosphere, while radiants at the antapex are produced by meteors overtaking the earth. The distribution of activity throughout a day is characterized by sine-like, diurnal variation with a maximum at 6 a.m. and a minimum at 6 p.m. Figure 1A and 1B illustrate this feature for a low power system. Seasonal variations tend to produce a low level of activity in the first three months of the year reaching a minimum in February and a maximum during the months of July and August. A more detailed discussion of duirnal and seasonal variation occurs in reference (7).

Nature of meteoric reflection. When a meteor trail is illuminated by radio waves from a transmitter, each electron in its trail backscatters like a Hertzian dipole causing a portion of the trail to act as an antenna, re-radiating the waves in a conical pattern as depicted in the sketches of Figures 2A and 2B. The received signal is the vector sum or field strength sum scattered back by each trail electron. In Figure 2A, the meteor path is normal to the path from the transmitter so that it reflects the radio waves back to the transmitter in the way that it occurs in radar-type measurements. Figure 2A also shows the footprint of the received signal which is returned to a band of locations on the earth's surface, providing the means for communication from the transmitter to any point in the shaded area. Figure 2B shows a different meteor position and inclination, resulting in a new region in which communication becomes possible. For any given receiving site within the shaded area, the average waiting time for a meteor with appropriate position and orientation to provide a reflected path from the transmitter will be approximately the same. If the area surounding the transmitter contains a number of evenly distributed receivers, the waiting time to establish

communication (if any receiver is accepted) will be reduced by the ratio of the average area accessible with each single meteor to the total area containing the receivers. Thus the traffic capacity tends to increase as the area served is increased.

Duration and frequency of occurrence of meteor trails.
The larger meteors produce trails of higher reflectivity. Because of the greater density of ionized particles, these trails also persist longer than those produced by small meteors. Figures 3A and 3B shows the general dependence of reflected signal amplitude and theoretical duration as functions of meteor size. Average waiting time between meteors is also given to aid in interpretation. Note that two regions appear, exhibiting distinctly different behavior. The small meteors produce a low density of ionized particles, permitting the incident wave to pass through with only part of the wave energy scattered. The high ionized particle density produced by large meteors results in total reflection of the radio wave; the expanding trail appearing as a conducting cylinder. The two types of behavior are referred to as "underdense" and "overdense" trails. As the particles diffuse the apparent size of this cylinder first increases and then diminishes as the particle density drops. Eventually, the overdense trail degenerates into the underdense case, but by this time the reflected signal is of little significance.

The duration of overdense trails occurring at intervals of a few minutes are on the order of a second and longer. However, disturbance of the trail geometry by high altitude winds often results in fading of these signals due to interference between waves reflected from different parts of the trail. Table II is a summary of the average burst duration and average interval between bursts derived from field test data.

Figure 4 is an extract from a chart recording obtained at a receiving site over a 60 mile path using a 10 Watt CW transmitter. It is shown here to illustrate the form of the signals obtained from overdense trails and underdense trails. Moving from left to right, the long duration seven second signal is typical of the signal from an overdense trail. The signal rises slowly to its maximum and exhibits the effects of fadirg. The second signal depicts the form of the signal from an underdense trail characterized by a steep rise in signal level followed by an exponential like decay. The next signal to the right is an example of a smaller type overdense trail and to the far right another overdense signal appears.

Interspersed between larger bursts of signal described above are many smaller bursts ranging from 0.1 to 0.2 seconds duration. In a data transmission system, signal processing enables these shorter signal bursts to be used. For example, the bandwidth shown on the chart recording is 2.8 KHz, by suitable bandwidth narrowing, signal-to-noise improvements of from 10 dB to 14 dB are achievable. The effect is to increase the amplitude and usable duration of the bursts of signal.

III. Radio Propagation Model

Basic transmission equation. The transmission equation for communications via meteor trails follows from the familiar radar equation governing free space and reflective losses. For monostatic, the equation becomes:

$$\frac{P_R}{P_T} = \frac{G_R G_T \lambda^2}{16 \pi^2 R^4} \rho \qquad (1)$$

where P_T = Effective transmitted power
P_R = Incident power at receive antenna
$G_R G_T$ = Power gains of antennas toward trail, relative to an isotropic radiator in free space.
R = Range to meteor trail in meters
λ = Wavelength in meters
ρ = Backscattering cross section in square meters

When the receiver and transmitter locations form bistatic geometry this relationship is modified to:

$$\frac{P_R}{P_T} = \frac{G_R G_T \lambda^2}{16 \pi^2 R^2_T R^2_R} \rho \qquad (2)$$

where R_T, R_R = Range in meters of trail from transmitter and receiver, respectively.
β = Effective cross sectional area in square meters for forward or oblique scatter.

The development of transmission equations describing both backscatter and forward scatter models is presented in more detail in reference [3] [4]. Figure 5 illustrates the path geometry for two projected cases. Angle 2ϕ in this case is termed the "Forward Scatter Angle" and can vary from typically 4°to 60°. The bisector of this angle is normal to the tangent plane for prolate spheriod geometry.

Forward scattering requires that the effective cross-sectional areas and the resultant transmission equations be modified from the back scatter relationships. With specular reflection and bistatic geometry, a meteor trail to be useable must be tangent to one of a family of ellipsoids (prolate spheroid) with foci at the transmitter and receiver. In addition, the trail must be at least the length of one-half of the first Fresnel zone at the point of tangency and observable by both the transmitter and receiver. As a result of the oblique geometry, the length of the first Fresnel zone increases considerably over backscatter dimensions. The greater length of the Fresnel zone means a larger number of electrons are contributing to the scattered energy and the received signal should be greater than for backscatter under the same circumstances. Oblique incidence is equivalent to an increase in wavelength, $\lambda \sec \phi$, thus a larger initial signal with an increase in duration will be received.

Field tests. The radio propagation model just presented is a simplistic approach used initially. To obtain more refined data, a field test program was initiated. The objective of the field tests was to obtain sufficient data to make a statistical estimation of expected path loss, and of signal bursts and frequency of occurrence over short distance paths. Plenty of data was available on radio-type measurements and on long-distance forward scattering communications paths. No data could be found on systems using near vertical incidence with "forward scattering angles" (this term is used here with some reservations in this case) of from 4 up to 60 degrees. This as angle 2ϕ in Figure 5.

The test radio links consisted of CW transmitter and dipole antennas at the transmitting end and, a dipole antenna and calibrated receiver at the receiving end.

The antennas. The antennas were spaced 3/8 λ above a reflecting ground plane. Using this arrangement the gain was reduced in the zenith direction and peaks out at angles of 45° around each station (see Figure 7).

Receivers. The AGC of the receivers was calibrated using a resistive termination and frequency standard. The rise and fall of the AGC voltage was adjusted so that the rise time was less than 1 millisecond and the decay time less than 5 milliseconds. The bandwidth of each receiver was measured using the frequency standard and holding the AGC voltage constant for both a high and low AGC voltage reading.

Transmitters. The output power of the transmitters was measured and the effects of short-term drift noted.

Test procedures. Test links were established using the equipments described with the transmitter moved in various directions around the receiving station, or baselines extending from 10 miles to 200 miles. These links were set up in various locations in the Western United States. These tests made it possible to assemble a large amount of data on the durations of bursts of signals, the average interval between bursts and the amplitude of the received signal over a time period of many months. Eventually this data was broken down into blocks and the slow process of counting, for example, the number of 0.1 second, 0.5 second, 2 seconds, 5 seconds and so on even up to 4-hour bursts was started. At the same time the interval between bursts of the same duration was being recorded also. Once this process was underway an attempt was made to reduce the data to manageable proportion using the computer. The end goal was to develop a set of design criteria to be used by system designers.

One observation made from these tests was that up to transmitter and receiver spacings of 100 KM the path loss was fairly constant, from then on the fall off was quite gradual. This effect is illustrated in Figure 6. The second observation was that the duration and interval between bursts began to assume a predictable pattern. For example, it was found a half-second burst could be expected to occur every minute, a 2-second burst every five minutes and a 10- to 30-second burst every 100 minutes. This data is summarized in Table II for average burst duration and interval between bursts.

Computer program. The original path loss computations were based upon the simple radar equations derived from the basic transmission equation for the backscatter case by Lovell and Clegg [4] and by Mckinley [3]. The following assumptions were made by the author:

a) That flat earth could be assumed for distances up to 150 kilometers.

b) That only overdense trails would be used for communciations since they presented near perfect reflectors for radio waves.

c) Broad beamewidth antenna would be used to illuminate large area in the E-region above and around each station. In addition, a large common area of E-region illumination would exist between transmitting and receiving station.

It was known as indicated earlier that the electron densities of trails varies typically from 10^{14} electrons/meter to 10^{18} electrons/meter and that the duration of the overdense trail varies from about 0.3 seconds up to 30 seconds and more. This had been verfied using CW transmitters and calibrated receivers. The decision then was to combine the experimental field work with a computer program to try simplifying the system design procedures.

Since the received power varies as $1/R^4$ where R is the distance in meters between the transmitter and reflecting trail and between the reflecting trail and the receiver, the effect of variations in the length of R brought about by increases in the spacing between the transmitter and the receiver over the distances contemplated for short range communication could be ignored. It had been found to be the case experimentally. This effect can be seen in the curve of path loss versus distance shown in Figure 6.

The formulas recorded by Sugar[7] have been used as the basis of the computer program. Here the effects of polarization antenna gain, variation in R1, R2, λ and $\emptyset$ are considered in more detail. The effect of the variation of some of these parameters upon duration, interval between bursts have been exposed as a result of the program.

In addition to this work, an attempt has been made to determine on a statistical basis, the duty cycle of a multi-station complex by taking into consideration some of the complex geometry of the meteor communication path. These considerations are not very significant at low frequencies. At higher frequencies, particularly in multiple access systems operating at 150 MHz, the results are more significant and to be a subject of another paper.

Trade-off study. The computer study was instrumental in producing the system trade-off study shown in Figure 8. This approach is directed at data transmission systems where meteor-burst type systems perform best. Using this type data a designer, for example, knowing the data rate, band width and transmitter power, can determine the duty cycle of the proposed system. This approach is a first attempt at refining the data for use and is still being studied.

IV. System Design Considerations

Duty cycle. One important aspect of meteor burst communications is the suitability of the medium for transferring the amount of information to be transmitted. The number and duration of useful trails as well as the phase amplitude characteristics of the medium are important. When modulation scheme, matched to the characteristics of the medium, is selected for a system, it then becomes possible to establish system duty cycle.

Duty cycle can be defined as Duty Cycle = Mean System Rate/Fixed Instantaneous Rate. When maximum use is made of the medium, the duty cycle represents the availability of the medium. This represents a proper definition for availability, since it includes the profitable or unprofitable use of those trails which overlap in time.

The duty factor for maximizing the amount of information transformed has received extensive treatment by various authors (12). Its use in this application has to be modified.

In data collection systems where a small amount of information is stored at a number of sites and only requires periodic transfer to a central location, access time and useful trail duration longer than a minimum interval is the first consideration. If the system is designed to use trails which are longer than the transfer time for a message and each trail is considered to support one transfer, then the number of useful trails becomes the dominating availability factor. Thus, duty

factor, as generally used, does not apply. The approach then is to match the information to the burst duration and interval between bursts.

Remote sites. At remote sites prime power requirements become a major consideration. The power source must be self-contained. Typically, with today's technology there are combinations of solar cell and storage battery; battery, and thermo electric generator and propane gas heat source. Unattended operations demand a reasonable level of reliability. Each of these considerations impacts on the cost. Cost also becomes a controlling consideration in view of a large number of remote sites with very small amounts of information to be transferred. These factors suggest closed-loop operation be utilized. Its direct cost for the type of environment where trail availability is critical will usually be less than the indirect costs imposed by open-loop operation. Typical trade-offs associated with open-loop and closed-loop operations are listed in Table III. Closed-loop operation was considered necessary for baseline system. There are occasions, however, when open-loop operations could be featured for certain sites and conditions, typically, a low rate data transfer system using low power transmitters overdence type trails.

Receiver noise. The noise level for VHF receivers is primarily determined by cosmic noise unless operation is contemplatd within high-noise environments of suburban or large city areas. Figure 9 illustrates that cosmic noise more so than receiver noise establishes the operating noise level for frequencies between 30 and 100 MHz. Cosmic noise will be used as limiting the reference noise limit.

At 30 MHz cosmic noise is defined as 1.38 X 10-23 X 30,000 X BW or -183.9 dBw/Hz. This is roughly 20 dB above an ideal receiver so that any reasonable noise figure in a receiver (6 to 12 dB) is acceptable.

Remote site operaton in rugged terrain reduces the effect of man-made noise levels and in addition offers substantial relief from the potential interference from undesirable radio signals from other sources.

In some instances, particularly in the short range backscatter type system, the terrain can be used to reduce interferrence by locating the antenna in a valley. Another approach is to locate the antenna in, for example, a chain link fence enclosure. This approach has been very effective in reducing man-made interferences. Man-made noise is more often than not vertically polarized and propagates along the ground. The fence effectively diverts this interference away from the antenna.

Antenna systems. The approach here as previously mentioned is to use low gain antennas to give broad coverage of the E-Region above and around the transmitting and receiving stations. As a result of the overlap of antenna patterns, a large areas of the E-region is illuminated by the receiving and transmitting antennas. The ionized trains form in a downward direction, and as the earth rotates, the direction of these trains appears to change giving rise to the diurnal variation noted earlier. In long-haul systems the positions of the transmitting and receiving antennas are changed from one side of the great circle path to the other to take advantage of the meteor trail hot-spots as they have been described (1). In the backscatter system any meteor trail which forms to the right or left of the line-of-site path between stations, behind the transmitting station or receiving station provided it meets the geometric orientation stated previously will provide a communication path.

Another advantage that has a bearing on multi-station networks all operating on the same frequency, is a trail aligned to provide a communications path between stations, for example, A and B may not be favorably aligned for station C. In some instances when a large number of stations have to be polled periodically; they can all be polled at the same time. Another possible advantage is that conventional LOS systems and backscatter systems can be operated simultaneously, in some cases on the same frequency, without interfering with each other.

Modulation schemes. Various types of modulation schemes have been tried, all have been found to lend themselves to this type of medium. These are as follows:

a) FM/FM telemetry.

b) Frequency following receiver. An FM receiver can be made adaptive to the extent that it can search out a signal and adjust its frequency accordingly. The phase locked loop is a common implementation of this technique.

c) Band dividing receiver. Rather than performing detection in a wide band, it is practical to divide the IF band pass into a number of smaller bands. Each band must be equal to twice the signal band width and the bands are made to overlap by the signal band width. In this manner, there is assurance that the signal will be totally enclosed by one of the sub-bands.

A selection of the sub-band to be used is made by a signal-to-noise measurement on each channel. Since each sub-band is twice the required band width, a 3 db penalty is incurred. This is considerably better than the 10 db loss from detecting in the wide band. The multiplicity of channels and the channel selection equipment, however, represents a considerable increase in complexity. The bank of sub-bands can be used to additional advantage. By connecting the output of alternate bands in phase opposition, a considerable negation of impulse noise can be achieved.

d) Pulse compression. The problems of instability can be resolved at the 140 BPS rate by using frequency sweep rather than shift. This system can resolve the stability problem and still retain an effective noise band of 140 cps. This is basically a pulse compression technique. The mixer is the detector and responds to a sweeping frequency by producing a beat note between the instantaneous frequency and the signal frequency seconds earlier. By using two sweep rates, the binary states can be conveyed. Since the two channels are detected in phase opposition, the system is largely immune to impulse noise. The noise bnd width is the width of the band passfilter and this need only be large enough to support the 140 BPS data rate.

Interrogation. The power in the base station can be increased considerably above the remote transmitter. It is desirable to achieve turn-on as fast as possible to best utilize the transmission time. In addition, the power to band width ratio of the down-link and up-link should be equal in order for the system to realize the greatest reciprocity.

Voice transmissions. Voice transmissions have been transmitted over the backscatter link using continuous loop tape recorders at both terminals driven from stable frequency sources, in one case using the local grid. The voice transmissions are recorded on tape and then continuously radiated. As the path opens up intermittantly, parts of the message are received and recorded on the receiver site tape recorder until all of the message is assembled and read. The method used here is perhaps better described as a meteor-trickle system rather than a meteor-burst system.

V. A Baseline System Design

System requirements. The system used here as an example is described as follows: The system is designed to transmit data sequentially using frequency shift keying as the modulation mode. The calculations are based on operation at signal/noise threshold.

Duty cycle and burst length. The system specification states the requirement to interrogate up to 300 remote stations. Each station has an average of 4 sensors with a desired interrogation of one to two times daily, between 1200 and 2400 daily transmissions are required. Since the frequency of interrogation may be a function of the sensor and the location, it is desirable to interrogate each sensor, rather than each remote station. To achieve this on a single interrogation per burst basis, requires operation on the 0.1 second burst level. There are an average of 2880 bursts of 0.1 second duration per 24-hour period. The message length must be somewhat less than .1 second. The path delay will be roughly .001 second. We can anticipate a turn on time of .001 to .01 seconds. The message length must be less than .09 second.

The average waiting time for a 0.1 second burst is 0.5 minutes. If the system uses a sequential interrogation program and waits for each reply before interrogating the next sensor, the probability of receiving a reply in the half minute period is 0.5.

Probability of not receiving the reply is (1-.5); probability of not receiving a reply in N periods is (1-.5)N; probability of receiving a reply in N periods is 1 - (1-.5)N.

or

$$P_{(a)} = 1-.5^N$$

The waiting time can thus be determined as N(.5) minutes.

For a .99 probability of message receipt:

$$.99 = 1-.5N$$

Therefore, the time to interrogate a sensor will average 0.5 minutes and will be greater than 3.35 minutes in no more than 1% of the interrogations.

Information, rates, and deviation. A typical transmission requires a parameter resolution of 1%. This is equivalent to an ensemble of 100 discrete levels. A 7-unit code would be required. To send this data in .05 seconds, a rate of 140 BPS is required. The data required for a 0.1% accuracy is 10 bits. At 140 BPS, this would require .071 seconds. The theoretical bandwidth required for an arbitrarily small error rate is described for a data rate C as:

$$C = BW \log_2 (1 + S/N)$$

or

$$BW = \frac{140}{\log_2 (1 + S/N)}$$

for

$S/N = 0$ db, $BW = 140$ cps

$S/N = 4.6$ db, $BW = 70$ cps

$S/N = 12$ db, $BW = 35$ cps

This same relationship applies to an analog channel. That is, using a 140 cps channel for distinguishing 100 discrete frequency states of a single tone:

$$\frac{I}{T} = 140 \log_2 (1 + S/N)$$

or for a S/N = 1,

$$T = \frac{7 \text{ Bits}}{140 \log_2 2} = .05 \text{ sec.}$$

Where T is the integration time required to distinguish the tone frequency within 1%.

Permissible Signal to Noise Ratio (S/N).

The determination of S/N threshold for the system is made on the basis of acceptable error rates in the output signal channel. The detected S/N can then be translated to predetection of IF S/N.

Assume a maximum message length is 10 bits. An error rate of 10^{-3} would result in one incorrect message in 100 transmissions. An error rate of 10^{-4} would result in one incorrect message in 1000 transmissions. Since the concern here is with threshold values, an error rate between 10^{-3} and 10^{-4} would seem adequate.

The error rate as a function of output S/N is shown in Figure 10 for a two-state signal. As can be seen, the desired error rate requires an output signal-to-noise ratio of 11 to 12 db. 12 db is taken as the acceptable threshold value. These establish the theoretical relations for information and bandwidth.

Cosmic noise. A 15-watt transmitter power (+ 11.7 dbw) is available. Cosmic noise is defined as $1.38 \times 10^{-23} \times 30{,}000 \times BW$ or -183.9 dbw/cps. This is roughly 20 db above an ideal receiver so any reasonable noise figure in the receiver (6 to 12 db) is acceptable.

Noise. Man-made noise is generally of an impulse nature and can be effectively eliminated (especially in a narrow band system) by wide band limiting and phase cancellation. The cosmic noise level of -183.9 dbw/cps is, therefore, taken as the system limitation.

The noise level in the 140 cps band pass is:

$$-183-.^{0} \text{ dbw} + 10 \log 140 = -162.4 \text{ dbw}$$

The required signal level is:

$$-162.4 \text{ dbw} + 12 \text{ db} = 150.4 \text{ dbw}$$

For the 15-watt transmitter and the path loss associated with the 0.1 second bursts, the received signal level is:

$$11.7 \text{ dbw} - 150 \text{ db} = -138.3 \text{ dbw}$$

This leaves a 12.1 db margin if the band pass is confined to 140 cps. Since the system must have all instabilities resolved if detection is confined to this band pass, the 12 db margin can be utilized to resolve this instability.

REFERENCES

1. Davis, G. W. H. Gladys, S. J., Lang, G. R., Luke, L. M. and Taylor, M. K., Proc Inst Radio Engineer, 1957, pp. 45 1666-1678, The Canadian JANET System.
2. Bartholome, P. J. and Vogt, I. M., IEEE Trans Communications Technology, Vo. Comm-16, April 1968, Comet - A meteor-burst System Incorporation ARQ and Diversity Reception.
3. McKinley, D. W. R. Magraw-Hill Book Company, In., 1961, Meteor Science and Engineering.
4. Lovell, A. C. B., Oxford at the Clarendon Press, 1954, Meteor Astronomy.
5. Sugar, G., Die Handbuch der Phsik, Vol. 52.
6. Jarvis, Alan, Navy Times, Dec. 29, 1980, Corps Closer to Field Computer Use.
7. Sugar, G., Proc IEEE, February 1964, Radio Propagation by Reflection from Meteor Traile.
8. Vincent, W. R., Woiram, R. T., Sifford, B. M., Jayne, W. E., Peterson, A. M., Proc IRE, pp. 45, 1693-1700, A meteor Burst System for extended Range VHF Communications.
9. Appleton, E. V., and Barnett, M. A., Proc Royal Society, 1925, pp. 109, 621-641, On Some Direct Evidence for Downward Atmospheric Reflection of Electric Rays.
10. Brest, G. and Tuve, M. A., Phys. Rev. 28, pp. 554-573, 1926, A Test of the Existance of the Conducting Layer.
11. Skellet, A. M., Physic Review, 1931 pp. 37, 1668, The Effect of Meteors on Radio Transmission through the Kennelly - Heavyside Layer.
12. Campbell, L. L., Hines, C. O., Proc IRE, December 1957, Bandwidth Considerations in a JANET System.

TABLE I
Estimates of the number of sporadic meteors incident on the earth

	Mass	Radius	Number of this mass swept up each day	Electrons per meter of trail length
Particles passing through the atmosphere and falling to earth	10	cm	10	
Meteors totally disintegrated in the upper atmosphere	10	1cm	100	------
	10	2cm	1000	------
	10	0.8cm	10000	10^{13}
	1	0.4cm	100000	10^{14}
	10^{-1}	0.2cm	10^{6}	10^{15}
	10^{-2}	0.08cm	10^{7}	10^{15}
	10^{-3}	0.04cm	10^{8}	10^{14}
	10^{-4}	0.02cm	10^{9}	10^{13}
	10^{-5}	80 microns	10^{10}	10^{12}
	10^{-6}	40 microns	10^{11}	10^{11}
	10^{-7}	20 microns	10^{12}	10^{10}
limit of radar measurement	10^{-8}	8 microns	10^{13}	10^{10}

Note; Micro-meteorites particles ranking in size from 0.2 micron to 4 micron are not shown. As far as known they produce little effect on the E-region. Estimates of the number swept up by the earth are 10^{8} per day.

TABLE II
Average burst duration, interval and electron trail density

Burst Duration	Interval Between Bursts	Typical Values of Electron Density Electrons/Meter
0.1 sec	0.5 min	10^{13} E/M
0.5 sec	1 min	10^{14} E/M
2 sec	5 min	10^{15} E/M
25 sec	100 min	10^{16} E/M
4 hrs	240 hrs	10^{18} E/M

TABLE III
Tradeoffs Between Open and Closed Loop Operation

Open Loop	Closed Loop
A prior schedule Simple On-Off Sequence	Interrogation-response Automated On-Off Sequences
Inflexible schedule after manual set	Flexible Scheduling
Selection of complete from incomplete incomplete message sequences implied	Retransmission control for messages
Receiver and automated circuitry not required	Receiver and automatic circuitry required
Continuous transmission for an interval	Periodic, on demand
Requires high probability of trails trails during an interval Implies greater system sensitivity/ power and antenna beam width	Presence of a trail Sensitivity/power based on average duty cycle requirements
Less costly for reasonable system parameters	More efficient and flexible
Potentially better system reliability	Better operational capability
Less communications reliability	Matched Communications System

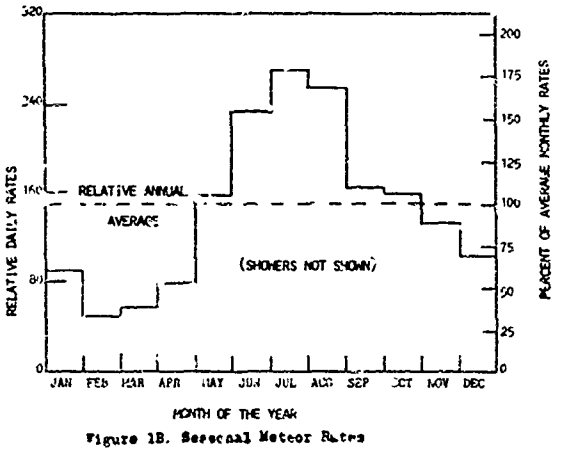


Figure 1B. Seasonal Meteor Rates

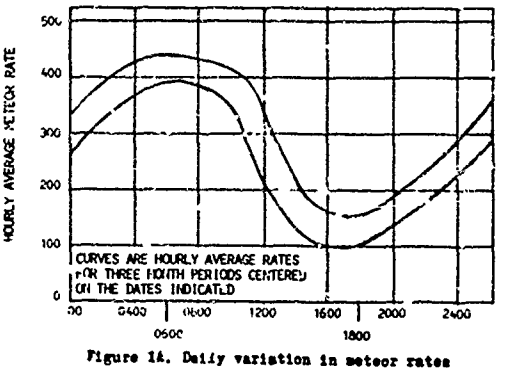


Figure 1A. Daily variation in meteor rates

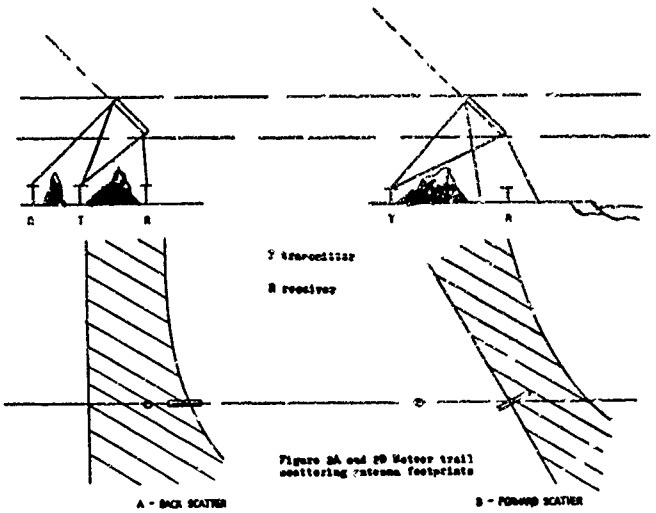


Figure 2A and 2B Meteor trail scattering antenna footprints

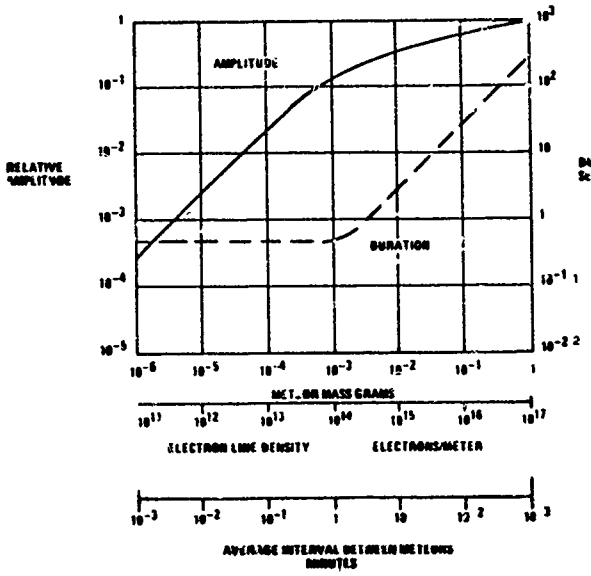


FIGURE 3A METEOR REFLECTION CHARACTERISTICS

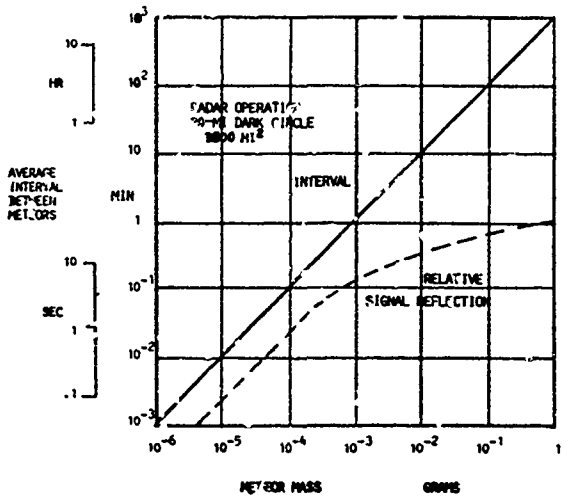


FIGURE 3B. FREQUENCY OF METEOR APPEARANCE

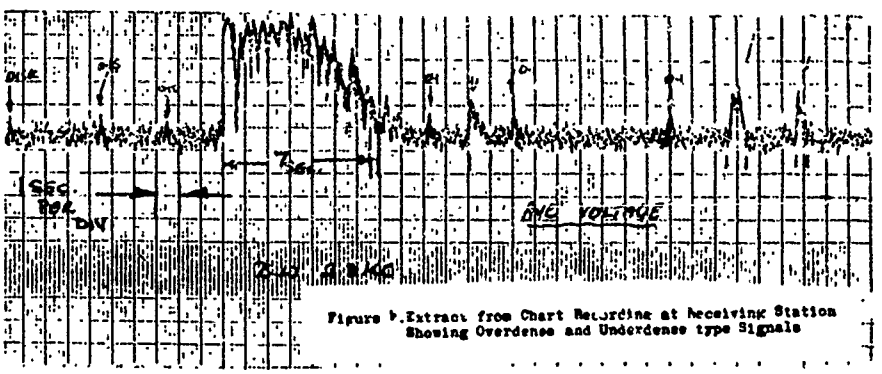


Figure 4. Extract from Chart Recording at Receiving Station Showing Overdense and Underdense type Signals

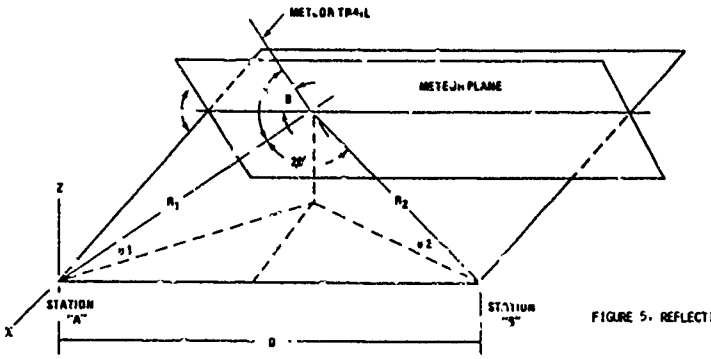


FIGURE 5. REFLECTION GEOMETRY AND IDENTIFICATION OF GEOMETRIC TERMS

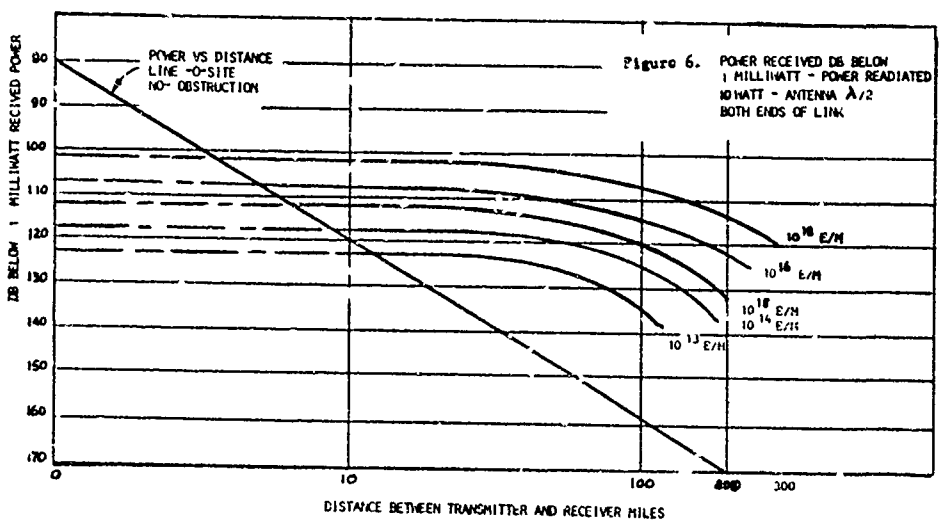
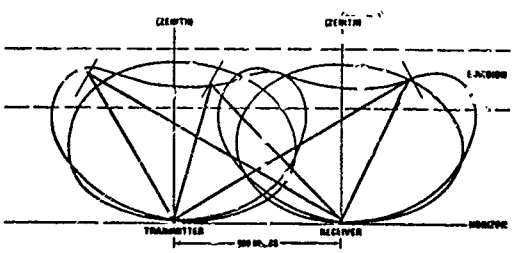
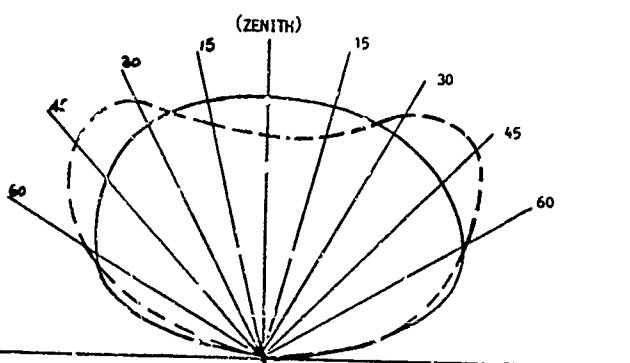


Figure 6. POWER RECEIVED DB BELOW 1 MILLIWATT - POWER READIATED 10 WATT - ANTENNA $\lambda/2$ BOTH ENDS OF LINK



SKETCH SHOWING HOW ANTENNA PATTERNS OVERLAP IN THE E-REGION



Dotted line, dipole $3/8\lambda$ above ground plain
Solid line , dipole $1/4\lambda$ above ground plain

FIGURE 7. METEOR BURST ANTENNA PATTERN

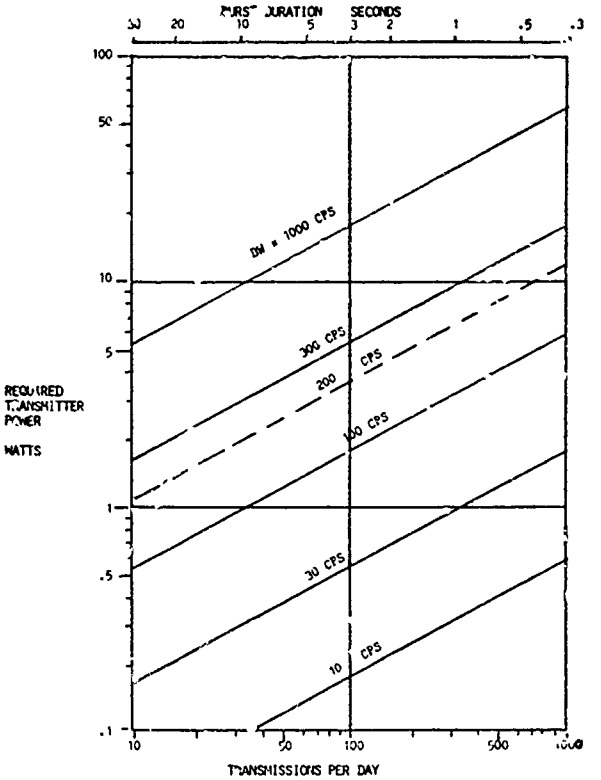


FIGURE. 8 SYSTEM TRADE-OFF EXAMPLE

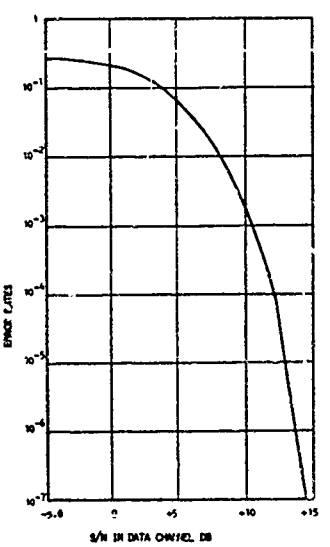


FIGURE 10. DIGITAL ERROR RATES FOR [illegible] TWO STATE SIGNAL

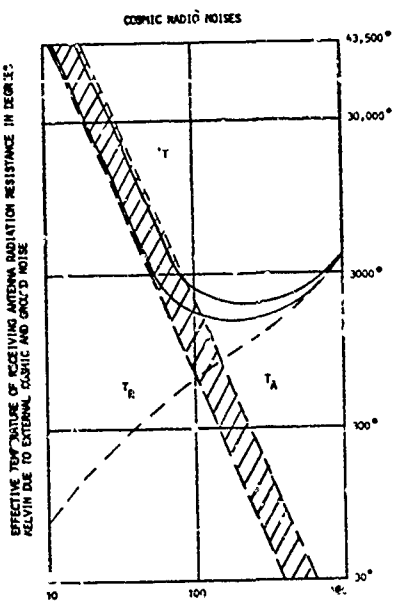


FIGURE 9. TOTAL NOISE TEMPERATURE T_T HATCHED AREA DIURNAL VARIATION
T_R RECEIVER NOISE TEMPERATURE T_A COSMIC NOISE TEMPERATURE

SEASONAL VARIATIONS IN THE AURORA IONOSPHERE

B. W. Reinisch, J. Tang and Huang Xueqin
University of Lowell Center for Atmospheric Research
Lowell, Massachusetts 01854

ABSTRACT

Automatic processing of digital ionograms from Goose Bay, Labrador (65° corrected geomagnetic latitude), permits fast evaluation of the diurnal and seasonal variations in the ionosphere. Ionogram data from 1980 are processed and displayed in ways that show the seasonal variations in foF2 and hminF. The effects of the mid-latitude trough that moves over the Goose Bay Station in the late evening is discussed.

1. INTRODUCTION

Automatic processing of ionograms has become possible with the availability of advanced digital sounders. Since about 1970 when the first digital ionograms were routinely recorded on magnetic tape we were developing software and hardware techniques for the extraction of the echo traces from the ionograms (Bibl et al, 1973). The initial concept was to decide immediately after transmission of each individual frequency which signals are echoes and only retain the amplitudes and heights of the identified echoes. For the archiving of routine ionogram observations and for the monitoring of ionospheric trends (Buchau et al, 1978), this method is acceptable if a sufficient number of echo points per frequency is allowed. But this approach is clearly inadequate for many scientific investigations in a disturbed ionosphere where the researchers want to see the complete range-versus-frequency display, i.e. the raw ionogram. Such a display not only reveals all the nuances in the signal characteristics but it also establishes a continuous transition to the familiar analog ionograms. Bibl and Reinisch (1978) described the on- or off-line printing of Digisonde ionograms using an optically weighted font to retain the digital resolution in the quasi-analog ionogram display (Figure 1).

A special objective is the automatic calculation of vertical electron density profiles. In this case, the vertical echo trace must be extracted and oblique, ducted and multiple echoes disregarded. During auroral night conditions the trace must be found within the spread F signals, relying on amplitude, doppler and incidence angle informations that are contained in the digital ionograms. This is only possible by examining the ionogram in its entirety.

In this paper we shortly describe the automatic ionogram scaling algorithm, developed on a main frame computer, and its application to some 5,000 ionograms from Goose Bay, Labrador. The diurnal and seasonal variation of the critical frequency and the minimum virtual height of the F2 layer are used as indicators for the variability of the auroral F region.

In 1971 the Air Force Geophysics Laboratory (AFGL) established the Goose Bay Ionospheric Observatory, equipping it with riometers, magnetometers, satellite receivers for total electron content measurements, and a Digisonde 128 for bottomside ionospheric sounding. Goose Bay is located at 53° geographic and 65° corrected geomagnetic latitude which means that the auroral oval (Feldstein and Starkov, 1967) reaches Goose Bay around local midnight during Q = 3 conditions (Gassmann, 1973), and correspondingly earlier during high magnetic activity. The mid-latitude F region trough, extending some 100-200 km equatorward from the southern edge of the oval is observed in the Goose Bay ionograms as a rapid decrease in foF2.

In 1978 the new Digisonde 128PS (Bibl and Reinisch, 1978) was installed at Goose Bay capable of measuring the signal polarization, incidence angles and doppler shifts. The Geomonitor (Reinisch and Smith, 1976) currently performing the real-time ionogram scaling at Goose Bay does not yet take advantage of the additional information available. The new scaling algorithm presented here will eventually be implemented for the on-line processing.

2. AUTOMATIC SCALING OF MULTIPARAMETER IONOGRAMS

Automatic scaling of ionograms requires proper tagging of the signals with regard to polarization and incidence angle. To monitor the overhead ionosphere only vertical echoes should be considered. The tape-recorded vertical ionograms from Goose Bay contain for each frequency-range pixel the polarization, incidence angle and doppler information in addition to the amplitude. An amplitude ionogram with mild spread F is shown at the bottom of Figure 1. The X-cusp emerges out of the O-trace at about 5 MHz, and the automatic separation and identification of the O and X trace from the amplitudes alone would be difficult. By printing only the vertical signals with O-polarization we obtained the ionogram in the upper half of Figure 1, which contains the data points the automatic scaling algorithm is using. Figure 2 shows a quiet daytime ionogram which is substantially simplified when the fxF1 and fxF2 cusps are removed (top of Figure 2). There are a number of bite-outs within the vertical O-trace which are caused by oblique or X-polarization echoes with higher amplitude than the vertical O-echo in these particular pixels. Some of these holes are replenished during the processing by checking the pixels surrounding the holes.

The scaling algorithm starts the trace identification by finding the center windows, independently for E and F regions (see Figure 2). Allowing for reasonable slopes, the trace is then determined by sliding the window to the right (increasing frequency), and later to the left of the center position. This integrating window method appeared to be successful even under relatively disturbed conditions. Without going into details of the scaling algorithm, it should be mentioned that the resulting h'(f) trace may not be a smooth function suitable for use in a electron density profile inversion program that is based on the standard lamination technique (see for example Doupnik and Schmerling, 1965). The trace data would have to be smoothed which is not a simple task because of the peaks in h' at foE and foF1. Huang and Reinisch (1981) successfully applied the profile-fitting method to automatically scaled topside ionograms. This method finds the monotonic electron density profile, described by a polynomial, that best reproduces the h'-traces in the least-squares sense. We are in the process of applying this approach to the automatically scaled Goose Bay ionograms, considering the multi-layer structure of the bottomside ionosphere.

3. VARIATION OF THE AURORAL F REGION

Four months of ionograms for January, April, July and September 1980 were processed with the new algorithm. With three ionograms per hour, or 72 per day, a total of about 6,000 ionograms was automatically scaled. We only discuss here two important F-region parameters, the F2 layer critical frequency foF2, and the minimum height of the F-layer h'F. The results for the four seasons are displayed in Figure 3, using the same optically weighted number font as the ionograms. The left most panel shows foF2, the middle panel h'F, and the right panel the magnetic activity. The 72 daily readings are arranged in one line, with consecutive days following each other. Days with no data were simply deleted, namely January 25, April 5, 6, 17, 18 and July 19, 31, without serious effect on the evolving pattern.

January shows a well defined presunrise minimum in foF2 of about 3.5 MHz at 07 AST, and the morning increase in foF2 is very steady during the entire month. The sharp break in intensity indicates the 12 MHz line. During the first half of January foF2 is larger than 12 MHz for three to four hours around local noon, occasionally exceeding 14 MHz. After January 20, foF2 stays below 12 MHz. This is also true for the first few days in April. Starting with April 8, the foF2 peak values are about 8 MHz occurring in the late afternoon around 1800 AST. Values close to or above 10 MHz are reached on April 11, 14, 19, 22 and 27, centered around 14 AST. July shows the maximum at 18 to 19 AST reaching values of 8 MHz. The presunrise minimum of about 4.5 MHz occurs around 03 AST for April and July but it is not as sharply defined as in January. In September the pattern is changing again. The presunrise minimum of about 4 MHz is well defined around 04:30 AST, and toward the end of the month the peak of 11.5 MHz occurs around noon.

The gross features of the F layer are fairly well described in this way. The F-region trough is not always easy to see. The sharp decrease in foF2 occurring between 20 and 23 AST during the months of April and September seems to indicate that the equatorward edge of the trough is moving over the station. In January and September the transition to night conditions occurs gradually and identification of the trough in the foF2 maps requires comparison with the individual ionograms, which we have not yet done.

The minimum virtual height of the F region shows some very systematic patterns. As expected, the height reaches its maximum at night, displaying a large degree of variation, while the day values are much more consistent. In January, the heights level off shortly after sunrise to about 230 km and show little variation during the day. The apparent height minima at 09 and 16 AST are the result of a mistake in the scaling algorithm, which mistook the high cusp at foE as F-region echoes. We have eliminated this mistake but have not yet corrected the figure. In April, July and September the heights decrease rapidly at 05 to 06 AST to about 220 km and form a shallow minimum of 210 km around noon.

The night values for h'F vary considerably from day to day and the controlling function of the magnetic activity especially in the early night hours becomes evident when one compares the magnetic A-index with the height values. It is not our intention in this paper to analyze the F-layer height variations in terms of magnetic activity but rather to present a technique that makes it easy to relate different geophysical parameters. A good example is the magnetic storm starting on 7 April. The virtual height reaches values of 350 to 450 km at night, and on April 11 minimum heights of more than 400 km occur already at 16 AST. Similar good correlation exists in January and July, while the situation in September is less clear.

CONCLUSION

Automatic processing of ionograms provides the means for a fast and detailed survey of the diurnal and seasonal variation in the ionosphere. Our study of the auroral F-region at Goose Bay, Labrador, showed a fairly regular diurnal variation that gradually changed with season. Abrupt day-to-day changes in foF2 and h'F are clearly identified by the method, demonstrating its superiority over the use of monthly median curves. After having verified the accuracy and usefulness of our processing technique we can now apply it to an entire year of data to better understand the seasonal variations and the effects of magnetic storms. We can also apply our new electron density profile inversion algorithm to the automatically scaled Goose Bay ionogram data. It would appear desirable that as soon as the inversion algorithm is completed to implement this technique for real-time processing of ionograms. The HF and VHF communication user could then obtain an up-to-date electron density profile at any time of the day using telephone/modem links.

ACKNOWLEDGEMENT

This work was in part supported by the U.S. Air Force under Contract F19628-80-C-0113. The authors appreciate the helpful discussion with K. Bibl, J. Buchau and W. N. Hall.

REFERENCES

Reinisch, Bodo W. and Sheryl Smith, "Geomonitor, Digital Real Time Processor for Geophysical Data," Final Report, AFGL-TR-76-0292, University of Lowell Center for Atmospheric Research (1976).

Bibl, K. and B. W. Reinisch, "The Universal Digital Ionosonde," Radio Science, 13, No. 13, pp. 519-530 (1978).

Buchau, J., W. N. Hall, B. W. Reinisch and S. Smith, "Remote Ionospheric Monitoring," Proceedings of the IES, paper 5-5 (1978).

Feldstein, Y. I. and G. V. Starkov, "Dynamics of Auroral Belt and Polar Geomagnetic Disturbances," Planet. Space Sci., 15, pp. 209-229 (1967).

Gassmann, G. J., "Analog Model 1972 of the Arctic Ionosphere," AFCRL-TR-73-0151, AF Surveys in Geophysics No. 259 (1973).

Doupnik, J. R. and E. R. Schmerling, "The Reduction of Ionograms from the Bottomside and Topside," J.A.T.P., 27, pp. 917-941 (1965).

Huang, X. Q. and B. W. Reinisch, "Automatic Calculation of Electron Density Profiles from Digital Ionograms, Part II. True Height Inversion of Topside Ionograms with the Profile-Fitting Method," submitted to Radio Science (1981).

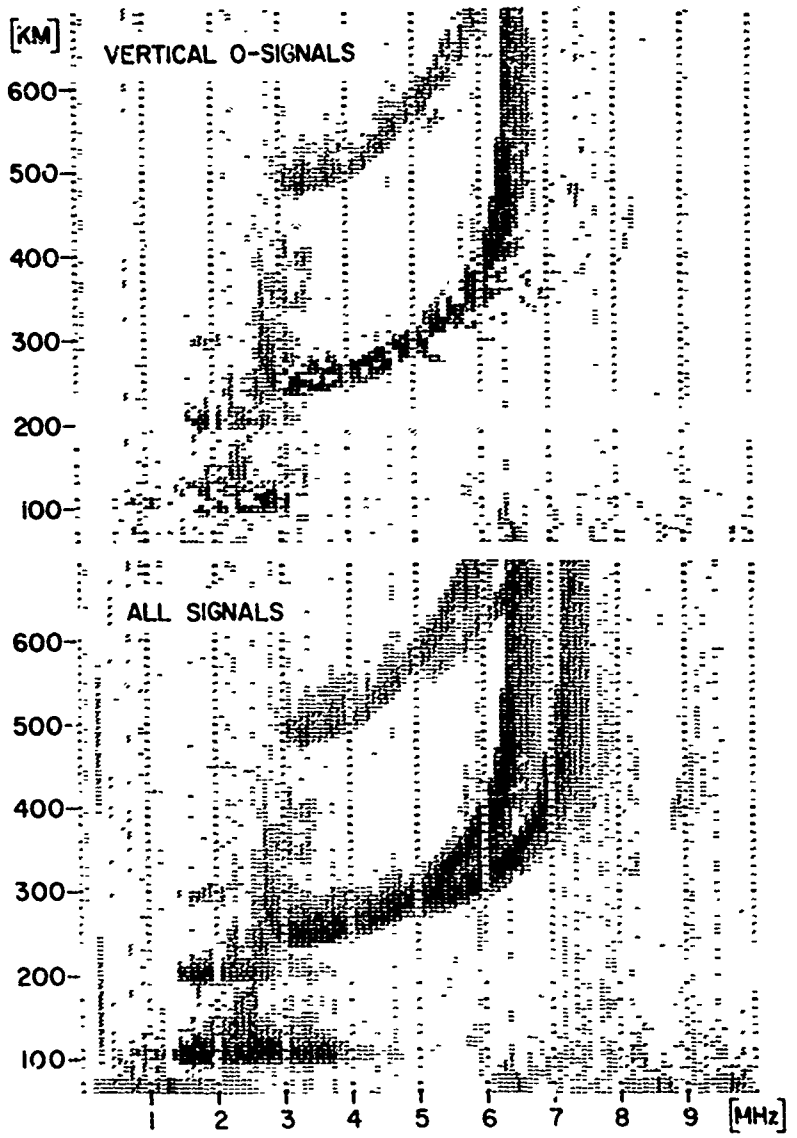
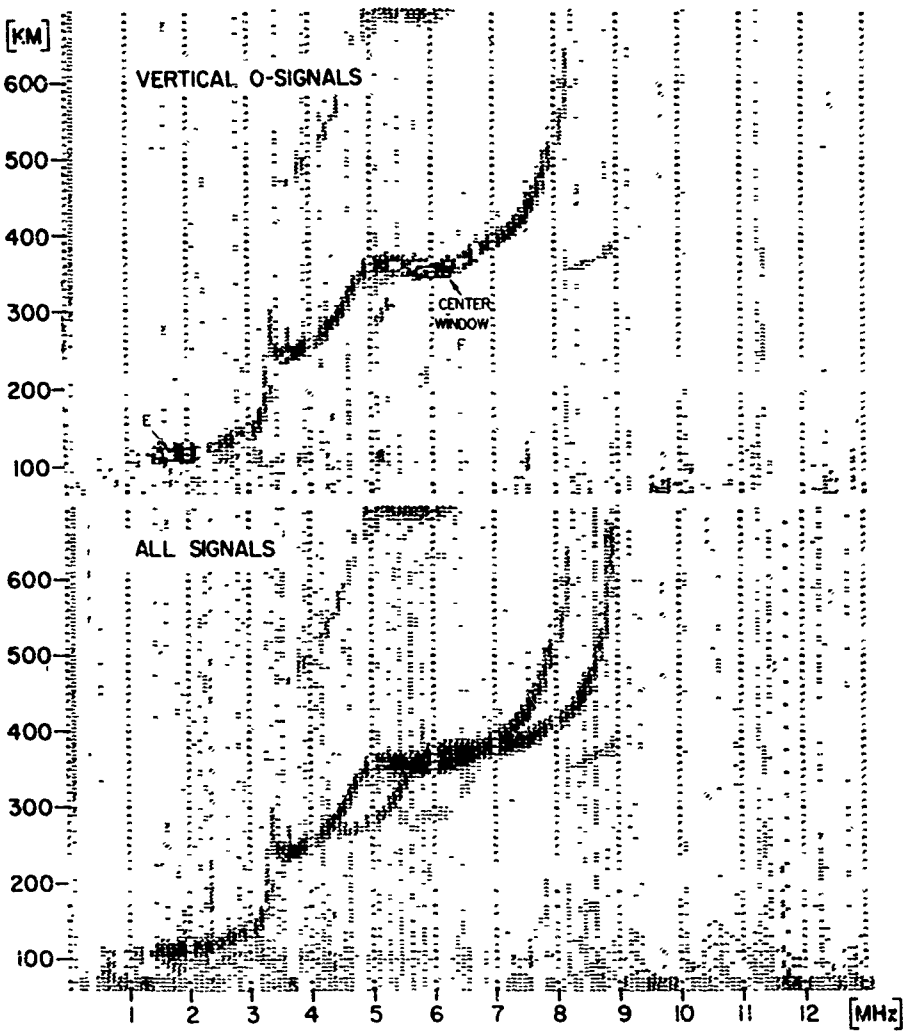


Figure 1



UNDISTURBED DAY IONOGRAM
GOOSE BAY 16 JUN 1980 17:20 AST

Figure 2

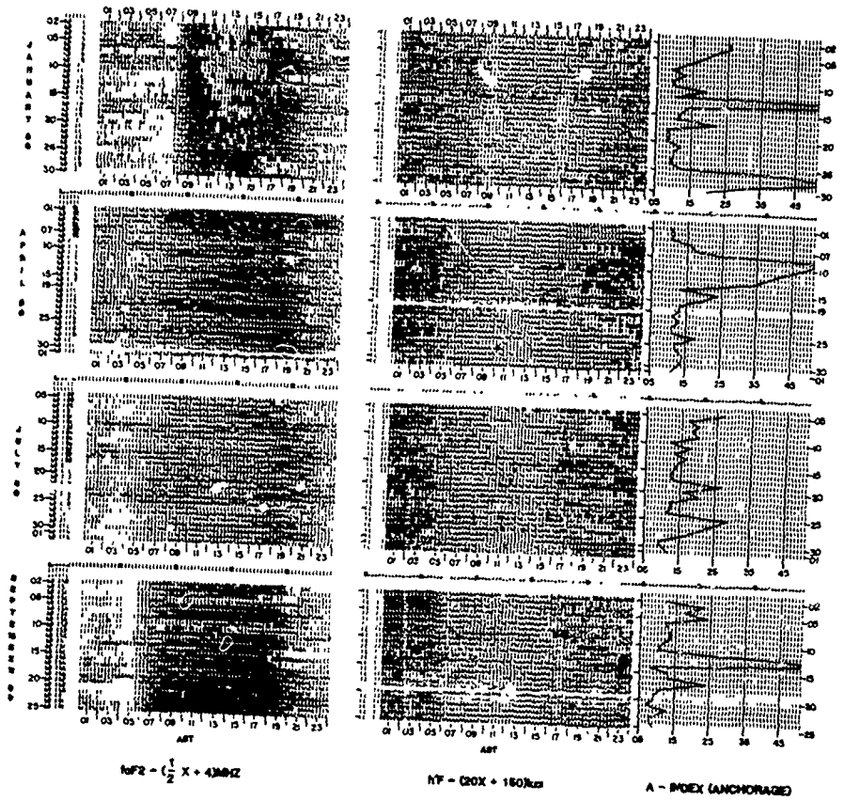

foF2 ~ (1/2 X + 4)MHz
hF ~ (20X + 160)km
A ~ PVDEX (ANCHORAGE)
AST
VARIATIONS IN AURORAL F-REGION

Figure 3

REMOTE SENSING OF HIGH-LATITUDE IONIZATION PROFILES BY GROUND-BASED AND SPACEBORNE INSTRUMENTATION

by

Richard R. Vondrak
Radio Physics Laboratory
SRI International
Menlo Park, CA 94025

ABSTRACT

Ionospheric specification and modeling are now largely based on data provided by active remote sensing with radiowave techniques (e.g., ionosondes, incoherent-scatter radars, and satellite beacons). More recently, passive remote sensing techniques have been developed that can be used to monitor quantitatively the spatial distribution of high-latitude E-region ionization. These passive methods depend on the measurement, or inference, of the energy distribution of precipitating kilovolt electrons, the principal source of the nighttime E-region at high latitudes. To validate these techniques, coordinated measurements of the auroral ionosphere have been made with the Chatanika incoherent-scatter radar and a variety of ground-based and spaceborne sensors.

INTRODUCTION

Evaluation of the effects of the ionosphere on radiowave systems requires an accurate specification of the spatial and temporal distribution of ionization. The most accurate ionization measurements are those made directly by in-situ spaceborne sensors, although they are limited to observations in the immediate vicinity of the rocket or satellite. For this reason, ionospheric specification and modeling are now largely based on data provided by active remote sensing by radiowave techniques (e.g., ionosondes, incoherent-scatter radars, and satellite beacons).

More recently, passive remote sensing techniques have been developed that can be used to monitor quantitatively the spatial distribution of high-latitude ionization. These passive methods utilize the measurement or inference of the energy distribution of precipitating kilovolt electrons, the principal source of the nighttime E-region at high latitudes. The energy distribution can be measured by rocket-borne or satellite-borne charged particle detectors. Another possibility is to infer the electron energy distribution from measurements of optical emissions at multiple wavelengths or from measurements of bremsstrahlung X-rays. From the energy distribution one can obtain the altitude profile of ionization by application of energy-deposition codes and atmospheric models.

The utility of these passive methods depends on the accuracy of the particle-emission and particle-atmosphere codes. To validate these techniques, coordinated measurements of the auroral ionosphere have been made with the Chatanika incoherent-scatter radar and a variety of ground-based and spaceborne sensors.

This paper illustrates the utility of the various techniques by showing two examples of simultaneous measurements with incoherent-scatter radar and passive remote sensors. In the first example the radar was coordinated with measurements by a rocket-borne array of charged-particle detectors. In the second, the radar observations were compared with data from ground-based multi-wavelength photometers. The final section of the paper discusses the accuracy and some of the limitations of the remote sensing methods.

PASSIVE REMOTE SENSING METHODS

Passive remote sensing of the ionosphere requires an ability to measure the principal source of ionization. In the high-latitude nighttime E-region, charged-particle precipitation generally supplies more than 90% of the total ionization. Other sources (galactic EUV and cosmic radiation, scattered EUV) produce a peak E-region density of only 2.5×10^3 cm^{-3}, while the E-region densities exceed 1×10^5 cm^{-3} even in the diffuse aurora. During most auroral conditions, kilovolt electrons are the principal source of E-region ionization.

The altitude profile of ionization produced by auroral particle precipitation is a function of the flux, energy distribution, and species of the precipitating charged particles. The production rate of ion-electron pairs in the atmosphere is proportional to the altitude profile of energy deposition. This profile can be computed by standard methods that have beer developed for both protons (Rather, 1970) and electrons (Rees, 1963; Berger, Seltzer, and Maeda, 1970). The ionization, n, in the E-region is related to the production, q, by the electron continuity equation:

$$\frac{\partial n}{\partial t} = q - \alpha n^2 + \vec{v} \cdot \vec{\nabla} n$$

where α is the effective recombination coefficient, and $\vec{v}$ is the velocity of the ionospheric plasma. If we assume equilibrium conditions and neglect transport, then Q is equal to the recombination loss αn^2. These assumptions are normally valid below altitudes of about 200 km. In this case, the accuracy in determining the ionization profile depends on the combined accuracies of the measured energy spectrum, the energy deposition codes, and the assumed profile of the effective recombination coefficient.

The photometric method uses the ratio of commonly observed auroral spectroscopic features to infer the characteristic energy of the electron flux producing the emissions. The physical basis of this method is that incoming electrons of different energies are stopped at different heights, where they excite atoms and molecules in proportion to the abundance of each species at that altitude. For example, an electron of low energy will produce emissions of atomic oxygen at 6300 Å in greater abundance than it will produce emissions of ionized molecular nitrogen at 4278 Å. Rees and Luckey (1974) have computed the ratios expected for incoming electron beams with Maxwellian energy distributions. If emission intensities are measured at two wavelengths, a fit to a two-parameter differential energy distribution can be made. For a Maxwellian distribution the parameters used are the total energy flux and the energy at the peak of the distribution (which is also referred to as the characteristic energy α). The mean energy of a Maxwellian distribution is twice the characteristic energy. Analyses of photometric and radar measurements have shown that this technique works for characteristic energies less than approximately 10 keV (Sears and Vondrak, 1981).

Measurements at visual wavelengths suffer from the restriction to only nighttime measurements, because of scattered sunlight. Two other wavelength regimes have been used for top side remote sensing: the ultraviolet and the X-ray (Imhof et al., 1974; Mizera et al., 1978). At these wavelengths, it is possible to monitor the dayside ionosphere with spaceborne sensors. Unfortunately, it has not yet been demonstrated whether such measurements can be used to infer unambiguously the differential energy distribution of precipitating electrons and to compute the associated altitude profile of ionization.

The capabilities and limitations of the optical remote-sensing techniques are listed in Table 1.

Table 1

SPACE-BASED OPTICAL REMOTE SENSING OF AURORAL PARTICLE PRECIPITATION

Wavelength	Albedo Corrections	Temporal Coverage	Possible to Infer Electron Energy Distribution?
Visual	Necessary	Night only	yes, if $\alpha \leq 10$ keV (assumed to be Maxwellian)
X-ray	Unnecessary	Day and night	yes, although some ambiguities ($1 < E_e < 100$ keV)
UV	Unnecessary	Day and night	Uncertain, probably impossible

IONOSPHERIC PROFILES INFERRED FROM ROCKET-BORNE ELECTRON FLUX MEASUREMENTS

It is possiole to validate the technique of inferring altitude profiles of ionization from space-based electron flux measurements. This is done by comparing these profiles with simultaneous measurements of the ionosphere made with another sensor. Such comparative "input-output" experiments have been performed many times with the incoherent-scatter radar facility at Chatanika, Alaska (Leadabrand et al., 1972) and rocket-borne detectors flown from the nearby Poker Flat Rocket Range.

As an example, ionospheric conditions during a coordinated rocket/radar experiment are shown in Figure 1. On 9 March 1978 a Terrier-Malemute sounding rocket was launched from Poker Flat over a

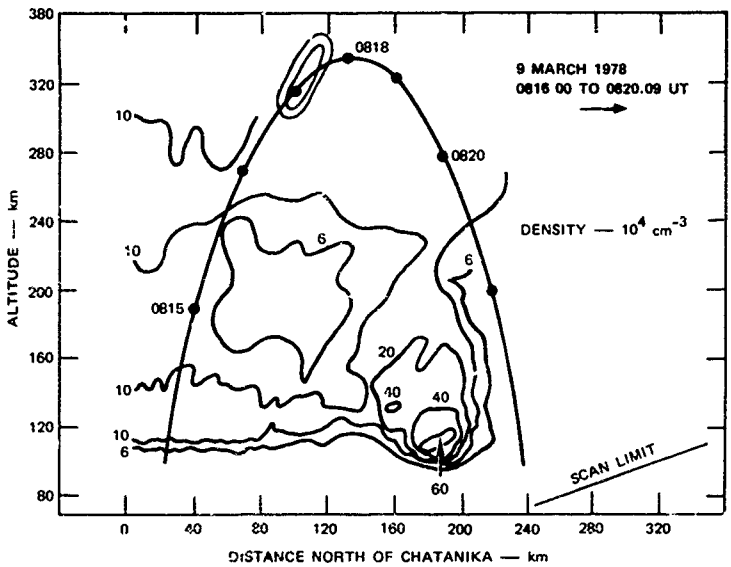


FIGURE 1 CONTOUR MAP OF IONOSPHERIC ELECTRON DENSITY MADE WITH THE CHATANIKA RADAR ON 9 MARCH 1978. The contours are in units of 10^4 cm^{-3} The rocket trajectory is also indicated

stable, premidnight auroral arc. As part of the Rice University scientific payload, it carried an array of 11 electrostatic analyzers and two Geiger-Muller tubes to measure the electron flux in 13 energy intervals from 0.04 to > 40 keV (Pulliam et al., 1981). The trajectory of the rocket is shown in Figure 1, as well as isodensity contours of the altitude and latitude of ionization measured simultaneously by the Chatanika radar. The radar was operated in an elevation-scan mode (Vondrak and Baron, 1976) so as to measure the electron density over more than 300 km of latitudinal extent in the E-region with a spatial resolution of about 10 km. Conspicuous in the ionization contours are the diffuse aurora overhead and to the north of Chatanika with E-region maxima of about 1×10^5 cm^{-3} at an altitude of 120 km, and the intense arc in the north with a maximum density of more than 6×10^5 cm^{-3} at 110 km altitude. The arc is imbedded in a region of enhanced ionization about 40 km wide and marks the northern boundary of any E-region ionization greater than 6×10^4 that was within the radar field of view. These measurements were made in the geomagnetic meridian plane, and represent a two-dimensional cross-section of the auroral features that extended in an east-west direction.

As can be seen in Figure 1, the rocket was above an altitude of 200 km from about 40 km north of Chatanika to 220 km north. At this altitude the incident kilovolt electrons are nearly unattenuated by atmospheric absorption and scattering. These rocket measurements of electron flux have been used as an input to an energy deposition code to compute the E-region ionization distribution below the rocket trajectory. Four-second averages of the particle spectra were used, resulting in a spatial resolution of about 10 km in latitude. Isodensity contours drawn from these profiles are shown in Figure 2. For comparison we show in Figure 3 the Chatanika radar measurements for the same region on a comparable spatial resolution. It can be seen that the principal features are present in both isodensity cross-sections, with good agreement both in general morphology and absolute values. To illustrate this further, we show in Figure 4 the inferred and measured profiles at two locations: within the diffuse aurora and within the auroral arc. It can be seen that they agree to within about 20%, an agreement typical of all the comparisons made during this experiment.

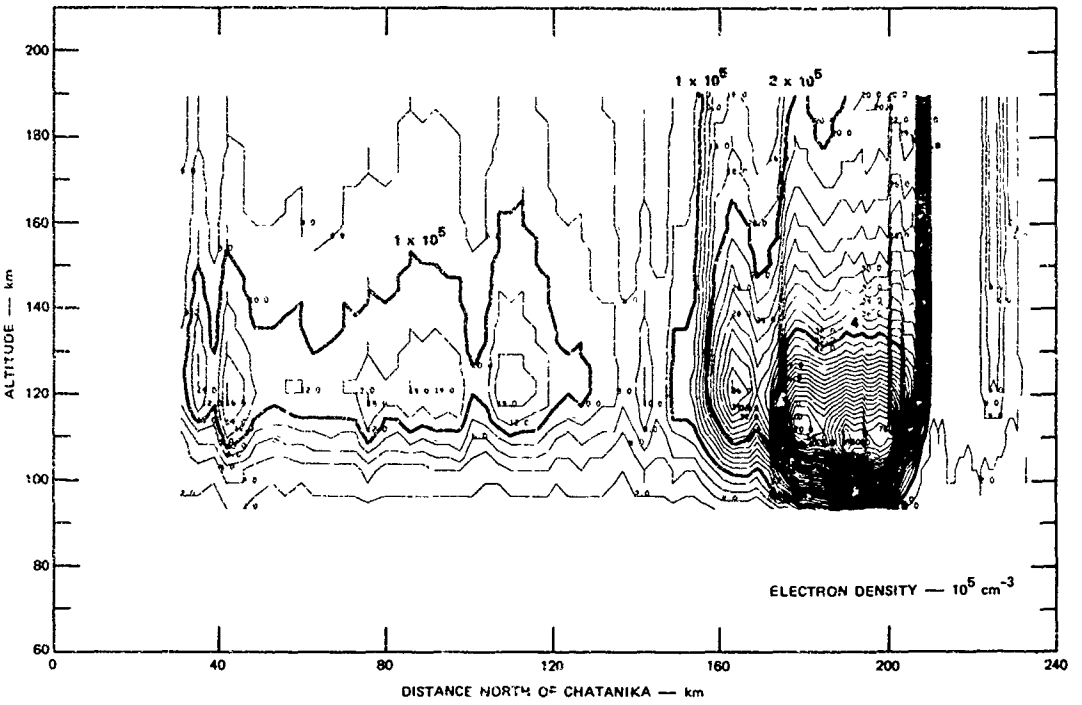


FIGURE 2 CONTOUR MAP OF IONOSPHERIC ELECTRON DENSITY INFERRED FROM KILOVOLT ELECTRON FLUXES MEASURED ON THE RICE UNIVERSITY ROCKET PAYLOAD

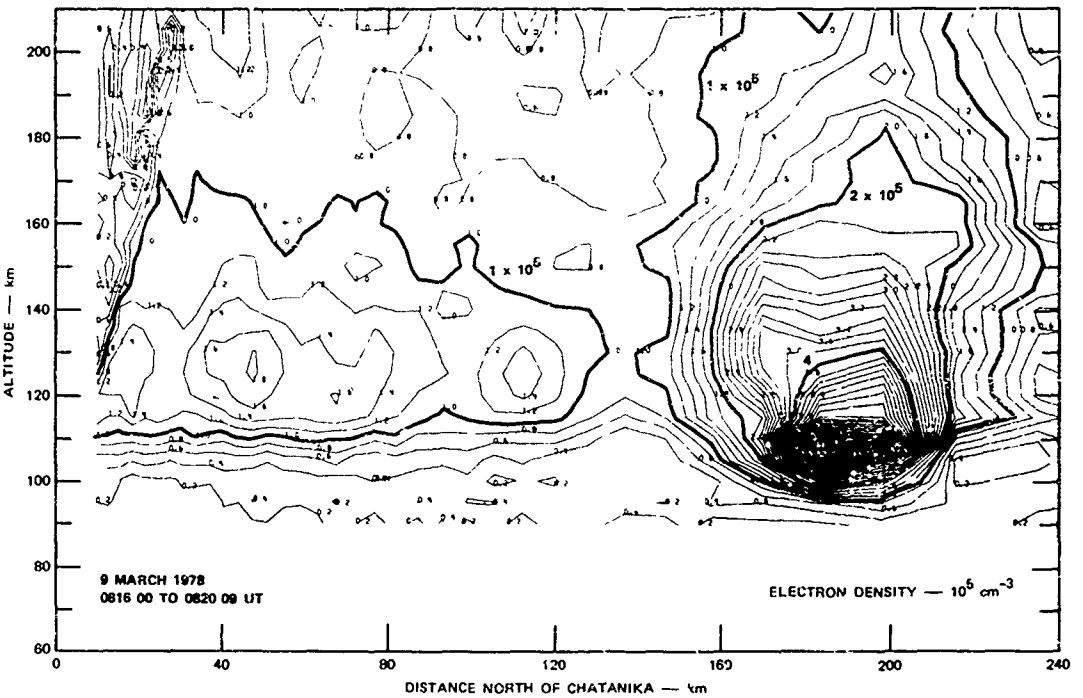


FIGURE 3 CONTOUR MAP OF CHATANIKA MEASUREMENTS OF IONOSPHERIC ELECTRON DENSITY SHOWN ON THE SAME SCALE AS FIGURE 2

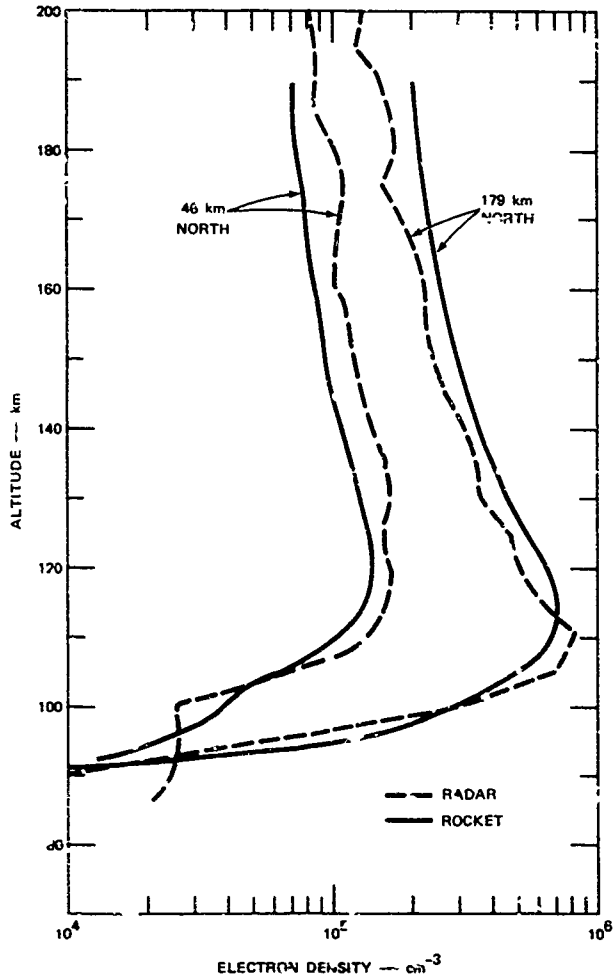


FIGURE 4 MEASURED (RADAR) AND INFERRED (ROCKET) ALTITUDE PROFILES OF IONIZATION DURING THE ROCKET EXPERIMENT ON 9 MARCH 1978. The profile at 46 km north is in the diffuse aurora; that at 179 km north is in the bright arc.

IONOSPHERIC PROFILES INFERRED FROM GROUND-BASED PHOTOMETER MEASUREMENTS

Spectrophotometric measurements can be used to deduce the characteristics of precipitating electrons with either ground-based or space-based sensors. Measurements made at Chatanika have shown their utility for inference of total electron energy input (Wickwar et al., 1975) and the mean energy of precipitating electrons (Vondrak and Sears, 1977; Sears and Vondrak, 1981). Topside spectrophotometric measurements made by the ISIS-2 satellite have been used to deduce total energy input (Murphree and Anger, 1978). Implicit in the measurements that have been presented by Vondrak and Sears (1977) is the information needed to generate ionization profiles with a particle-atmosphere code. These measurements were made during an early-morning fairly intense diffuse aurora on 20 February 1976, with both the incoherent-scatter radar and multiwavelength photometers measuring in the geomagnetic zenith. The photometric measurements at 6300 Å and 4278 Å were used to infer the electron differential energy distribution under the assumption of a Maxwellian beam (Rees and Luckey, 1974). Measurements were made with a time resolution between 10 and 20 seconds.

Examples of two profiles measured during the period described by Vondrak and Stars (1977) are shown in Figures 5 and 6. In the first profile the photometer inferred a characteristic energy of 2.1 keV and a total energy flux of 2.4 ergs/cm^2-s. This is shown to agree excellently with the simultaneously measured Chatanika profile. The only major disagreement is below 95 km altitude. In the lower E-region the ionization is often due to non-Maxwellian sources (Nagata et al., 1975; Vondrak et al., 1980).

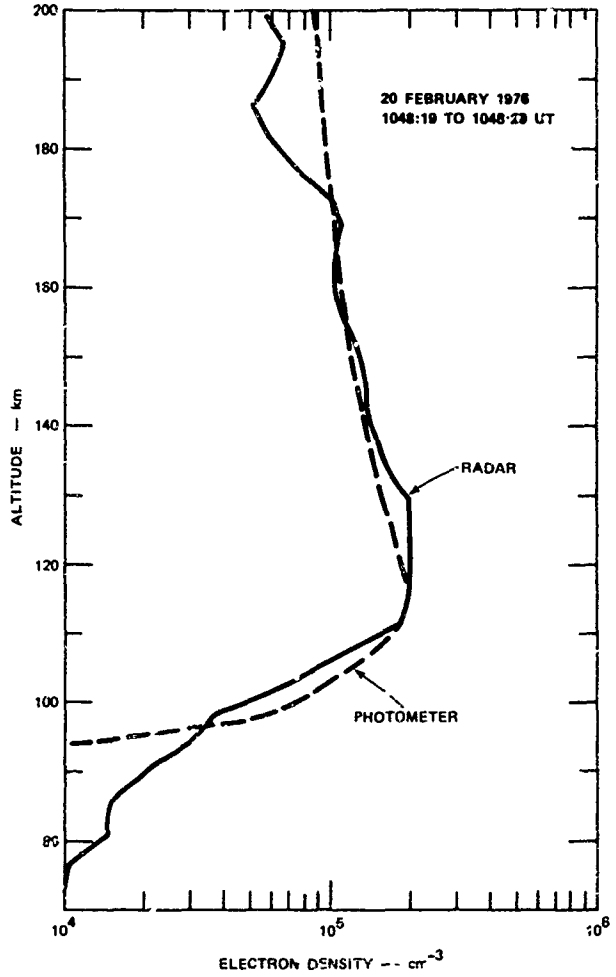


FIGURE 5 IONIZATION PROFILES MEASURED BY THE CHATANIKA RADAR AND INFERRED FROM PHOTOMETRIC OBSERVATIONS BETWEEN 1048:19 AND 1048:29 ON 20 FEBRUARY 1976

A profile comparison from this period that does not agree as well is shown in Figure 6. Here the photometer inferred an incoming electron flux of about 11 erg/cm^2-s with a characteristic energy of 4.6 keV. This produces an E-region maximum of about 4×10^5 cm^{-3} at 105 km altitude. The measured profile had a similar maximum, but at an altitude of about 95 km. The Maxwellian profile that is the best fit to the radar measurements is also shown in Figure 6. This distribution has a characteristic energy of 12 keV. However, it fails to give a perfect fit at altitudes above and below the peak. This disagreement indicates that the incoming electron energy distribution at this time was probably not a simple Maxwellian.

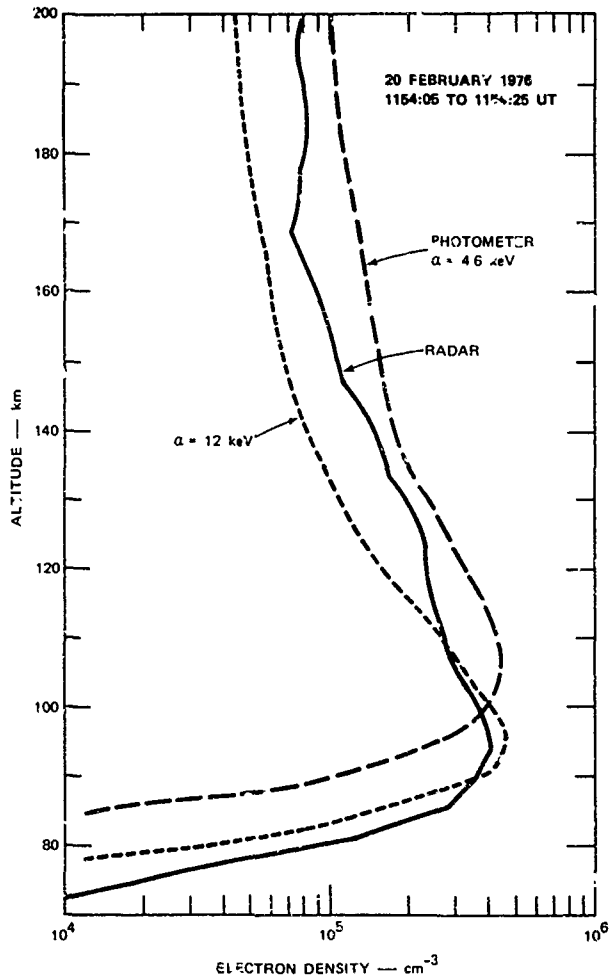


FIGURE 6 IONIZATION PROFILES MEASURED BY THE CHATANIKA RADAR AND INFERRED FROM PHOTOMETRIC OBSERVATIONS BETWEEN 1154.05 AND 1154 25 ON 20 FEBRUARY 1976

The comparisons in Figures 5 and 6 also illustrate the interesting point that the E-region maximum density is generally proportional to the square root of the incident electron energy flux, and only weakly dependent on the characteristic energy. The principal effect of changing the characteristic energy is to alter the altitude of the E-region maximum. Because comparisons have shown that the energy characteristics can be obtained to within 20% to 30% by photometric means (Wickwar et al., 1975; Vondrak and Sears, 1977), the key E-region parameters can be deduced with comparable accuracy.

DISCUSSION

The data comparisons shown in this paper indicate that passive remote sensing can be used to deduce with reasonable accuracy the key parameters of the high-latitude E-region. At low altitudes (below 80 km), the main uncertainties are inadequate knowledge of the effective recombination coefficient. At higher altitudes (above 180 km), transport of ionization by the convection electric field is often comparable to the local production term in the electron continuity equation. For these reasons these techniques have general utility only within the altitude interval of 80 to 180 km.

The accuracy of these techniques has been validated in only a preliminary way. The overall accuracy depends on the combined uncertainty of measurements of the incident electron distribution and of the particle-ionosphere codes. The individual uncertainty of each of these is probably 20% to 50% depending on quality of instrumentation and auroral conditions. The optical techniques require in addition the deduction of the incident electron energy spectrum. The accuracy of this is uncertain, but experience indicates combined statistical and systematic accuracies of about 20% to 30%.

Application of these techniques to operational spaceborne particle detector systems can yield the two-dimensional (altitude, latitude) cross-sections of E-region ionization beneath the satellite trajectory. Data from spectrophotometric imaging systems can, in principle, be used to specify the instantaneous three-dimensional distribution of high-latitude nighttime ionization. Finally, if techniques are devised for the reliable interpretation of X-ray or ultraviolet measurements, such ionospheric specification can be extended to the dayside.

ACKNOWLEDGMENTS

I am grateful for the contributions of my colleagues at SRI International, Lockheed Palo Alto Research Laboratories, and Rice University. In particular, I wish to thank R. D. Sears of LPARL and H. R. Anderson of Rice University for the use of their data in this paper. M. McCready and R. Robinson provided valuable assistance in the analysis of the radar and rocket measurements. Special contributions were made by the site crew at the Chatanika Incoherent Scatter Radar Facility, which is operated by SRI International with the support of National Science Foundation grant ATM-78-23658. Optical measurements were supported by various contracts from the Defense Nuclear Agency; the rocket measurements were supported by National Aeronautics and Space Administration contract NAS6-1667 and NSF grant ATM-78-21761.

REFERENCES

Berger, M. J., S. M. Seltzer, and K. Maeda, "Energy deposition by auroral electrons in the atmosphere," J. Atmos. Terr. Phys., 32, 1015, 1970.

Eather, R. H., "Ionization produced by auroral proton precipitation," Ann. Geophys., 26, 609-613, 1970.

Imhof, W. L., G. H. Nakano, R. G. Johnson and J. B. Reagan, "Satellite observations of bremsstrahlung from widespread energetic electron precipitation events," J. Geophys. Res., 79, 565, 1974.

Leadabrand, R. L., M. J. Baron, J. Petriceks, and H. F. Bates, "Chatanika, Alaska, auroral zone incoherent scatter facility," Radio Sci., 7, 747-756, 1972.

Mizera, P. F., J. G. Luhmann, W. A. Kolasinski, and J. B. Blake, "Correlated observations of auroral arcs, electrons, and X-rays from a DMSP satellite," J. Geophys. Res., 83, 5573, 1978.

Murphree, J. S. and C. D. Anger, "Instantaneous auroral particle energy deposition as determined by optical emissions," Geophys. Res. Letters, 5, 551, 1978.

Nagata, T., T. Hirasawa, M. Takizawa, and T. Tohmatsee. Antarctic substorm events observed by sounding rockets: ionization of the D- and E-region by auroral electrons," Planet. Space Sci., 1321, 1975.

Pulliam, D. M., H. R. Anderson, K. Stamnes and M. H. Rees, "Mechanisms of auroral electron acceleration and atmospheric interaction inferred from rocket-borne observations," submitted to J. Geophys. Res., 1980.

Rees, M. H., "Auroral ionization and excitation by incident energetic electrons," Planet. Space Sci., 11, 1209, 1963.

Rees, M. H. and D. Luckey, "Auroral electron energy derived from ratio of spectroscopic emissions: 1. Model computations," J. Geophys. Res., 79, 5181-5186, 1974.

Sears, R. D. and R. R. Vondrak, "Optical emissions and ionization profiles in an intense pulsating aurora," J. Geophys. Res., in press, 1981.

Vondrak, R. R. and M. J. Baron, "Radar measurements of the latitudinal variation of auroral ionization," Radio Sci., 11, 939, 1976.

Vondrak, R., D. Evans and T. Moore, "Precipitating proton and electron contributions to ionization and conductivity in a midnight diffuse aurora," EOS, 61, 1089, 1980.

Vondrak, R. R. and R. D. Sears, "Comparison of incoherent scatter radar and photometric measurements of the energy distribution of auroral electrons," J. Geophys. Res., 83, 1665, 1977.

Wickwar, V. B., M. J. Baron, and R. D. Sears, "Auroral energy input from energetic electrons and Joule heating at Chatanika," J. Geophys. Res., 80, 4364-4367, 1975.

APPLICATION OF X-RAY IMAGING TECHNIQUES TO AURORAL MONITORING

David M. Rust and Paul Burstein
American Science and Engineering, Inc.
Cambridge, Massachusetts

ABSTRACT

The precipitation of energetic particles into the ionosphere produces bremsstrahlung X-rays and K-alpha line emission from excited oxygen and nitrogen. If viewed from a spacecraft in a highly elliptical polar orbit, this soft (0.3 - 3.0 keV) X-radiation will provide an almost uninterrupted record of dayside and nightside auroras. A grazing incidence X-ray telescope especially designed for such auroral monitoring is described. High photon collection efficiency will permit exposure times of approximately 100 seconds during substorms. Spectrophotometry will allow users to derive the energy spectrum of the precipitating particles. If placed in a 15 R_E orbit, the telescope can produce auroral X-ray images with 30 km resolution. Absolute position of X-ray auroras can be established with a small optical telescope co-aligned with the X-ray telescope. Comparison of X-ray and optical images will establish the height and global distribution of X-ray aurorae, relative to well-known optical auroras, thus melding the new X-ray results with knowledge of optical auroras.

INTRODUCTION

The ISIS-2 and DMSP series of satellites have opened an entirely new era in morphological studies of the aurora, enabling one to scan the global distribution of auroral luminosity (Lui and Anger, 1973; Akasofu, 1976). The introduction of auroral scanning from satellites has greatly improved our understanding of global particle precipitation patterns, although it is not possible not possible to follow developments on the sunlit side of the earth.

The auroral zone is expected to be a strong emitter of X-rays between 0.3 and 3.0 keV. The X-rays are bremsstrahlung emitted directly by electrons precipitating from the magnetosphere and Kα X-ray lines emitted by excited oxygen and nitrogen atoms. With suitable instrumentation, these X-rays may be used to detect auroras on a global scale with no interference by sunlight, clouds or snowfields.

Mizera et al. (1978) placed a 1.4 - 20 keV X-ray scanning collimator aboard a DMSP satellite. The resolution of the instrument was too coarse to isolate emission from individual auroral features, and, because of the 100-minute orbital period, temporal developments could not be distinguished from spatial structure. However, Mizera et al. were able to show that the X-ray emission from auroral regions beneath the satellite agreed well in spectral form and absolute magnitude with the bremsstrahlung expected from precipitating electrons sampled on the satellite.

From current theoretical work and from extrapolations of higher energy observations, it is clear that most of the X-radiation from precipitating particles will be in the 0.3 -3.0 keV band. In this paper we describe an instrumental program for using X-rays in this energy band to make subtly detailed pictures of the auroral zone.

The potential for discovery in X-ray pictures of the earth should not be underrated. Apollo mission photographs of the earth, DMSP mosaics of the auroral zones and Skylab solar images have transformed astrogeophysics. The X-ray emission of the earth may similarily transform magnetospheric physics. In the following sections, we describe some of the scientific problems that such X-ray pictures can help to solve.

SCIENTIFIC BACKGROUND AND PROBLEMS

Scale and Effects of Particle Precipitation

The total particle energy precipitated into the earth's atmosphere is estimated to vary from about 10^{15} ergs/sec in simple quiet auroral arcs, to more than 10^{18} ergs/sec in diffuse and discrete auroral forms during moderate geomagnetic disturbances associated with auroral breakups.

This precipitation energy is distributed over the auroral oval region whose instantaneous size varies between 10^{14} and 10^{16} cm^2 (larger area corresponds to the more active times). These particles have energies from a few tens of eV to ~ 100 keV. Protons and electrons are precipitated in various proportions. These particles primarily ionize the atmospheric constituents and produce optical arcs, changes in ionospheric conductivities, and other thermodynamic effects that affect the meteorology and chemistry of the upper atmosphere and radio signals passing through the auroral zone.

It has turned out to be quite difficult, even with DMSP auroral mosaics from space, to relate the satellite measurements of energetic particles to familiar auroral forms. Precipitating electrons and protons are detected over much wider areas than discrete auroral forms. A basic question to be answered by wide-field imagery is: what corresponds to the auroral arcs in the energetic particle recordings from low orbiting satellites?

Particle acceleration processes operate on a global scale, but the scale width of the acceleration regions is very narrow, and we are still far from understanding what the various auroral arcs are showing us in terms of particle and field distributions. Simultaneous, large-field optical and X-ray observations at ~ 10 kilometers resolution should bring a dramatic improvement inour understanding of global particle acceleration and precipitation.

Direct measurements of precipitating particles on auroral magnetic lines of force habe shown that the particle energy spectra are complicated and variable. The current belief is that the low energy (~ 1 keV) component is a consequence of a parallel electric field that exists at 1-3 R_E from the surface of the earth. The higher energy components show characteristics of the plasma sheet and the outer Van Allen particle population and are thus possible accelerated from sources more distant than the ionosphere. To trace each component from injection and acceleration in the magnetosphere to precipitation in the ionosphere, spectral observations with a large-field remote sensot plus in situ measurements are needed.

Substorms

A substorm represents an instability of the entire magnetosphere during which time energy, in the form of charged particles, is efficiently dissipated into the ionosphere (Akasofu, 1968). The particle precipitation can be viewed as the final process of a sequence of complex events that starts at the plasma source, i.e., the ionosphere and the solar wind. The auroral oval exhibits systematic shifts during substorms. We know the nightside aurora exhibits poleward surges during substorms. The dayside aurora simultaneously moves equatorwards. These are statistical results established by ground-based and satellite optical observations. They do not tell us global details about substorm behavior on the relevant time scales (~ 10 mintues). With global snapshots, we could better understand the large scale magnetosphere distortions and geotail developments that cause substorms. We could learn whether the whole auroral oval shifts sunward in substorms and whether the oval recovers as a unit. Isolated precursors of substorm development could be studied and their behavior could help to explain the substorm triggering mechanism.

Proton Precipitation in Substorms

In general, electrons and protons precipitate together with roughly equal intensity, but during substorms keV proton and electron fluxes may be uncorrelated or anticorrelated on a small scale (Bernstein and Wax, 1970). Existing spacecraft measurements have not allowed us to distinguish between space and time variations in these cases. However, it is clear that the range of variations between keV proton and electron fluxes implies that a wide variety of physical processes are operating simultaneously over fairly narrow ranges in magnetic latitude.

Hultqvist (1974) reported that keV proton fluxes vary by much less in latitude than keV electron fluxes. In a strong substorm, the proton flux will vary by less than a factor of two on the nightside, while the electron flux may increase by more than an order of magnitude. However, on the dayside, the proton flux may vary greatly, especially in the polar cusp. Even in quiet periods, there are large variations in proton flux in the cusp. Perhaps the precipitation events which are predominantly protons are telling us that the magnetospheric electron population has already been depleted by a process which acts most efficiently on electrons. One clue that supports this interpretation is the correlation between almost pure keV proton precipitation on the dayside and Kp. Most proton aurorae are seen at high Kp values.

What is needed to investigate the electron depletion model of pure proton auroras is monitoring of the total energy in precipitating electron and the size of the electron precipitation zone during disturbed times on both the day and night sides.

Polar Cusp Phemomena

Dayside aurorae have received increased attention primarily because of the broad magnetospheric implication of their association with the polar cusp. Studies of this region relate to the fundamental questions of whether the magnetosphere is open or closed, how plasma enters the magnetosphere from the solar wind, and how magnetospheric current systems are configured.

Of particular relevance to the above questions is the response of the dayside aurora to substorms and interplanetary magnetic field (IMF) variations. Equatorward movements of dayside aurora can be regarded as indicating erosion of the dayside magnetosphere to the tail. Customarily this erosion process is seen as a consequence of increased merging of interplanetary and magnetospheric field lines. Alternatively, dayside auroral movements could result from internal reconfiguration of large-scale magnetospheric current systems.

There have been a number of investigations of the equatorial shift of dayside aurorae in association with substorms, but there are disagreements on the magnitude of such shifts and timing of the shifts with respect to substorm onset. There have also been many investigations of the relation of dayside aurora or cusp position to the direction and magnitude of the interplanetary magnetic field. These studies have been statistical in nature, often involving long-term (~ 45 min) averaging of B_z data and low time resolution (~ 90 min) satellite data on cusp position. Consequently one may question whether claimed correlations are direct results of merging (dayside field erosion), or result less directly via the substorm process. Until global observations are available, the question of the meaning of the polar cusp movements will probably remain unanswered.

SAR Arcs and Proton Precipitation

Stable auroral red (SAR) arcs are broad regions of predominently 6300 Å emission that appear during the rapid recovery phase of magnetic storms. Their 6300 Å intensity is almost constant for ~ 10 hours. The emission is probably due to excitation by thermal electrons (E ~ 1 eV) rather than by direct energetic particle bombardment (Cole, 1965). Nevertheless, SAR arcs may be a valuable diagnostic of energy transfer to the upper atmosphere from 10 - 100 keV ring current protons.

Cornwall et al. (1970) have advanced a recovery phase theory of SAR arc formation in which ring current protons dissipate most of their energy into ion cyclotron wave turbulence and proton precipitation losses. Cornwall et al. suggest that the energy-draining turbulence occurs when the plasma-pause expands into the stable ring current proton belt. In the ensuing precipitation process, the protons lose about half their energy to ion cyclotron waves. The waves, in turn, are absorbed by thermal electrons which conduct heat to the ionosphere and excite SAR arcs on the lines of force connected to the destabilized proton region. Because of the very low energies involved, no X-ray emission should be excited by the heated electrons.

The precipitating protons, whose energies are in the 5 - 50 keV range, should give rise to K-line X-rays in addition to hydrogen Balmer line emission. Weak hydrogen emission has been detected in association with SAR arcs (Kleckner and Hoch, 1973). The protons should precipitate along a narrow region on lines of force just inside the plasmapause. The resultant hydrogen emission is frequently on the threshold of sensitivity for optical band experiments.

Although the expected proton fluxes are low, they may efficiently excite hitherto unobserved Kα X-radiation near SAR arcs. Since there is little likelihood of precipitating electron contamination in SAR arcs, any Kα line emission would be a tracer of proton precipitation. A comparative study of 6300 Å SAR arc position with K-line emission arcs would be very useful, therefore, in tracing the location and dynamics of the plasma-pause as it expands into the ring current proton belt.

Albedo X-Rays and Atmospheric Density

The only published observations of the earth's radiation at 0.3 - 3.0 keV were made by Rugge et al. (1978) with a proportional counter on HEAO-1. They detected only solar albedo X-rays because their field of view was confined to the equatorial region. The count rates recorded during quiet solar times were an order of magnitude lower than the flux expected from a moderate aurora. However, as Rugge et al. point out, the albedo flux is dependent upon the integrated

atmospheric density above the reflection level. This level is at approximately 100 km, so the atmosphere above that level could be monitored around the northern hemisphere as a function of solar activity by a sensitive X-ray instrument.

In the next section we describe an instrument that will be sufficiently sensitive to address the problems raised in the preceeding paragraphs.

EXPERIMENTAL APPROACH

X-Ray Imaging

Over the past twenty years, there have been major improvements in the ability to image sources of soft X-ray emission using the techniques of grazing incidence optics. Focussing soft X-ray telescopes have been flown on sounding rockets, small satellites, manned spacecraft, and (most recently) on the High Energy Astrophysical Observatory (Miller et al. 1977). New instruments are being designed now for sub-arc second resolution (Davis et al. 1979).

Table I. Auroral X-Ray Telescope Specifications

X-ray optical system	Wolter Type I grazing incidence telescope
Collecting area	94 cm^2 @ 31.6 Å
Wavelength range	4 - 40 Å (0.3 - 3.0 keV)
Energy resolution	170 eV
Visible-band imager	Doublet lens, coaxial with X-ray optics
Collecting area	2.85 cm^2
Wavelength range	6300 Å (OI) and 4278 Å (N_2^+)
Field of view	1.25 x 1.5 degrees
Detector	Back-illuminated, thinned, charge-coupled device
Detector dark current	< 1 electron/sec/pixel (at -60°C)
Preamp sensitivity	20 photoelectrons rms
Dynamic range	> 10^3
Signal integration time	10 - 1000 sec nominal
Resolution (4 x 4 pixels)	1.2 arc min (30 km at 15 R_E)
Viewing requirements	Up to 15° offset from earth center
Filters	7 band pass filters
Calibration	Radioactive source on filter wheel
Size	26 x 28 x 63 cm
Weight	20.12 kg
Thermal control	Passive and active (± 5° C)
Memory	700 K bytes
Telemetry/storage rate	5 Kbps average
Power	16.2 watts

Parameters of an auroral X-ray telescope (AXT) experiment are summarized in Table I. The energy range covered is from 300 eV to 3.0 keV. The low energy cutoff is set by a filter which excludes visible and XUV radiation. The filter passes the Kα line emission from oxygen and nitrogen. Other filters attenuate Kα radiation to lessen pulse pile-up, which can interfere with measurements at the high energy end of the spectrum. The upper energy limit is set by the grazing angle and nickel coating of the mirror and is the energy at which the reflection coefficient goes to zero.

The striking appearance of X-ray emitting astronomical bodies is best illustrated by Figure 1, which shows the solar corona in the 0.3 - 3.0 keV energy band of the AXT. The smallest features seen in the solar X-ray image are about one two-hundredth of the solar radius. Coincidentally, the smallest features that the proposed auroral X-ray telescope will resolve at 15 R_E are about one two-hundredth of the earth's radius. Figure 1 may suggest how the X-ray emitting earth will appear.

X-ray imaging will open up many new possibilities in auroral research. X-ray aurora can be detected on the dayside and nightside hemispheres. X-rays can be used for quantitative study of

the mid-latitude dayside aurorae, which occur only during very large magnetic disturbances and are difficult to detect optically. Daytime and nightime injection can be studied at any time of the year without interference from sunlight, the moon, or clouds.

X-ray imaging will remove all of the limitations encountered heretofore in remote sensing. For example, the overlaid rectangle on Figure 2 shows the large instantaneous field of view of the AXT. The small block, corresponding to ~ 30 x 30 km, shows that the resolution of the instrument at 15 R_E is high enough to produce images of DMSP quality.

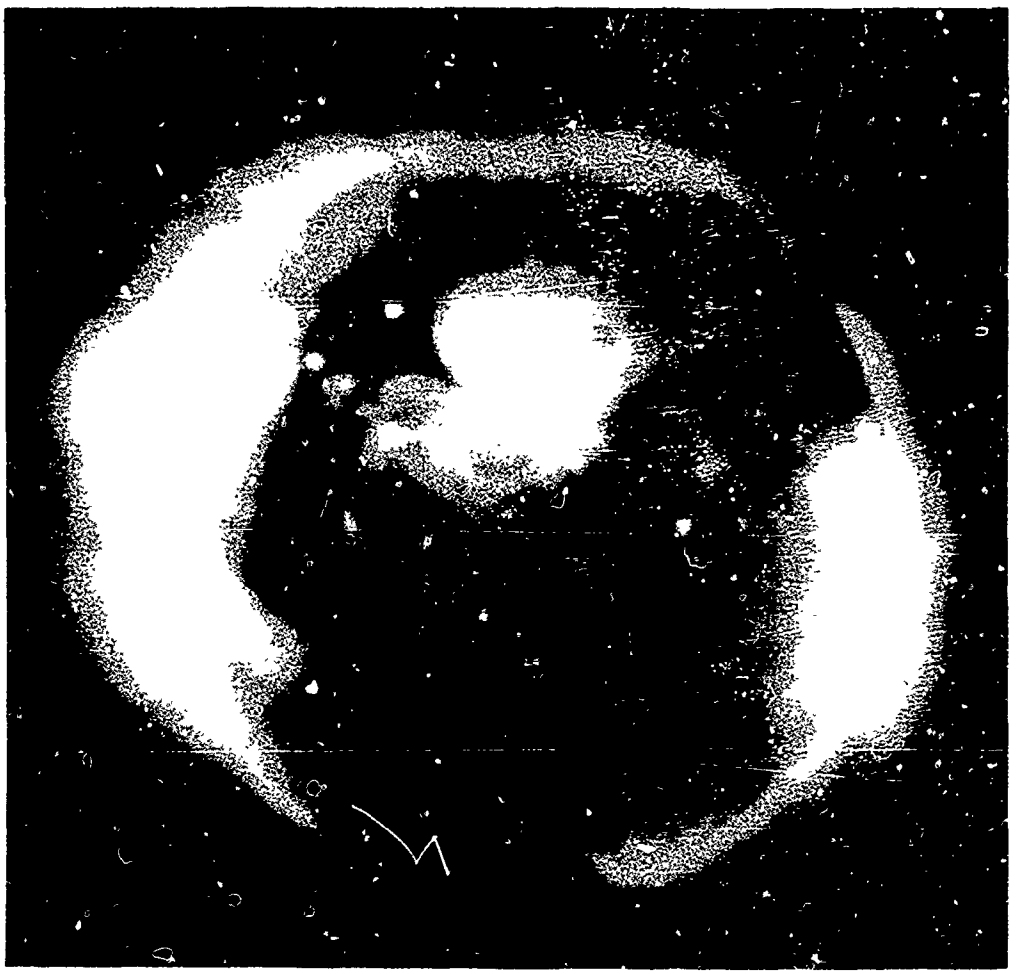


Figure 1. A soft X-ray image of the solar corona obtained at the time of June 30, 1973 eclipse by the AS&E telescope experiment (S-054) on Skylab.

Figure 2. The auroral oval, from a mosaic of DMSP images. The field of view of the AXT, from 15 R_E is outlined in narrow white lines.

X-Ray Detection

Charge-coupled devices (CCDs) are monolithic arrays of solid-state photodetectors originally developed for detecting visible light images. We at AS&E and others have already used CCDs to detect X-rays in much the same way as one uses photographic film, as a total energy detector (Rhenda and Lowrance, 1975; Burstein et al., 1978). However, the capabilities of the CCD have recently been extended to detecting individual soft X-ray photons (Burstein et al., 1978; Schwartz et al., 1979). Figure 3 shows that the spectral resolution between 0.3 and 3.0 keV will be better than that of a proportional counter when the CCD intrinsic noise is ~ 20 electrons rms.

The CCD can be used either for single photon energy detection or as a total energy detector. As a total energy detector the CCD is used exactly as for visible light detection. The output of a particular picture element ("pixel") is a charge which is proportional to the total amount of energy deposited in the pixel. If the intrinsic noise level associated with a pixel is small compared to the signal, then that signal will be recognizable above background. Individual photons of visible light do not liberate sufficient charge in a CCD to be detected as such. X-ray photons are much more energetic and can be detected individually.

If only one X-ray strikes a pixel, then the charge in that pixel will be a function of the photon energy, as in a solid-state detector. An X-ray photon interacting in the bulk silicon of the CCD will create one electron-hole pair for each 3.6 eV of its energy. As long as the probability of any single pixel receiving more than one photon during an integration period is small, the CCD may be used as a non-dispersive X-ray spectrometer. The charge created by each X-ray is stored in the pixel where the interaction occurred until the device is read out. The pixel address is the position, and thus the device is a position-sensitive X-ray spectrometer. The data can be sorted into as many energy intervals as desired, consistent with resolution.

Interpretation of X-Ray Spectra

We believe that X-rays from aurorae will be easier to interpret than optical or UV emissions since bremsstrahlung is generated by the primary precipitating electrons themselves, while optical and UV line emissions result primarily from secondary electron interactions with many atomic and molecular species distributed in a variable and sometimes poorly understood way along the precipitation path.

Expected Count-Rate and Background

Auroral Emissions

On the basis of observations by Mizera et al. (1978) and the theoretical work of Luhmann and Blake (1977), one can estimate the count rates expected for various auroras. Since the aurora is an extended source, that is, each pixel is fully illuminated by auroral structure, the flux F on one pixel is:

$$F = A(\lambda)\ I\ \frac{d^2}{f^2}$$

where $A(\lambda)$ is the effective collecting area of the telescope. $A(\lambda)$ includes the effect of 2000Å of aluminum prefilters and the response curves of the telescope and CCD. $A(\lambda) \sim 30$ cm^2 at ~ 0.5 keV (see Table I). I is the X-ray flux in photons cm^{-2} sec^{-1} keV^{-1}. We take $I \approx 5 \times 10^5$ photons cm^{-2} sec^{-1} keV^{-1} sterad^{-1} in a moderate aurora.

The pixel dimension d is 30×10^{-4} cm and the telescope focal length is 35 cm (ref. Table 1). Then, the count rate in 170 eV bands at ~ 0.5 keV will be ~ 0.4 counts per second (see Figure 4) per "macropixel", which is a 4 x 4 block of pixels averaged during readout of the CCD.

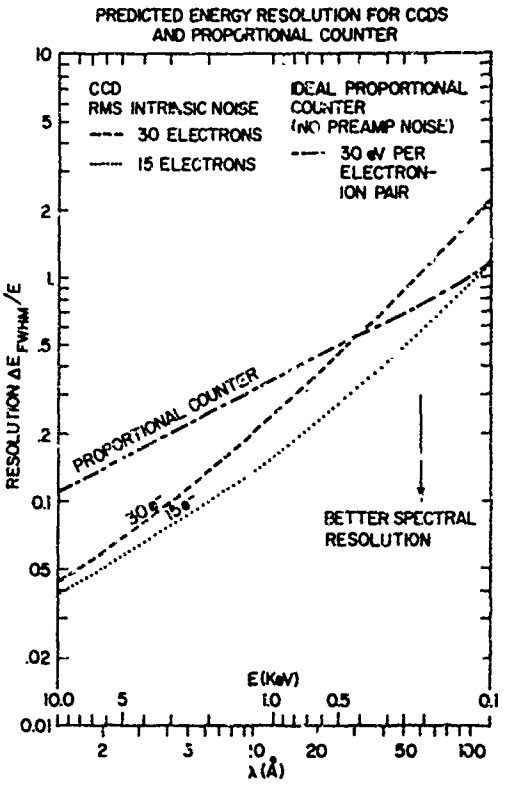


Figure 3. Predicted energy resolution for CCDs and a proportional counter.

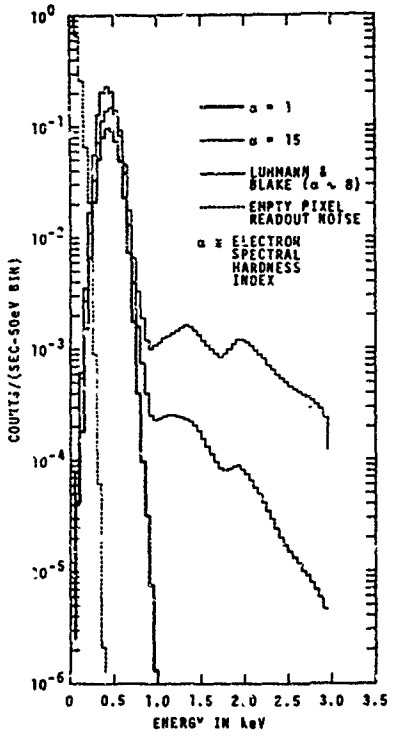


Figure 4. Expected pulse height distribution for various input spectra. The total AXT response including telescope and filter transmission curves has been taken into account. Detector response and pulse height spreading are also included.

The above estimate applies to moderate aurorae which occur about 10 days per month. Brighter aurorae, with 10 times this flux, are not uncommon. In a moderate aurora, 400 integrations (100 seconds) should yeild 40 counts in the Kα lines. Longer exposures or larger macropixels will be required to obtain three-sigma count levels at 1-3 keV. Integration times will vary between 10 seconds over the brightest aurorae and 1000 seconds, or more, in faint auroras. Integration times will be determined either by explicit command sequences, or by the total signal level, as recorded continuously in a separate register. In bright aurorae, the integration time should be lessened because the brightest features also are the most rapidly changing.

Expected Pulse-Height Distribution

Figure 4 shows the expected response of the AXT to two X-ray spectra from idealized electron beams with α of 1 and 15 (α is the index of spectral hardness). Also shown is the AXT response to the Frank and Ackerson electron spectrum analyzed by Luhmann and Blake (1977). We estimate that α = 8 in this case. While the K-lines dominate the spectra, it is clear that the 1 - 3 keV response characteristics will differentiate among these spectra. Time integrals of 100 - 1000 seconds over this band easily yield >3σ results.

The minimum detectable flux (3σ) is of the order of 10^4 photons $sec^{-1}cm^{-2}keV^{-1}sterad^{-1}$ (including the K-lines) over a time scale of ~ 1000 seconds on the dark side of the earth. The dayside minimum is higher because of the earth's solar X-ray albedo, as described below.

For any moderate aurora the AXT can give a reliable estimate of the index of the electron spectrum and the total electron energy. In general, the strength of the K-lines relative to the continuum will yield information on the proton/electron ratio for any event.

X-Ray Background

X-ray noise is contamination from non-auroral sources which either mask or overwhelm the auroral X-rays. These X-ray sources are:

- o solar X-rays scattered or reflected from the atmosphere
- o cosmic ray induced X-rays in the upper atmosphere
- o interaction of charged particles with the various parts of the telescope which result in an X-ray striking the detector
- o celestial X-rays scattered from the atmosphere.

With the exception of the first item the rest of the list may be safely ignored since their contributions are negligible. The solar X-ray albedo (Rugge et al., 1978) merits further attention because it provides a background which may contribute as much as 10% to the auroral X-ray signal.

The non-flaring range of solar X-ray fluxes is 10^{-3} to 10^{-2} ergs/cm^{2}sec over the 8 - 20 Å band and 0.03 - 0.1 ergs/cm^{2}sec over the 44 - 60 Å band. For each wavelength interval the ratio of cross-sections for scattering σ_{sc} and absorption σ_{abs} yields the albedo flux. At wavelengths greater than the 31 Å K-edge of nitrogen, the solar X-ray albedo is high because σ_{abs} decreases there. Using the cross-sections of Hayakawa et al. (1971), the expected albedo fluxes are in the 10^{-9} - 10^{-8} erg/cm^2-s-sr-Å range. This yields a detected flux of 0.01 photons/sec-macropixel. This level is significant only in comparison to weak auroras. At shorter wavelengths, the auroral flux will be well above the solar albedo.

INSTRUMENTATION

Grazing Incidence X-Ray Telescope

To record auroral X-ray spectra we plan to use techniques recently devised at AS&E (Burstein et al., 1978; Burstein, 1979) for using a CCD as a non-dispersive imaging X-ray spectrometer. An X-ray telescope consisting of seven confocal mirror pairs will provide the required collecting area.

The AXT (see Figure 5) consists of the following major components: A 35-cm focal length Wolter Type I grazing incidence X-ray mirror assembly, a coaxial visible aurora imager, a CCD focal plane camera with a radiative cooler system, a ten position filter wheel/ shutter. a non-closable contamination control door, and video processing, control and telemetry electronics.

The X-ray optical system is similar to others successfully flown on many missions. However, a thorough design study was untertaken to devise a telescope especially suited to this application. The relatively low fluxes expected and the relaxed resolution requirements led us to the current

design which features very high collecting area, short focal length and less stringent (therefore, cost-reducing) requ rements ror surface finish and alignment. Each of the seven nested concentric mirror pairs has a superpolished nickel-plated beryllium parabolic forward element followed by a hyperboloidal element with the same surface finish. Mirrors are figured by the single point diamond turning technique. Initial polishing is done conventionally; the ultimate surface finish of approximately 15 Å RMS is achieved by a proprietary superpolish process.

In front of the X-ray mirror aperture is a two-layer thermal prefilter each layer consisting of 1000 Å of pure aluminum on a nickel support mesh. These filters serve three major purposes: they prevent contaminants from entering the telescope through the forward apertures, they filter out unwanted white light, and they serve as a thermal barrier to prevent heat loss through the forward apertures.

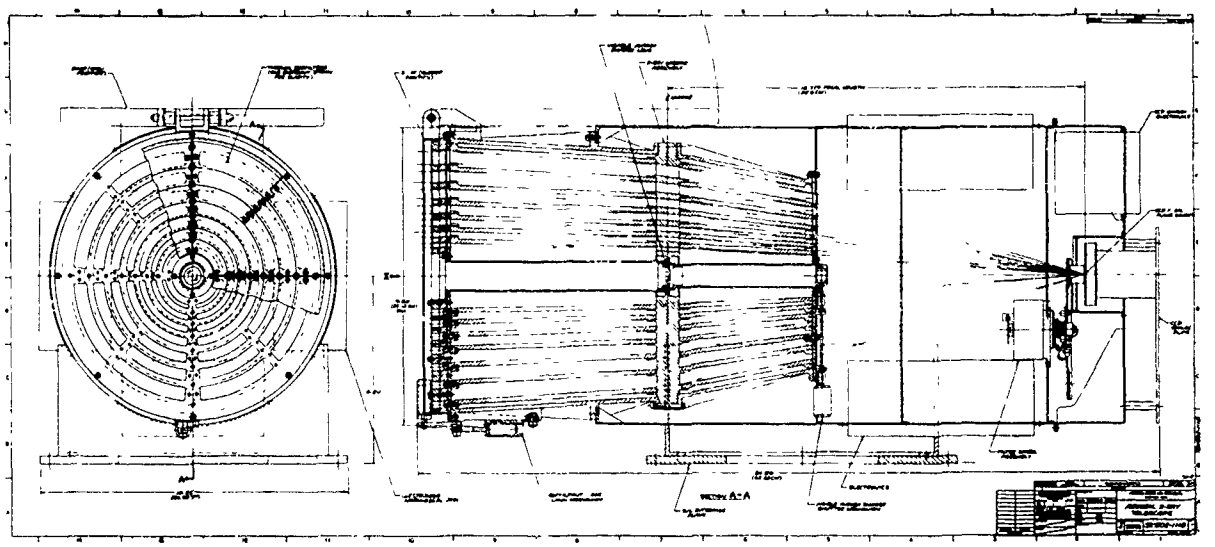


Figure 5. Engineering layout of the AXT.

CCD Camera

The heart of the camera system is a buried-channel 256 x 320 picture element ("pixel") RCA CCD at the focus of the telescope. By suitable clocking, the CCD pixels are combined on the chip into "macropixels". In normal operation, the instrument's 10 arcsec spatial resolution is achieved by macropixels consisting of 4 x 4 CCD pixels. Larger macropixels can be obtained on command when desired to reduce telemetry traffic or for other reasons when the resulting loss in spatial resolution is acceptable. It is possible to provide a selection of macropixel sizes of 4 x 4, 5 x 5, and 6 x 6 CCD pixels, though pixel configurations ultimately available may be rectangular as well as square. The CCD integrates photons for an exposure period sufficiently short to make the probability of multiple photon incidences into a macropixel small. The time required to expose and read out a 64 x 80 macropixel frame is approximately 0.25 seconds.

The use of macropixels is an important feature of the proposed design. It reduces to a minimum the proportion of X-ray photons that divide their energy among two or more locations, while still achieving the required spatial resolution of the instrument. X-ray photons that divide their energy among two or more locations contribute to the apparent low-energy background and add a low energy "tail" to the apparent energy distribution of incident X-rays. The use of macropixels reduces this effect to the point where it ceases to be a concern.

Pointing

To achieve the desired scientific objectives it is required that the AXT-carrying spacecraft have a despun platform with relatively precise pointing, knowledge and stability. The knowledge requirement is for 6 arc min accuracy. The stability requirement is for 30 arc-seconds or better over a period of 100 seconds. A pointing accuracy of 0.1 degrees is sufficient for the instrument.

CONCLUSION

We have discussed an instrumental program to measure X-radiation due to the precipitation of energetic particles into the earth's ionosphere. Bremsstrahlung X-rays come directly from the primary precipitating electrons; Kα lines arise almost exclusively from excitation by protons and the primary electrons. Therefore, the auroral soft X-ray spectrum (0.3 - 3.0 keV) will provide an easily interpreted, remote measure of particle precipitation.

Instrumentation consists of a grazing incidence X-ray telescope especially designed for auroral research. The high collection efficiency of the proposed telescope will permit exposure times during substorms of approximately 100 sec. The detector, a highly efficient charge-coupled device, will allow non-dispersive X-ray spectroscopy in either the day or night hemisphere with spatial and spectral resolutions high enough to investigate substorm structure, stable auroral red (SAR) arcs, polar cusp precipitation, and polar processes thought to occur in the magnetosphere.

Absolute positions of X-ray emitting features will be established by comparisons with optical features. For this purpose we have included a small optical imager in the X-ray telescope. The imager uses the same detector as the X-ray telescope and operates at 6300Å (O I) and 4278Å (N_2+). The optical imager will be precisely aligned with the X-ray imager and will allow timely comparisons with ground-based observations. It will also establish the height and global distribution of X-ray aurorae relative to well-known optical aurorae, thus melding the new X-ray results with established auroral knowledge.

ACKNOWLEDGEMENTS

We are grateful to our colleagues R. Mastronardi, S. Kahler, A. Krieger and K. Kubierschky at American Science and Engineering for their advice during many parts of this study. George Parks (University of Washington) and Robert Eather (Boston College) helped with identification of the scientific problems to which the AXT could be applied. The work was partially supported by NASA contract NASW-3426.

REFERENCES

Akasofu, S.-I., "Polar and Magnetospheric Substorms," D. Reidel, Dordrecht, Holland, 1968.
Akasofu, S.-I., Space Sci. Rev., 19, 169, 1976.
Akasofu, S.-I., Astrophys. and Space Sci. Lib., 47, 1977.
Bernstein, W. and R.L. Wax, J. Geophys. Res., 75, 3915, 1970.
Burstein, P., A. Krieger, M. Vanderhill, and R. Wattson, Proc. SPIE, 143, 1978.
Burstein, P., A. Krieger, D. Harrison and P. Gauthier, Ann NY Acadm. Sci., 342, June, 1980.
Cole, K.D., J. Geophys. Res., 70, 1689, 1965.
Cornwall, J.M., F.V. Coroniti and R.M. Thorne, J. Geophys. Res., 76, 4428, 1971.
Davis, J.M., A.S. Krieger, J.K. Silk and R.C. Chase, Proc. SPIE, 184, 96, 1979.
Eather, R.H., Rev. of Geophys., 5, 207, 1967.
Hayakawa, D., T. Kato, F. Makino, H. Ogawa, Y. Tanaka and K. Yamashita, Proc. Res. Inst. Atmospherics, Nagoya Univ., 18, 1, 1971.
Hultqvist, B., Ann. Geophys., 30, 223, 1974.
Kleckner, E.W. and R.J. Hoch, J. Geophys. Res., 78, 1187, 1973.
Luhmann, J.G. and J.B. Blake, J. Atmos. Terr. Phys, 39, 913, 1977.
Liu, A.T.Y. and C.D. Anger, Planet. Space Sci., 21, 799, 1973.
Miller, R., G. Austin, D. Koch, N. Jagoda, T. Kirchner and R. Dias, IEEE 1977, Nuclear Sci. Symp., 1977.
Mizera, P.F., J.G. Luhmann, W.A. Kolasinski, and J.B. Blake, J. Geophys. Res., 83, 5573, 1978.
Renda, G. and J. Lowrence, JPL SP43-21, 91, 1975.
Rugge, H.R., D.L. McKenzie and P.A. Charles, Space Research XIX, 243, 1978.
Schwartz, D.A., R.E. Griffiths, S.S. Murray, M.V. Zombeck, and W. Bradley, Proc. SPIE, 184, 247, 1979.

IONOSPHERIC AND AURORAL MEASUREMENTS FROM SPACE

USING VACUUM ULTRAVIOLET EMISSION

R.E. Huffman, D.E. Paulsen, F.J. LeBlanc, and J.C. Larrabee
Air Force Geophysics Laboratory (LK) Hanscom Air Force Base, MA 01731

ABSTRACT

Naturally occurring vacuum ultraviolet (VUV) emission from aurora and other ionospheric regions of the atmosphere can be used for global remote sensing applications from space. The measurement of continuous (diffuse) aurora and the tropical airglow belts on the recent DOD satellite S3-4 demonstrates the potential of such methods. These observations were made by the experiment Vacuum Ultraviolet Backgrounds, which covered the wavelength region from 1100 to 2900 Angstroms with both a spectrometer with 1, 5, or 25 Å bandwidth and a photometer using interference filters centered at 1216, 1340, 1550 or 1750 Å. Aurora are observed on every pass of the polar-orbiting satellite as an enhancement above the local background level due to airglow and/or solar scattered radiation. These enhanced emission regions are seen both with and without solar illumination. The paper will describe these observations and the relative intensities of VUV spectral features such as the nitrogen Lyman-Birge-Hopfield and Vegard-Kaplan bands, oxygen 1304 and 1356 Å lines, and the hydrogen Lyman alpha line. Comparison will be made with Defense Meteorological Satellite Program data including the operational line scanner visual images and electron precipitation counter results.

INTRODUCTION

The use of the vacuum ultraviolet (VUV) region of the spectrum for auroral and ionospheric observations from space is attractive for a number of reasons, including lack of cloud or ground albedo interference. This naturally occurring emission is due to collisional excitation in auroras and recombination processes in the tropical disturbed ionospheric belts. The satellite experiment Vacuum Ultraviolet Backgrounds[1] on the Space Test Program S3-4 flight has obtained data which furthers understanding of the potential of the VUV for these uses. It is the purpose of this paper to illustrate the use of VUV measurements in auroral and ionospheric studies using our S3-4 measurements.

The ultraviolet instrumentation on the satellite OGO-4 first clearly demonstrated VUV auroral[1] and equatorial[3,4] ionospheric measurements. Since these observations in the late 60's, satellite observational data has been sparse until our S3-4 flight in 1978. There have however been several suggested applications of VUV emission data to ionospheric measurements.[5,6,7]

EXPERIMENTAL

The VUV Backgrounds sensors obtained measurements from a polar, sun-synchronous orbit during the period from late March through mid-September 1978. The experiment consists of a photometer which may have any one of four interference filters in place and a dual spectrometer covering a VUV range (about 1100-1900 Å) and a UV range (about 1700-2900 Å) simultaneously. The photometer filters were peaked at 1216, 1340, 1550, and 1750 Å. and there was also a variable aperture to change the sensitivity. All sensors were pointed in the earth-center or nadir direction at all times. The experiment has been more completely described recently.[1]

An example of a complete revolution is shown in Fig. 1. The photometer is set at the 1340 Å filter, which is sensitive to the oxygen atomic lines at 1304 and 1356 Å. However, the filter also passes the hydrogen Lyman alpha geocorona at 1216 Å to a certain extent. Also, some nitrogen Lyman-Birge-Hopfield (LBH) bands are observed in the dayglow.

The difference in level between the night (left) and day (right) sides of the figure are clear. In addition, aurora are seen as large increases above the slowly varying solar controlled background in both the solar-illuminated and dark regions of the orbit. The observation of aurora is a normal feature of the data, being seen with all filters and on every pass. There are differences in location, intensity, and shape from orbit to orbit both due to normal variability and due to the relation of the S3-4 ground track to the auroral oval location.

The tropical ultraviolet airglow bands are seen near the night equator as two enhancement regions. In agreement with previous observations [3,4] these bands are found to be centered about

the magnetic dip equator and to be variable in intensity. The local time at the equator is approximately 2230 throughout our observations.

AURORAL OBSERVATIONS

An enlarged view of a pass across the auroral oval is shown in Fig. 2. The dashed curve is for the photometer with the 1550 Å filter in use. This filter observes primarily the nitrogen LBH bands. These bands extend from about 1325 Å to 2200 Å and they are prominent in aurora and in the dayglow. The filter does not transmit at wavelengths shorter than about 1480 Å and therefore the other prominent auroral emissions at 1216 Å (H), 1304 Å (O) and 1356 Å (O) are not seen by this filter.

A series of maxima are seen in Fig. 2. The poleward narrow structures are considered to be arcs and the broader peak on the equator side is correlated with the diffuse or continuous aurora. The figure shows one-second averages of the data, and more arcs are found when 0.1 second averages are used. The general Gaussian shape of the continuous aurora does not change, however. The edges of the auroral oval for Q = 3 are shown by the figure to be in good agreement with the observed structure. In addition, visible region Defense Meteorological Support Program (DMSP) images showing auroral structure obtained about one hour before and 45 minutes after the VUV data have continuous aurora in the hatched areas, which are plotted at the same magnetic latitudes.

The spectrometer scans its wavelength range sequentially, returning to the same wavelength about every 22 seconds. Intensities of several prominent features are shown in Fig. 2. The hydrogen Lyman alpha signal is high and decreases slowly as the satellite proceeds toward the night-side equator. There is however a slight rise near the peak of the continuous aurora. The oxygen 1304 Å line becomes very bright in the aurora and is centered more on the largest arc. It is broader than the 1550 Å aurora due probably to multiple scattering by atomic oxygen. This fact complicates modelling of the 1304 Å emission. Finally, a few points are shown for the oxygen 1356 Å line and for a feature near 1500 Å which is due to atomic nitrogen and nitrogen LBH bands. Thus, the spectrometer data confirm the auroral location.

A search has been made for accidental near-simultaneous observations of the aurora in the DMSP visible images and with the VUV Background experiment. Fig. 3 illustrates the best case found to date, where the time difference is about 10 minutes. Dr. J.H. Whalen of AFGL is responsible for identifying this image and also for the DMSP data shown in Fig. 2.

In this figure, the white dots are the positions of S3-4 at the UT minutes indicated. The point at 1530 is located at the bright auroral arc. Also observed in this night image are the Antarctic coastline and widespread other emission from scattering of moonlight by clouds and snow. The auroral arc eventually becomes indistinct to the right of the crossing due in part to the ground clutter.

The VUV observation is shown in the lower half of Fig. 4, where the 1550 Å photometer record across the auroral zone shown in Fig. 3 is displayed. In this case, a strong arc is seen near 1530, which is identified with the single arc seen in the DMSP image. However, an extensive continuous auroral region and several other arcs are seen as well in the VUV. We interpret the DMSP sensitivity limit to be insufficient to observe this additional structure. The auroral structure not seen in the DMSP visible image extends to at least 500 km on either side of the bright arc, as shown by the ground track scale in Fig. 4. In Figs. 2 and 4, the diffuse auroral intensity peak is about 200 Rayleighs.

It is to be noted that the VUV records are not subject to low altitude and ground albedo scattering, because the atmosphere below about 100 km altitude is opaque to the VUV due to absorption by molecular oxygen. This fact makes it possible to observe aurora in the VUV in the daytime, as shown elsewhere in this paper. The important point also to be made is that the VUV records at night are also not cluttered by ground and cloud albedo effects. The trace in Fig. 4 is clean in this regard. An increase in signal variability seen at low count rates is statistical in nature. Thus, VUV observations at night are easier to interpret than visible region images.

The oxygen 1304 Å line and hydrogen Lyman alpha from the spectrometer are shown in the top half of Fig. 4. There is a clear enhancement of about 80% above the geocoronal level for Lyman alpha. When the geocoronal background is subtracted, the difference curve clearly peaks closer to the continuous aurora than to the arc region. This enhancement in the hydrogen emission is evidence for proton auroral excitation. As in Figure 3, the 1304 Å line appears to be associated with the bright arc. It is clear that different VUV lines behave differently and contain separate diagnostic information for auroral studies.

The profiles across the auroral zone obtained in the VUV may vary greatly. Some examples are shown in Figure 5. Note the lower curve, which is a daytime observation with the 1550 Å filter. The aurora is clearly seen above the dayglow.

Correlation of the VUV observations with the SSJ/3 electron spectrometer measurements on the DMSP-F2 satellite has been possible in a few cases of near simultaneous auroral observations. The auroral zone location is similar in the two sets of data for the few cases available.

EQUATORIAL OBSERVATIONS

The presence of the tropical UV airglow belts in these records is unusual since the observational altitude is between 200 and 250 km at the night equator. Photometer and spectrometer data are shown in Figure 6. The spectrometer records show increased oxygen 1304 and 1356 Å emission near the night equator.

This radiation has been found to be due to oxygen recombination in the F-region, and it has usually been assumed that the emission originates at altitudes above the present measurements. Our observations may be due to extension of the night F-region slightly below our altitude or to multiple scattering from emission originating above the satellite. The intensities we observe are several times less than the OGO-4 values. In Figure 1, the intensities are 120 and 30 Rayleigh for the southerly and northerly peaks, respectively. In Figure 6, the comparable values are 170 and 20 Rayleigh.

REFERENCES

1. Huffman, R.E., F.J. LeBlanc, J.C. Larrabee, and D.E. Paulsen, Satellite vacuum ultraviolet airglow and auroral observations, J. Geophys. Res., 85, 2201-2215, 1980.
2. Chubb, T.A., and G.T. Hicks, Observations of the aurora in the far ultraviolet from OGO-4, J. Geophys. Res., 75, 1290-1311, 1970.
3. Barth, C.A., and S. Schaffner, Ogo-4 spectrometer measurements of the tropical ultraviolet airglow, J. Geophys. Res., 75, 4299-4306, 1970.
4. Hicks, G.T., and T.A. Chubb, Equatorial auroral airglow in the far ultraviolet, J. Geophys. Res., 75, 6233-6248, 1970.
5. Meier, R.R., and C.B. Opal, Tropical UV arcs: Comparison of the brightness with f_oF_2, J. Geophys. Res., 78, 3189-3193, 1973.
6. Tinsley, B.A., and J.A. Bittencourt, Determination of F region height and peak electron density at night using airglow emissions from atomic oxygen, J. Geophys. Res., 80, 2333-2337, 1975.
7. Chandra, S., E.I. Reed, R.R. Meier, C.B. Opal, and G.T. Hicks, Remote sensing of the ionospheric F layer by use of O I 6300-Å and O I 1356 Å observations, J. Geophys. Res., 80, 2327-2332, 1975.

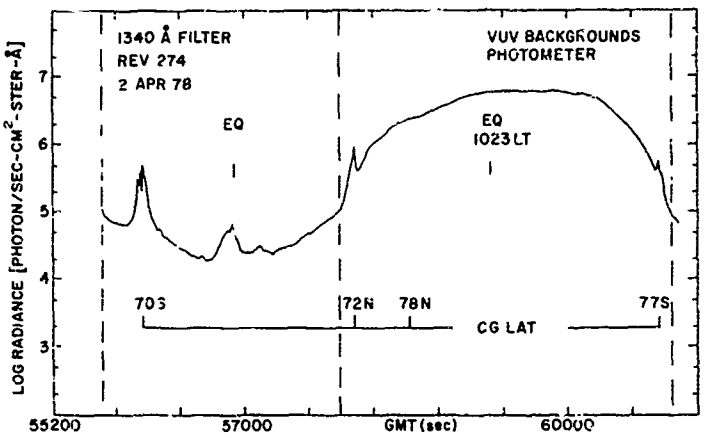


Fig. 1 VUV Backgrounds experiment global survey, one revolution

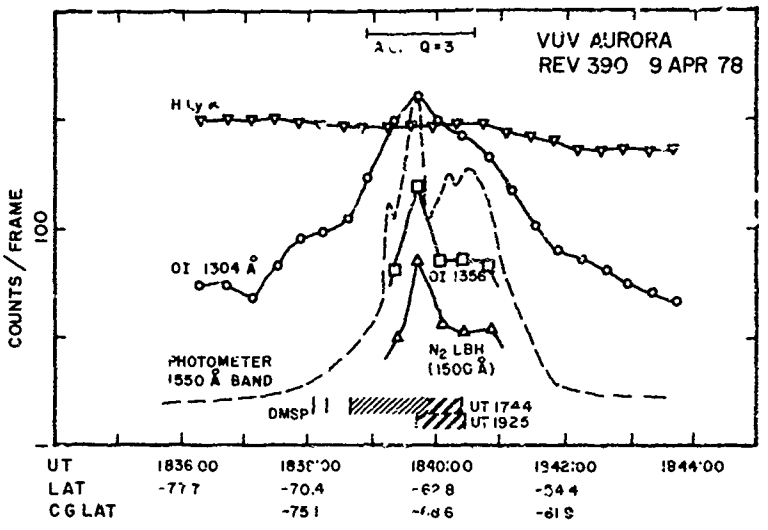


Fig. 2 VUV Backgrounds experiment auroral profile

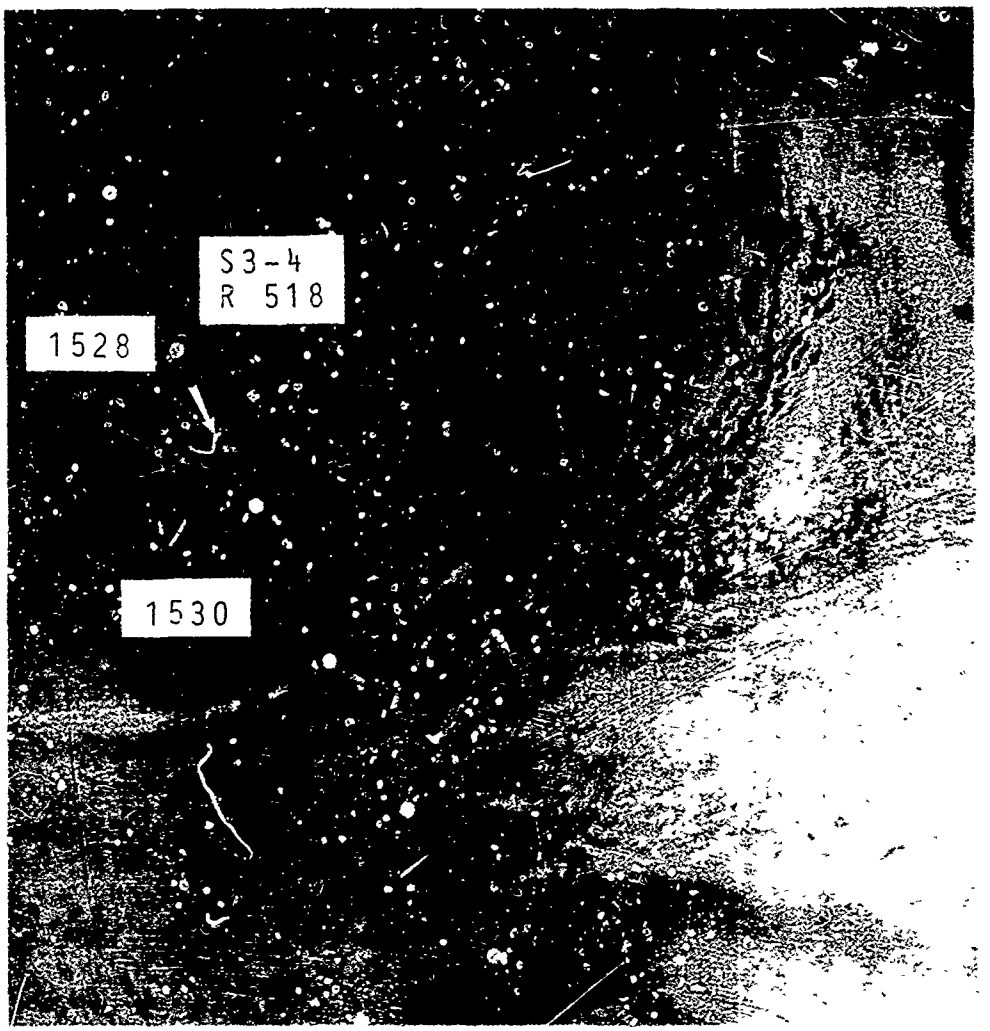


Fig. 3 Defense Meteorological Support Program (DMSP) image of auroral zone with S3-4 Rev. 518 pass shown, 17 April 1978. Local times shown for S3-4 with VUV Backgrounds experiment. DMSP image obtained about 10 minutes earlier.

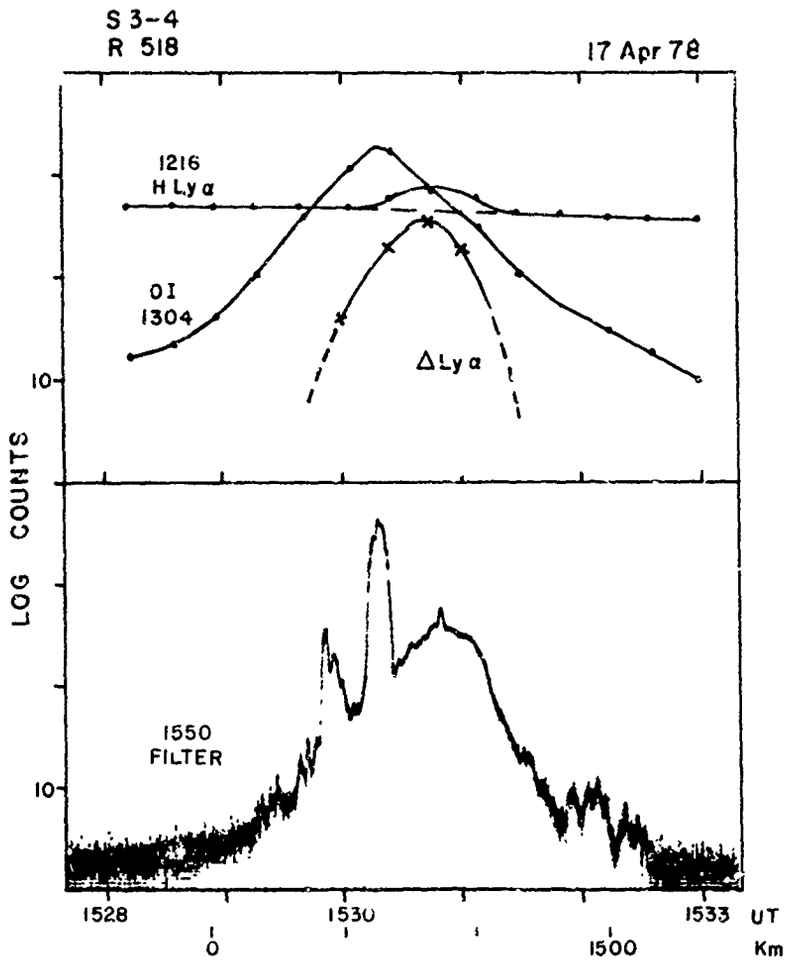


Fig. 4 VUV Backgrounds experiment auroral data corresponding to DMSP picture or previous figure. Bottom: 1550 Å photometer filter; Top: Spectrometer data

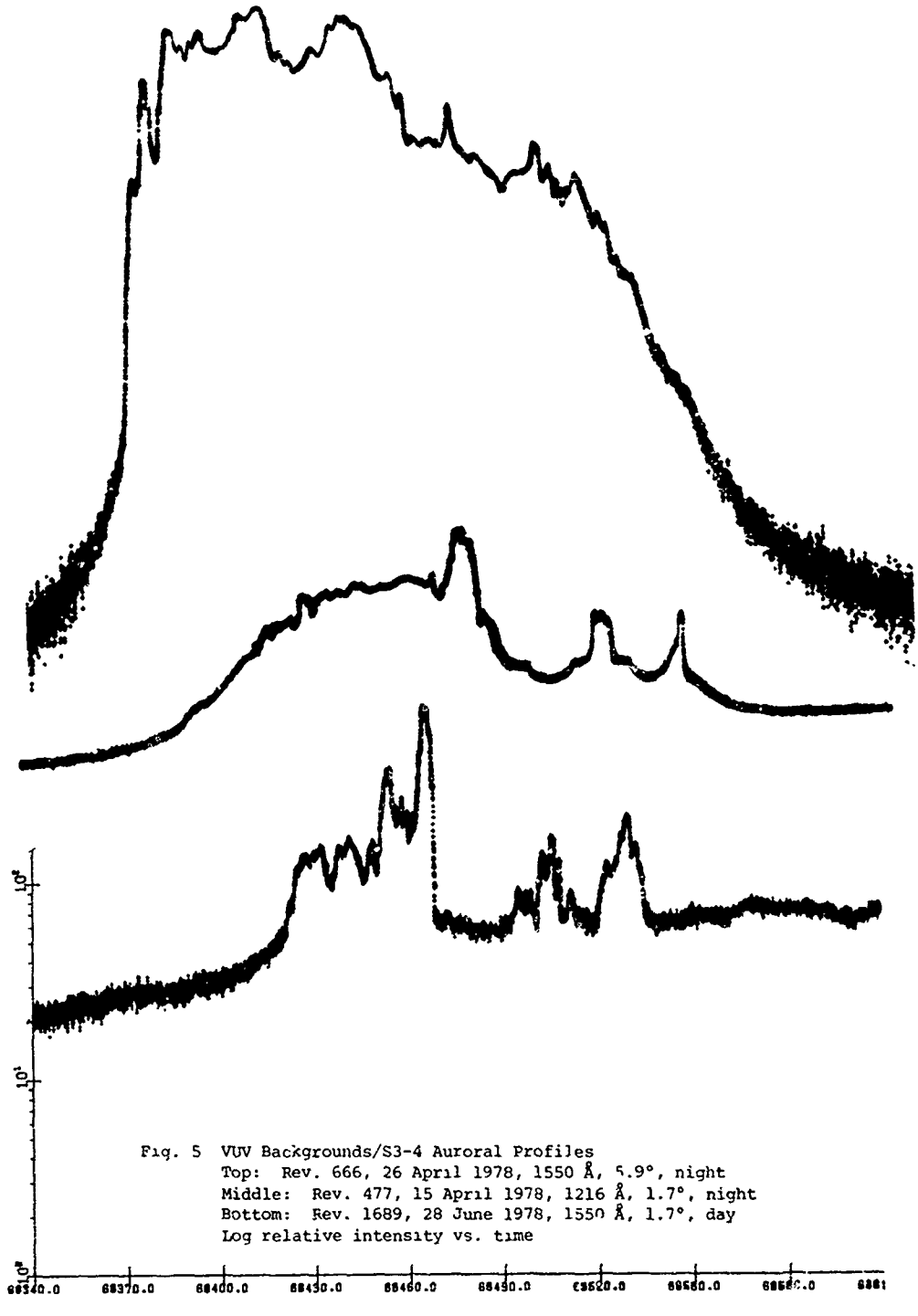


Fig. 5 VUV Backgrounds/S3-4 Auroral Profiles
Top: Rev. 666, 26 April 1978, 1550 Å, 5.9°, night
Middle: Rev. 477, 15 April 1978, 1216 Å, 1.7°, night
Bottom: Rev. 1689, 28 June 1978, 1550 Å, 1.7°, day
Log relative intensity vs. time

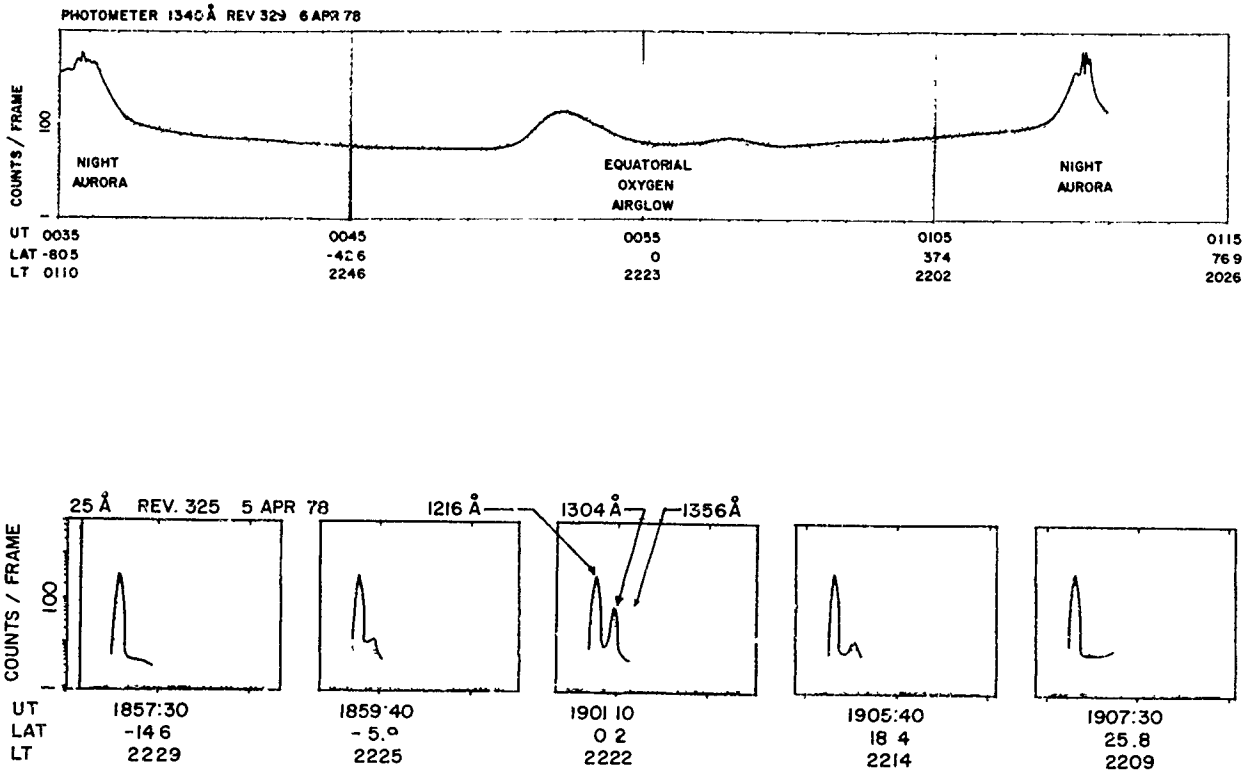


Fig. 6 VUV Backgrounds experiment observations of tropical UV airglow
Top: Photometer with 1340 Å filter; Bottom: Spectrometer

THE ROLE OF CURVATURE OF IONOSPHERIC IRREGULARITIES IN TRANSEQUATORIAL PROPAGATION AT VHF

Jerry A. Ferguson
Naval Ocean Systems Center
San Diego, CA 92152

INTRODUCTION

The theory of spread-F by Booker and Ferguson (1978) employs backscattering from long field-aligned irregularities of ionization in the F-region. This theory can be adapted for transequatorial propagation in the VHF band. It involves scattering by irregularities in planes perpendicular to the earth's magnetic field combined with coherent behavior along each field line. In the case of transequatorial scattering, curvature of the field lines results in caustics and foci. These results may prove sufficient to explain the signal strengths observed experimentally.

CAUSTICS AND FOCI

The calculations of Booker and Ferguson (1978) were for straight irregularities over a flat earth. This simplification is satisfactory for backscattering in the east-west plane at the equator. For transequatorial propagation, however, the curvatures of the earth and field lines must be considered. The scattering model to be employed assumes coherence along the irregularity axis and diffraction transverse to the axis. In order to consider the effect of the curvature of the field lines we have examined the specular component of rays scattered from a field line. The earth's magnetic field is approximated by a dipole so that the field lines are defined by

$$r = r_E \cos^2\theta \tag{1}$$

where r_E is the equatorial radius and r is the radius at a latitude θ. Eq. (1) can also be written

$$H = (H_E + a)\cos^2\theta - a \tag{2}$$

where the H's are altitudes above the earth's surface and a is the earth's radius. The geometry for the specular scattering diagrams is shown in Figure 1. Altitude

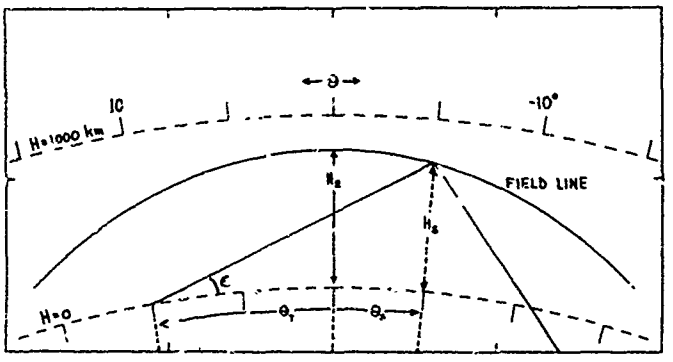


Figure 1. Geometry for specular scattering diagrams

references are provided by the dashed lines representing the earth's surface and a height of 1000 km. On each of these lines are marks every 5° of latitude; we

take northern latitudes as positive. In Figures 2 through 6 we display results for field lines defined by equatorial altitudes of 400, 800 and 1200 km with the 400 km case shown at the bottom. We launch rays at an elevation angle ε from a transmitter at latitude θ_T towards the field line. At the field line the specular scattering angle is determined and the appropriate scattered ray is traced. We also define a scattering latitude θ_S and altitude H_S. In the following diagrams we do not show rays for which H_S is less than 250 km or whose elevation angle at the transmitter is less than zero. Furthermore, for clarity, we do not consider reflection from the earth's surface.

Figure 2 shows ray paths for transmitters at the equator and at 5°. For the equatorial transmitter the scattered rays diverge except near the extreme scattering latitudes. When the transmitter is moved to 5° we see that the rays scattering from points south of the equator show convergence. For the 400 km field line the scattered rays converge to a focus within the boundaries of the figure.

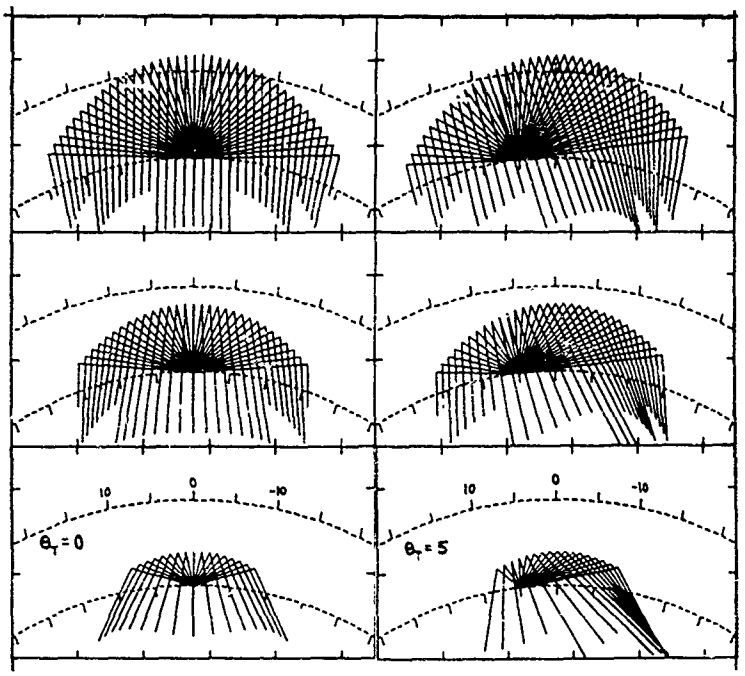


Figure 2. Transmitter at the equator (left) and at 5° (right)

Transmitter latitudes of 10° and 13° are shown in Figure 3. Now we can clearly see caustics and foci. For the transmitter at 10° the caustic associated with the 400 km field line is at the earth's surface. It is also clear that, for a field line between 400 and 800 km, the rays must converge to a focus at the earth's surface. This focusing occurs for H_E = 681 km. For the 400 km field line the caustic intersects the earth's surface at -8.8°. This indicates that the strongest field strengths may not be found at the conjugate point as might at first be expected. It is also interesting to note that the rays which form the caustic at the earth's surface are scattered south of the equator at -3°. When the transmitter is located 3° further north, caustics associated with field lines at all equatorial altitudes between 400 and 1200 km intersect the earth's surface. We let θ_C be the latitude of these intersections. For H_E at 400 and 1200 km we find θ_C at -9.1° and -12.2° and θ_S at -3.9° and -4.3°, respectively. We note that θ_C is north of the conjugate point at -13° and that θ_S has moved only 1° from that obtained with θ_T = 10°.

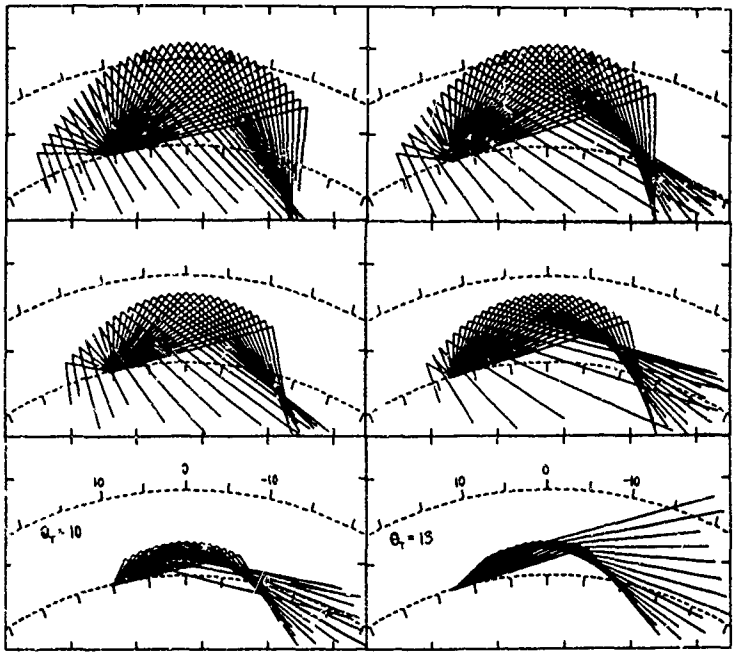


Figure 3. Transmitter at 10° (left) and 13° (right)

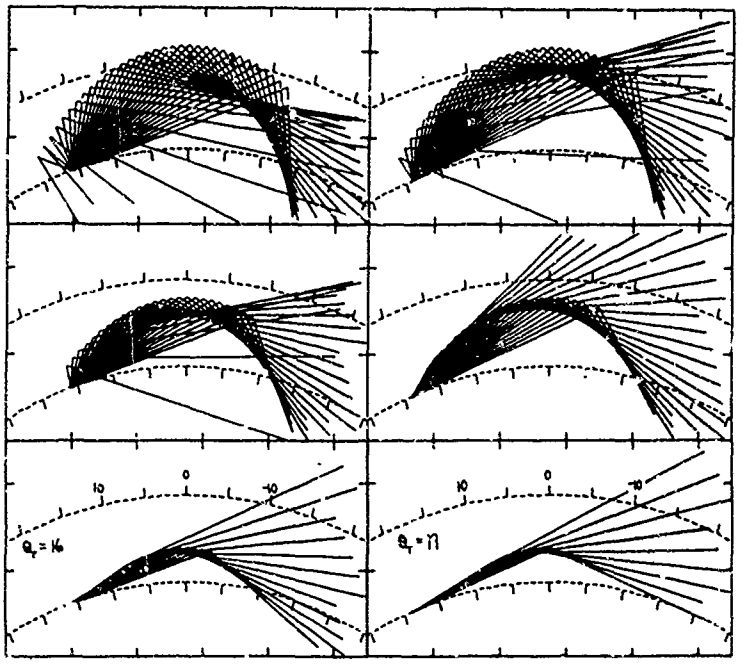


Figure 4. Transmitter at 16° (left) and 19° (right)

Figure 4 shows results for transmitters at 16° and 19°. These results continue those of Figure 3 except that the caustics for the lower field lines do not reach the earth's surface.

In Figure 5 the transmitters are at 22° and 25°. The ray paths for H_E = 400 do not appear because of the restrictions on H_S and elevation angle described earlier. It can be seen that, for θ_T somewhere between 22° and 25°, the intersection of caustics with the earth's surface disappears. We note that specularly scattered rays still reach the earth's surface, so that non-zero field strengths can be expected at the conjugate to the transmitter.

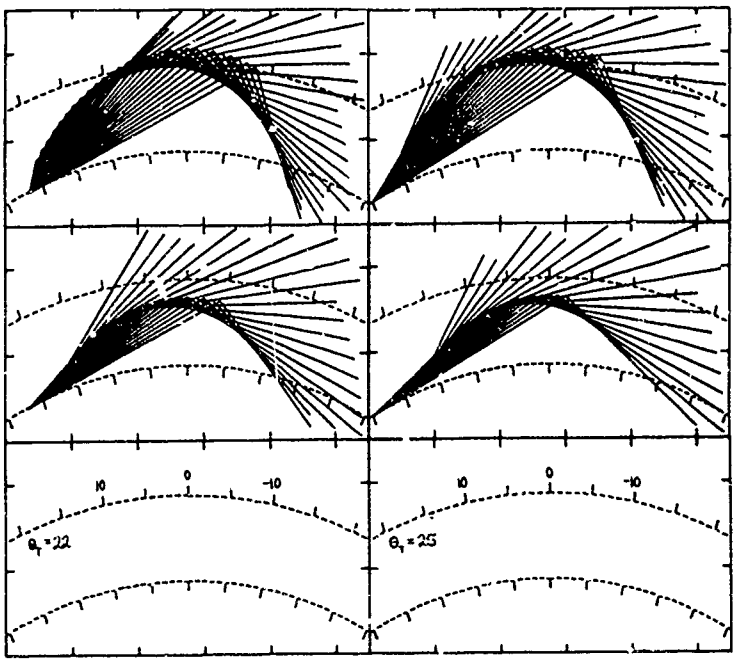


Figure 5. Transmitter at 22° (left) and 25° (right)

The values of θ_C and θ_S for Figures 2 through 5 are presented in the following Table. Also shown are the transmitter locations for each field line for which the caustic first intersects the earth's surface. In these cases the value of θ_C is the conjugate of θ_T and θ_S is at the equator. These cases also produce foci at θ_C. One notable result in this table is the extremely small variation of θ_C and θ_S obtained relative to the variations of θ_T.

H_E	θ_T	θ_S	H_S	θ_C	
1200	11.4	0	1200	-11.4	
	13.0	-4.3	1157	-12.2	
	16.0	-6.1	1114	-12.7	
	19.0	-6.9	1091	-12.9	
	22.0	-7.4	1074	-13.1	
	25.0	-7.6	1068	-30.0	($\varepsilon < 0$)
800	10.5	0	800	-10.5	
	13.0	-4.3	760	-11.3	
	16.0	-5.4	736	-11.7	
	19.0	-5.8	727	-11.7	
	22.0	-6.1	719	-11.8	($\varepsilon < 0$)
	25.0	-6.1	719	-11.6	
400	8.2	0	400	-8.2	
	10.0	-3.0	382	-8.8	
	13.0	-3.9	369	-9.1	
	16.0	-4.1	365	-9.0	
	19.0	-4.2	364	-9.0	
	22.0	-4.1	365	-8.8	($\varepsilon < 0$)
	25.0	-4.0	367	-8.7	

CURVED COORDINATES

In order to evaluate the scattered field in a curved coordinate frame, we need expressions for the distances from the transmitter to the scattering point, R_1, and from the scattering point to the receiver, R_2. It will be convenient to introduce a set of coordinates as follows:

$$u = \frac{r}{\cos^2\theta} \; ; \; v = \frac{\sin\theta}{r^2} \tag{3}$$

where r and $\pi/2 - \theta$ are the usual spherical coordinates and θ is the latitude. The third coordinate ϕ is the longitude. In this system of coordinates a field line is represented as a constant in u. Solving for r and θ we obtain

$$r = u(1 - f^2); \quad \theta = \sin^{-1} f \tag{4}$$

where f is a solution to the quartic equation

$$u^2 v = \frac{f}{(1 - f^2)^2} \tag{5}$$

In general, this equation has two real roots and two complex roots. The curve for one of the real roots passes through the origin and therefore satisfies Eq. (4). For convenience in deriving the distance equations we will normalize all radial coordinates to the earth's radius, a. It is notationally convenient to retain the expressions $\sin\theta$ and $\cos\theta$ instead of the corresponding expressions in terms of f.

Let the transmitter be located on the earth's surface at a latitude θ_1 and a longitude ϕ_1. The square of the distance from the transmitter to a point (r, θ, ϕ) is given by

$$\begin{aligned} R_1^2 &= 1 + r^2 - 2\, r \cos\gamma_1 \\ &= 1 + u^2 \cos^4\theta - 2\, u \cos^2\theta \cos\gamma_1 \end{aligned} \tag{6}$$

where

$$\cos\gamma_1 = \sin\theta_1 \sin\theta + \cos\theta_1 \cos\theta \cos(\phi_1 - \phi) \tag{7}$$

We are interested in the value of R_1 in the vicinity of an irregularity. Let us define the axis of the irregularity by u_0 and $\phi = 0$. Let us define a new variable

$$t = u - u_0 \tag{8}$$

where $|t| \ll 1$. We assume ϕ to be small, giving the following approximation to Eq. (6)

$$\begin{aligned} R_1^2 \approx 1 + {} & u_0^2 \cos^4\theta - 2\, u_0 \cos^2\theta \,(\sin\theta_1 \sin\theta + \cos\theta_1 \cos\theta \cos\phi_1) \\ & - 2\,[u_0 \cos\theta_1 \sin\phi_1 \cos^3\theta]\phi \\ & + 2\,[(u_0 \cos^2\theta - \sin\theta_1 \sin\theta - \cos\theta_1 \cos\theta \cos\phi_1) \cos^2\theta]t \end{aligned} \tag{9}$$

where terms of order ϕt, ϕ^2 and $(t/u_0)^2$ have been ignored. Let us now define

$$P_1^2 = 1 + u_0^2 \cos^4\theta - 2\, u_0 \cos^2\theta \,(\sin\theta_1 \sin\theta + \cos\theta_1 \cos\theta \cos\phi_1) \tag{10}$$

The value of P_1 is the distance from the transmitter to a point on the axis of the irregularity. This axis is defined by u_0 so that we take the terms in θ in Eq. (10) to give only terms in v. P_1 is much larger than the terms $u_0\phi$ and t which are the cross-sectional dimensions of the irregularity.

To lowest order in v, the last two terms in Eq. (9) simplify to give

$$R_1 \approx P_1 - \frac{\cos\lambda_1 \sin\phi_1}{r_1}\, u_0\phi + \frac{u_0 - \cos\lambda_1 \cos\varphi_1}{r_1}\, t \tag{11}$$

where

$$r_1 = (1 + u_0^2 - 2\, u_0 \cos\theta_1 \cos\phi_1)^{1/2} \tag{12}$$

A similar result applies to the distance to the scattering point from a receiver on the ground at latitude θ_2 and longitude ϕ_2.

Before writing out the required integrals, let us examine the terms in w and t in Eq. (11). Let the line from the transmitter to the center of the irregularity at the equator make an angle ν_1 with the equatorial plane. This line has length r_1, Eq. (12), in the normalized coordinates. Let the projection of this line onto the equatorial plane be ρ_1. Let the angle that this line makes with the meridian containing the irregularity be β_1. From the geometry just described it can be shown that

$$\frac{u_0 - \cos\theta_1 \cos\phi_1}{r_1} = \frac{u_0 - \cos\theta_1 \cos\phi_1}{\rho_1} \quad \frac{\rho_1}{r_1} = \cos\beta_1 \cos\nu_1 \tag{13}$$

$$\frac{\cos\theta_1 \sin\phi_1}{r_1} = \frac{\cos\theta_1 \sin\phi_1}{\rho_1} \frac{\rho_1}{r_1} = \sin\beta_1 \cos\nu_1 \tag{14}$$

Eq. (11) can be written as

$$R_1 = P_1 + \underset{\sim}{n}_1 \cdot \underset{\sim}{\tau} \tag{15}$$

where

$$\underset{\sim}{n}_1 = (\cos\beta_1 \cos\nu_1, -\sin\beta_1 \cos\nu_1, -\sin\nu_1) \tag{16}$$

$$\underset{\sim}{\tau} = (t, w) \tag{17}$$

A similar result applies for the receiver coordinates. These results can be used to show that the scattered field at the receiver is proportional to

$$E \sim a u_0 \iiint \frac{\Delta\epsilon}{\epsilon} \frac{1}{a^2 P_1 P_2} \exp\{-jka[(P_1+P_2)+(\underset{\sim}{n}_1+\underset{\sim}{n}_2)\cdot\underset{\sim}{\tau}]\}\, adt\, adw\, d\sin\theta \tag{18}$$

where $\Delta\epsilon/\epsilon$ is the fractional fluctuation of the permittivity due to the irregularity.

The irregularity is assumed to be uniform along its length so that the integrals over t and w can be separated from the integral over θ. The former integrals then give the Fourier transform of $\Delta\epsilon/\epsilon$, denoted by $D(\underset{\sim}{k}_1 - \underset{\sim}{k}_2)$. Thus, Eq. (18) can be written

$$E \sim \frac{u_0}{a} D(\underset{\sim}{k}_1 - \underset{\sim}{k}_2) \int_{-L}^{L} \frac{1}{P_1 P_2} \exp\{-jka(P_1 + P_2)\}\, d\sin\theta \tag{19}$$

where $\underset{\sim}{k}_1 = k\underset{\sim}{n}_1$, $\underset{\sim}{k}_2 = -k\underset{\sim}{n}_2$ and 2L is the length of the irregularity. It can be seen that the previously described caustics and foci must be included in the remaining integral. In particular, we are interested in the quantity $P_1(\theta)+P_2(\theta)$ which we will denote by $F(\theta)$ where $\theta_1, \phi_1, \theta_2, \phi_2$ and u_0 are constant.

SHORT PATHS

Let us examine the behavior of $F(\theta)$ by considering specific examples, beginning with the case of $\theta_1 = 10°$. In order to show the behavior of F in meridians which do not contain the transmitter, contour maps of constant values of F in a plane of latitude vs. longitude have been prepared. These maps are for fixed values of $\theta_1, \phi_1, \theta_2, \phi_2$ and H_E with the restrictions $\theta_1 = -\theta_2$ and $\phi_1 = \phi_2$. A line of constant longitude is to be taken as one containing an irregularity. The contour levels are shown for values of F separated by 10 km. Also shown is a curve representing the locations on the field line for which the elevation angle at either the transmitter or receiver is zero. Thus, only the area which includes the equator and is enclosed by the two curves can be considered in the scattering description. This area will be described as the scattering region. For H_E = 300 km, contours of F are shown in the left side of Figure 6, in the meridian containing the transmitter and receiver, we see minima near 8° and -8°. However, these minima are outside the scattering region. In this meridian the value of F at the

equator is a maximum. It is also a saddle point of F. The rays which arrive at the receiver, scattering from irregularities in the meridian of the receiver and transmitter, come from these three extrema in F. The ray diagrams of Figure 3 verify this. However, it is also apparent that, for ϕ near 3° and -3°, the value

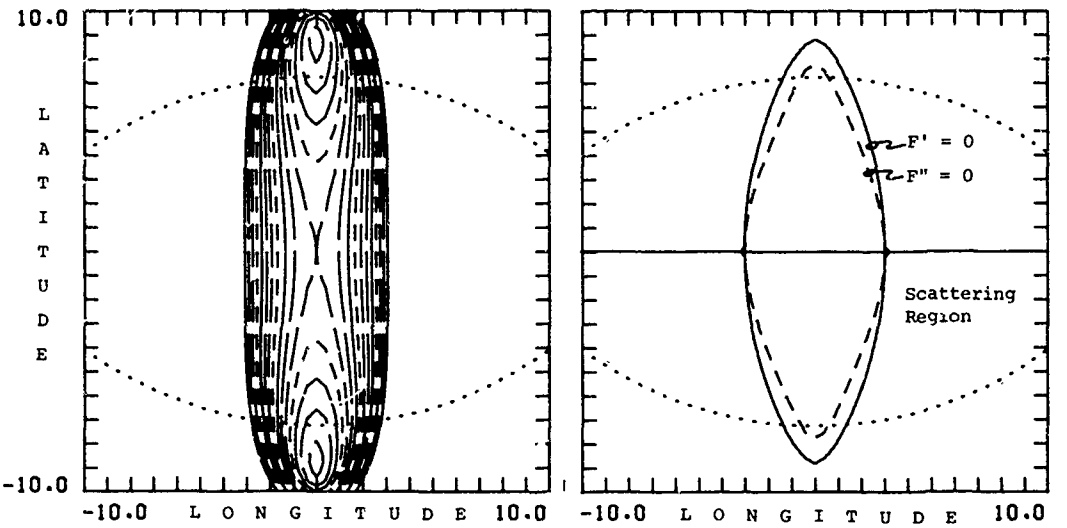


Figure 6. Contours of F (left) and F' = 0, F" = 0 (right) for a transmitter at 10° and H_E = 300 km

of F is approximately constant over a large range of θ. Contours of $F' = dF/d\theta = 0$ and $F'' = d^2F/d\theta^2 = 0$ are shown in the right side of Figure 6. Because $\theta_1 = -\theta_2$ and $\phi_1 = \phi_2$ we find F' = 0 at all points along the equator. There is another oval-shaped contour in the range of ϕ between +3° and -3°. This indicates that, for a fixed value of ϕ in this range, there are three extrema in F. The contour of F" = 0 shows the location of inflection points in F. We see that both derivatives are zero at the equator for $\phi = \pm 3°$.

Sets of contours for H_E = 500 km are shown in Figure 7. Features similar to those observed for H_E = 300 km are found. The minima in F are much shallower and the area enclosed by the contours of F' = 0 and F" = 0 is smaller. These minima are within the scattering region. Another set of contours, for H_E = 700 km, is shown in Figure 8. The minima of F in the meridian of the transmitter and receiver have now disappeared. This is particularly clear in that F' vanishes only at the equator and F" is not zero anywhere within the scattering region. In this case F is a minimum at the equator.

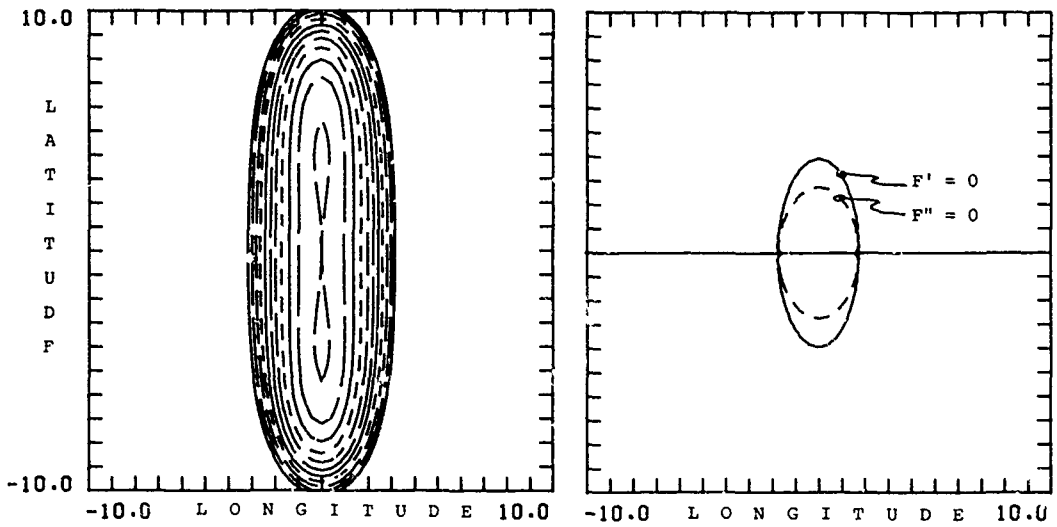


Figure 7. Contours for a transmitter at 10° and H_E = 600 km

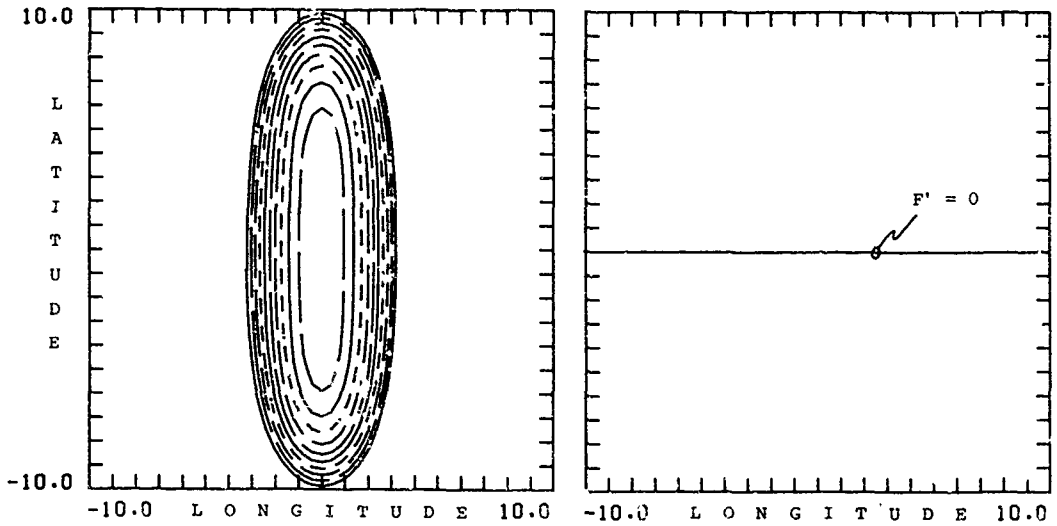


Figure 8. Contours for a transmitter at 10° and H_E = 700 km

The contours of Figure 6-8 are consistent with the ray diagrams of Figure 3. For H_E between 500 and 600 km, in the meridian of transmitter and receiver three rays are scattered to the receiver. These rays correspond to the three extrema observed in the contour diagrams of F. For H_E = 700 km only one ray is scattered to the receiver. For H_E = 681 km a large number of the rays scattered in the meridian of the transmitter and receiver arrive at the receiver (focusing).

The coincidence of F' = 0 and F" = 0 at $\phi = \pm 3°$ for H_E = 300 km causes focusing of rays at the receiver due to scattering from irregularities out of meridian of the transmitter and receiver. In this case the signals are arriving at the receiver at azimuthal angles of ±16.8° from the great-circle path.

The evaluation of the integral in Eq. (19) can be accomplished by locating the points of stationary phase, namely where F'vanishes. It is also clear that some cases will require evaluation of the integral at points where F" vanishes. Since focusing occurs latter case, contributions from these points will dominate the terms due to just F' = 0. Because of the complexity of the functional form of F(θ) it is not possible to determine when F'vanishes analytically except in the simplest cases.

LONG PATHS

We now consider cases where the transmitter is at a latitude greater than 20°. Because these paths are so long, the minimum equatorial field line altitude for which there is an appreciable scattering area is between 500 and 600 km. Another consequence of the path length is the spatial expansion of the oval shaped contour of F' = 0, seen in Figure 6. This expansion puts the oval, and the accompanying contour of F" = 0, outside the scattering area. These results are illustrated in Figure 9 which shows contours for $\theta_1 = 22°$, $\phi_1 = \phi_2 = 0°$ and H_E = 600 km. Thus, the only significant contributions to the integral in Eq. (19) are from equatorial values of the integrand. Raising the equatorial altitude of the field line from 600 to 800 km produces very little change in this result.

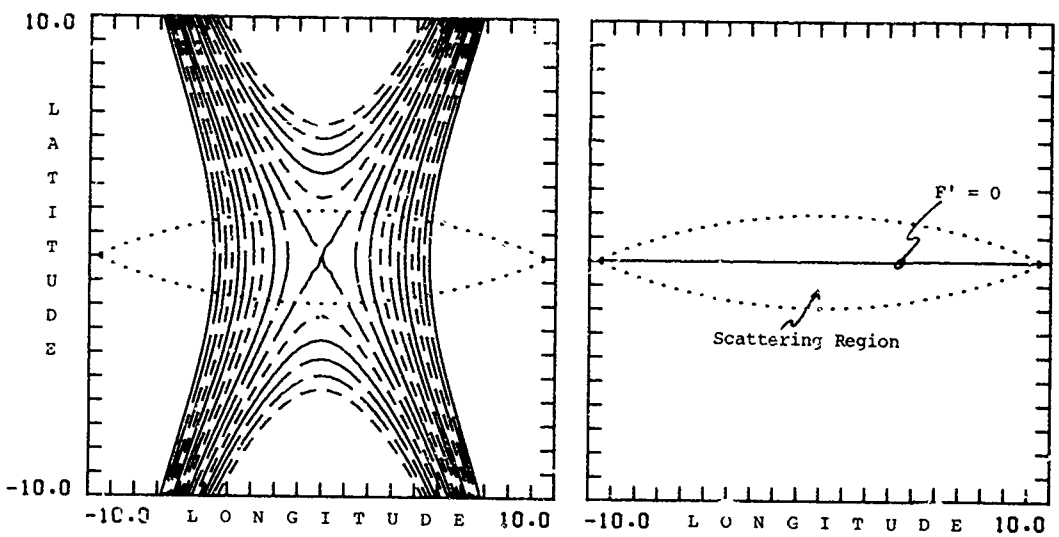


Figure 9. Contours for a transmitter at 22° with H_E = 600 km

DISCUSSION

Calculations of received field strength under conditions of transequatorial propagation in the VHF band need to be made incorporating the focusing phenomena described in this paper. It is clear that higher field strengths will be obtained than for scattering by straight field-aligned irregularities and much higher field strengths than for omnidirectional scattering

REFERENCES

Booker, H. G. and J. A. Ferguson, A theoretical model for equatorial ionospheric ionospheric spread-F echoes in the HF and VHF bands, J. Atmos. and Terr. Phys. 40, pp. 803-829, 1978.

SCINTILLATION STUDIES IN THE IONOSPHERIC TROUGH

J.W. MacDougall
Centre for Radio Science and Faculty of Engineering Science
The University of Western Ontario
London, Canada N6A 3K7

J.A. Koehler
Institute of Space and Atmospheric Studies, Department of Physics
University of Saskatchewan
Saskatoon, Canada S7N 0W0

ABSTRACT

During two weeks in August 1980 a two station experiment was run using a long east-west baseline. One station was at London, Ontario and measured scintillation activity and east-west F region convection in a region centred on 80°W, 39°N (geographic). The second station was at Southend, Saskatchewan and measured F region convection and scintillation indices in a region centred on 104°W, 50°N (geographic).

Scintillations were also being measured from Saskatoon and from London using radio stars. Auroral convection was monitored during active periods over a region centred on 103°W, 59°N. These comprehensive measurements are combined to show to what extent scintillations in the trough region are affected by convective processes.

INTRODUCTION

This paper examined some of the properties of scintillation activity in the region of the ionospheric trough. The original concept of this project was to examine the role of the ionospheric electric field in scintillations activity. One envisages two ways which the electric field may affect scintillations: 1. Strong electric fields can cause ionospheric irregularities to form by means of gradient-drift or two-stream mechanisms. These mechanisms have recently been reviewed by Fejer and Kelley (1980). 2. Electric fields might convect irregularities from a source region to the observing region. A southward convection of 55 m/sec means in about 6 hours irregularities would be transported 10°. Thus a patch of irregularities observed in midlatitudes might have been formed in the auroral zone.

In order to experimentally examine these roles of electric fields in scintillation behavior both electric field data and scintillation data covering a large region were desirable. Using the variations in scintillation activity at several widely spaced stations the size of a scintillation patch would be determined. These patches would then be compared with electric fields which might have played a role in their formation or movement. From previous studies, which will be discussed later, we thought that scintillation patches usually had east-west dimensions 1000 km or greater. Our information on the north-south dimensions seemed to indicate that although sometimes the patches were only the order of 100 km in width, usually they were wider.

Based on these 'facts' we planned an experiment with east-west spacings greater than 1000 km and several different north-south spacings. The opportunity for this experiment came in August 1980 and the following array of observations was obtained (see map, figure 1):

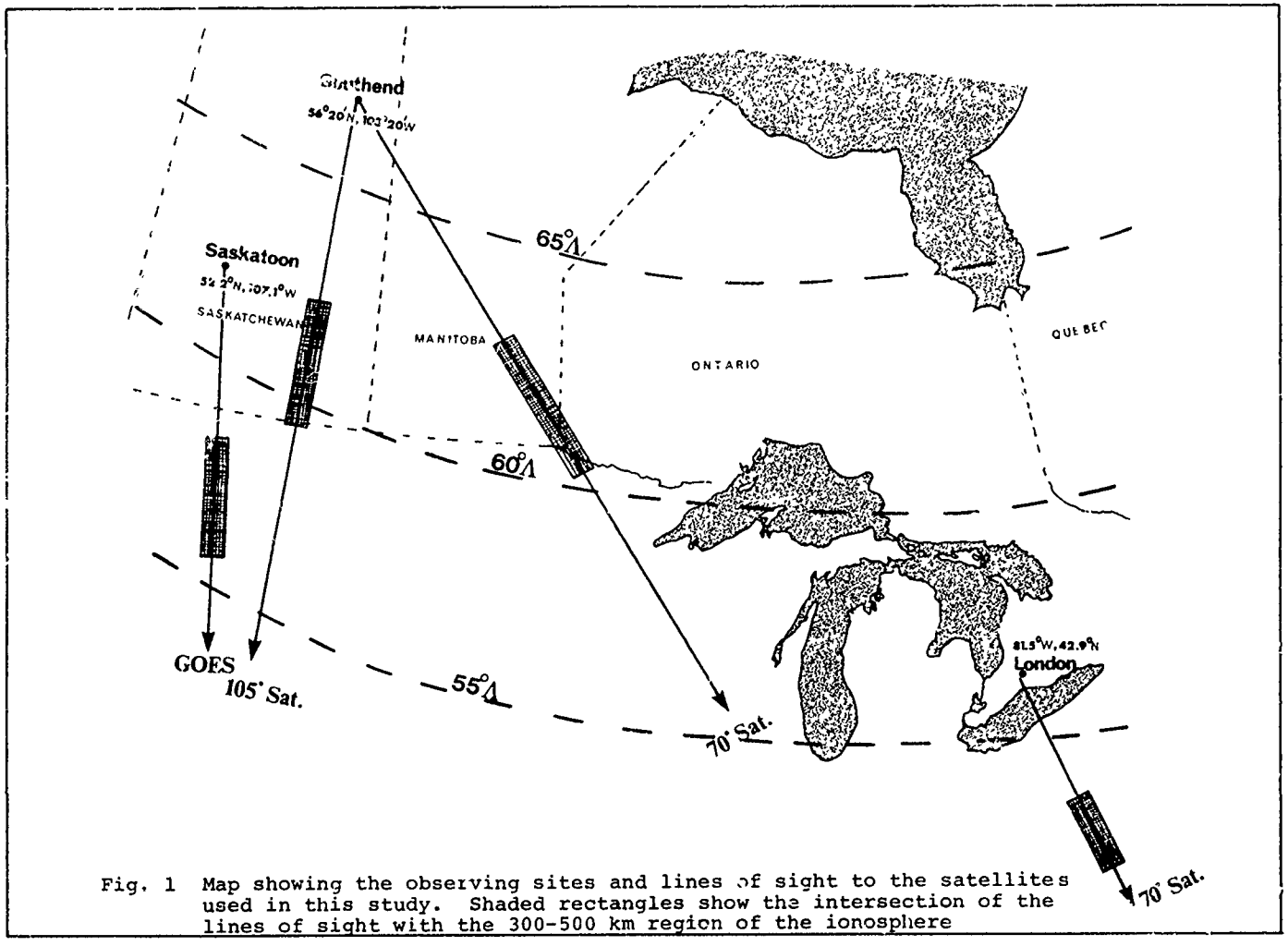


Fig. 1 Map showing the observing sites and lines of sight to the satellites used in this study. Shaded rectangles show the intersection of the lines of sight with the 300-500 km region of the ionosphere

(a) Scintillation indices from Southend in northern Saskatchewan using geostationary satellites ATS-3 at 105°W and ATS-5 at 70°W on a frequency of 137.35 MHz. The receiver was switched between these two satellites in alternative half hours. The line of sight to the two satellites intersected the F region ionosphere over the shaded regions shown in figure 1. A 300 km intersection is the north ends of the shaded regions and a 500 km intersection is the south ends. Earlier studies (see MacDougall, 1981a) showed that most nighttime scintillation irregularities are observed in this height range.

(b) Convection measurements from Southend. The scintillations in measurement (a) were also observed on spaced receivers in order to measure the convective motion. The analysis is similar to that described in MacDougall, 1981b and is referred to as the 'long-line' technique. The satellite at 105°W gives a measurement of east-west convective motion only. The geometry of the measurement using the satellite at 70°W makes the measured drift a mixture of east-west and north-south convective motion. Combining the two measurements should give both the east-west and the north-south convection. From the convection one can easily calculate the ionospheric electric field.

(c) Scintillation activity monitored from near Saskatoon, Saskatchewan using the GOES series of geostationary satellites on 136.38 MHz. The shaded region is as per measurement (a).

(d) Scintillation activity monitored from near London, Ontario using both the GOES series of satellites (not shown on figure 1) and the ATS-5 satellite at 70°W.

(e) Scintillation activity monitored from near London, Ontario using radio stars Cassiopeia A and Cygnus A as sources. These sources were observed near their upper transit locations (approximately overhead) and gave continuous nighttime observations between 2000 and 0600 hours local time during August.

(f) Convection measurements from near London, Ontario using the ATS-5 satellite at 70°. The measurement is similar to (b) above and to a first approximation gives the east-west convection although there will be a small component of north-south convection in the measurement.

(g) Auroral irregularities and north-south motion over Southend. Figure 1 does not show the details of this last measurement. A pair of 40 MHz transmitters and a pair of receivers were set up with a suitable geometry for detecting field aligned E region irregularities which might be formed by gradient-drift or two stream instability mechanisms. The region with a suitable geometry was in the vicinity of Southend. Since, as will be seen later, the scintillation measurements indicate too small a scale size for good comparison to be made to results from this experiment, this measurement will not be described in detail.

RESULTS

Patch Sizes

In the above sets of measurements, 20 days of scintillation activity was monitored in five areas (the four shaded regions shown in figure 1 plus the nighttime radio star measurements over London). Indices were subjectively scaled on a scale of 0 - 2 in increasing amplitude of scintillations. The five sets of measurements were cross-correlated in pairs. Since scintillation activity has a strong 24 hour periodicity, being most common at night, all these cross-correlations showed a pronounced diurnal oscillation of which figure 2a is typical. Note that the central maximum is no higher than other peaks in the oscillation. Scintillation activity on the same day is therefore no more highly correlated than is activity on randomly chosen days.

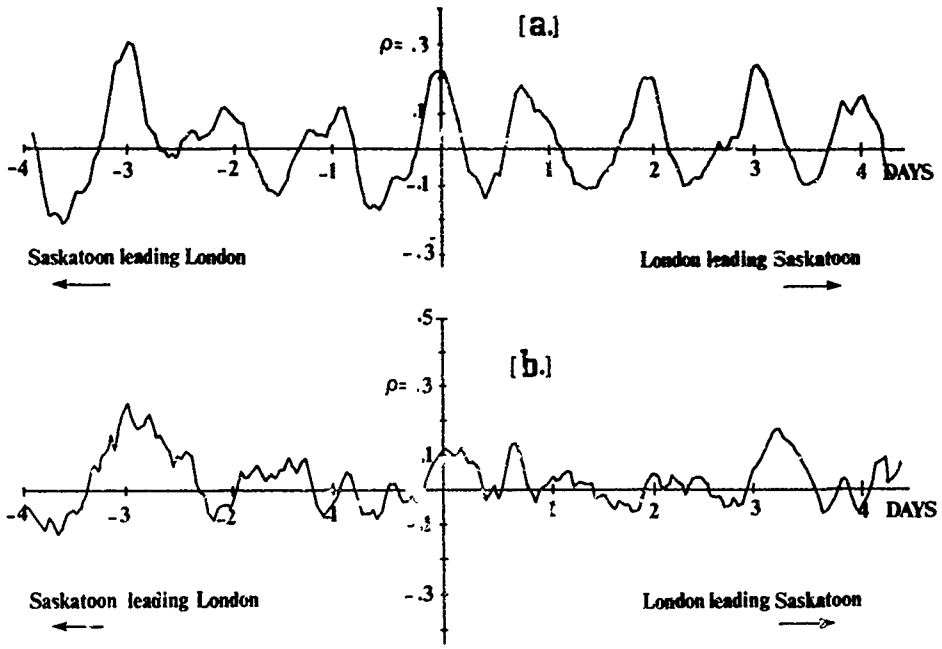


Fig. 2a Cross-correlation function between scintillations activity observed at London and Saskatoon

Fig. 2b As for 2a except scintillation activity has been normalized to remove the diurnal variation

Since the strong diurnal variation in scintillation activity is tending to hide any less pronounced correlations, the data was normalized by dividing each hourly scintillation index by the average index for that hour. A normalized index greater than one means higher than average scintillation activity during that hour. Cross-correlating these normalized indices for most of the station pairs gave results in which there is no evidence of significant cross-correlations. Figure 2b is typical of the results produced and does not show any evidence of significant correlation at zero days time shift.

The cross-correlations indicated that the patches must be of relatively small size. The smallest spacings in the array of measurements were the approximately 500 km east-west spacing between the 105° satellite and the 70° satellite observed from Southend, and the approximately 300 km north-south spacing between the 105° satellite observed from Southend, and the GOES satellites observed from Saskatoon. The cross-correlation functions for this 500 km east-west spacing is shown in figure 3, and there is a significant positive central correlation maximum. The cross-correlation function for the 300 km north-south spacing is shown in figure 4 and there is no obvious correlation around zero time shift. From these results we conclude that during August 1980 the size of the average scintillation patch was slightly smaller than 500 km in the east-west direction, and much smaller than 300 km in the north-south direction. This conclusion is also in agreement with the cross-correlation for the 70° satellite observed from London and radio star scintillations also from London which again showed no correlation for approximately a 300 km north-south spacing.

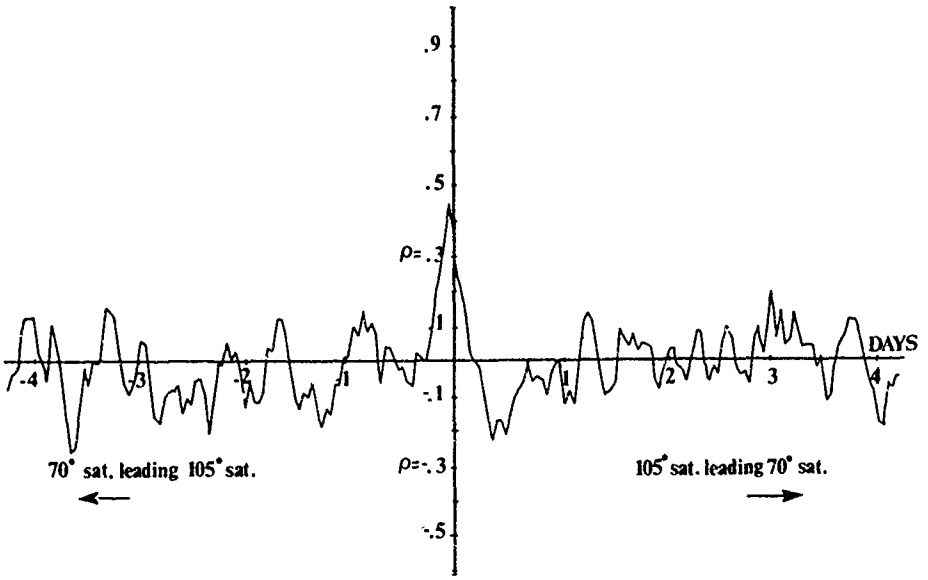


Fig. 3 Cross-correlation function between normalized scintillation activity observed from Southend using the two satellites

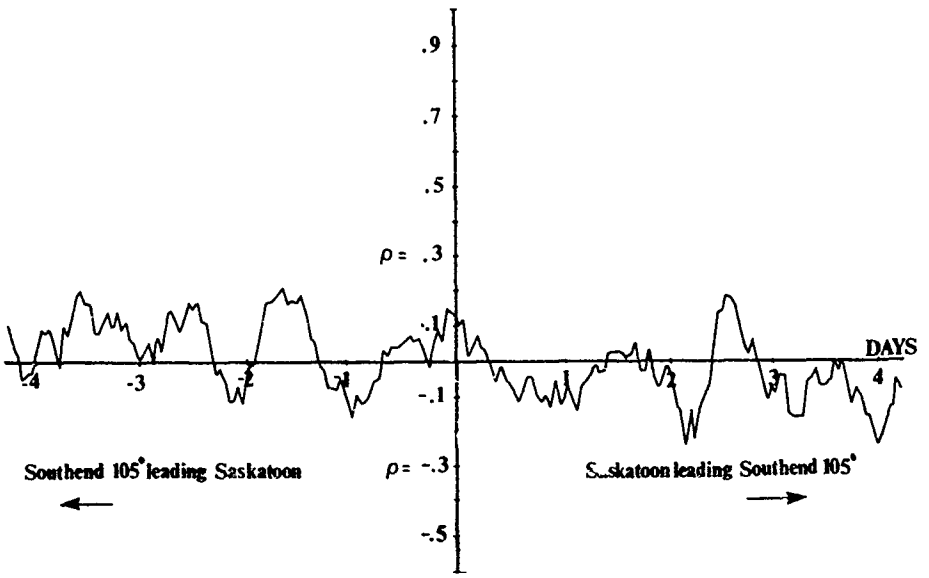


Fig. 4 Cross-correlation function of normalized scintillation activity observed from Saskatoon using GOES satellites and from Southend using the 105°W satellite (ATS-3).

Discussion of the Results for Patch Dimensions

As mentioned in the introduction, when this experiment was planned we expected larger sizes for the patches of irregularities than the < 300 km N-S, ~500 km E-W dimensions as determined above by the cross-correlations. A number of mid-latitude studies (Kent, 1959, Singleton and Lynch, 1962; Frihagen, 1963; Munro, 1963) had all found patch dimensions 200 - 400 km north-south and greater than 1000 km east-west. An earlier cross-correlation between London and Boston scintillations (the 'Boston' scintillation data was kindly supplied by J. Aarons of AFCRL) gave a cross-correlation greater than .7 for the spacing of about 800 km east-west. This earlier result was for equinox. We also did an equinoctial study on scintillation patches (MacDougall, 1981a) using the scanning effect of a moving satellite passing east or west of the observer. This enabled us to plot the positions of individual irregularities within the patches. These plots showed both large diffuse patches, and smaller more compact ones. An example of a moderate sized patch (many were smaller) is shown in figure 5.

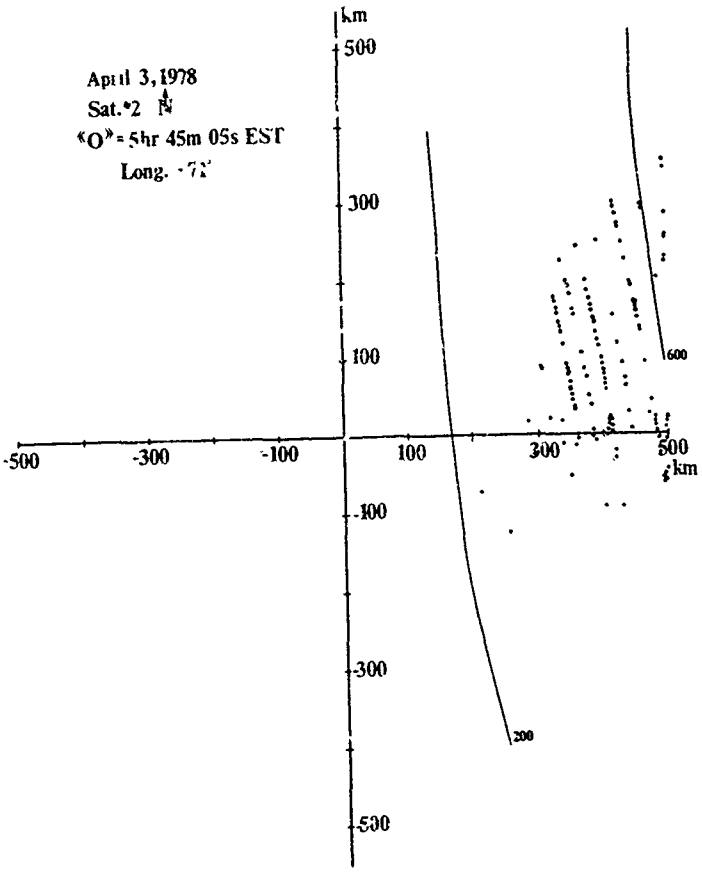


Fig. 5 A moderate size patch of scintillation irregularities seen by using the scanning effect of NNSS satellites. The two lines marked 200 and 600 show approximately the limits of the 'picture'. The scales show the geographically oriented coordinate system with the observer at the origin. The legend on the figure indicates that this was satellite #2 northbound and that it crossed the observer's latitude (the x axis) at 71°W at the time shown.

In equinox, therefore, although we found small patches (at about 55°Λ) the correlations showed that usually the patches were approximately 400 x 1000 km in size. In summer (August) it seems that this is not true, at least near 60°Λ, and small patch sizes predominate in the region of this study.

The correlation functions also can give us an indication of the time scales associated with the patches of irregularities. Autocorrelation functions were calculated for the normalized measurements of scintillation activity from the various stations. A typical autocorrelation function is shown in figure 6. The time scale of the patches is obviously of the order of a few hours.

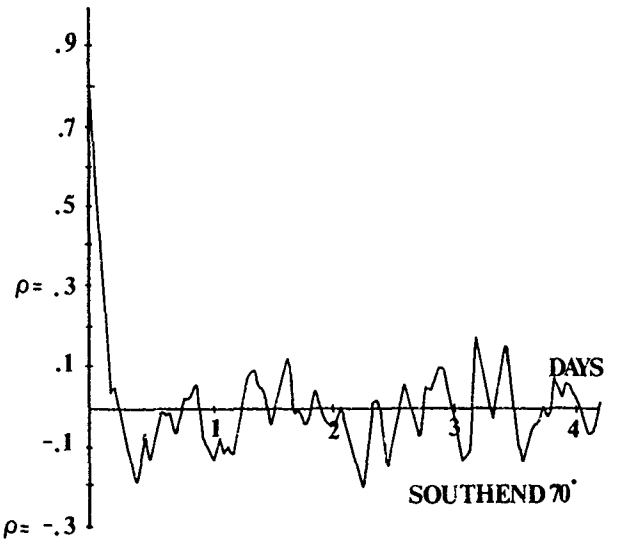


Fig. 6 Autocorrelation function of the normalized scintillation activity observed using the 70°W satellite from Southend

Before discussing possible mechanisms for formation of the scintillations there is one further comment on the scintillation statistics from this study. The scintillations observed using the two satellites from Southend always showed strong activity at night and moderate activity during daytime with the day-to-day variation in activity being a small modulation of the average behavior. At the more southerly stations the day-to-day variations were much larger relative to the average behavior. It is difficult to compare numerical values of scintillation levels for the various receiving locations used in this study since the elevation angles of the line of sight to the satellites are different for each of the locations, and (see for instance figure 3 of JSSG, 1968) scintillations would be expected to be much stronger for the small elevation angles at Southend. The latitudinal variation of the average scintillation behavior does appear to be compatable with that in JSSG, 1968 which shows a large increase of average scintillation index at about 60°. In this study we are not dealing with this average background level but with the fluctuations, or patches, which are superimposed on it.

Electric Field Effects in Scintillation Patches

The original intention of this study was to examine the role of electric field in causing scintillations. Our immediate conclusion is that electric fields appear to play a relatively minor role.

Electric field data was obtained by looking at the convection of the scintillation irregularities past a pair of antennas spaced apart 600 meters at Southend. (The layout at London for this measurement is described in MacDougall, 1981b). Convection results from Southend for a typical quiet day are shown in figure 7.

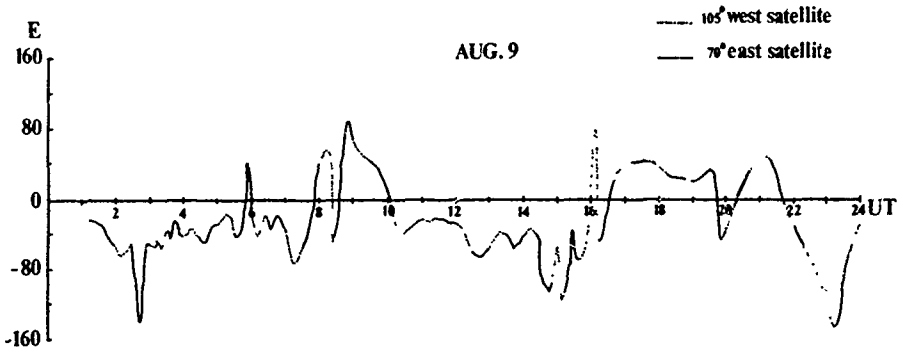


Fig. 7 East-West ionospheric convection measured from Southend using satellites as indicated

During disturbed days there was strong nighttime westward convection before local midnight. The average daily variation for quiet days was approximately just a 40 meters/sec westward convection. This may be seen during most of the day shown in figure 7. Figure 7 also indicates two further facts about the convection:

1. North-south convection is small.
2. The convection is well correlated over east-west distances of the order of 500 km.

The evidence for these two facts is that, as indicated by the symbols, the line in figure 7 is composed of alternate half hour segments using the two satellites observed from Southend. When using the 70° satellite the measured drifts are an approximately equal sum of north-south and east-west convection, whereas using the 105° satellite the drifts observed are almost purely due to east-west convection. The continuity of the curve therefore indicates that north-south convection is small relative to the east-west.

The curve also shows that although the two observing regions, see figure 1, are about 500 km apart, there is no obvious discontinuity in the patterns when switching from one satellite to the other and therefore east-west scale sizes for the electric field patterns are of greater dimension than 500 km. A comparison of the Southend and the London drift measurements however does show very noticeable differences (figure 8 shows an example). It is not yet clear to what extent these differences are an east-west or a north-south effect.

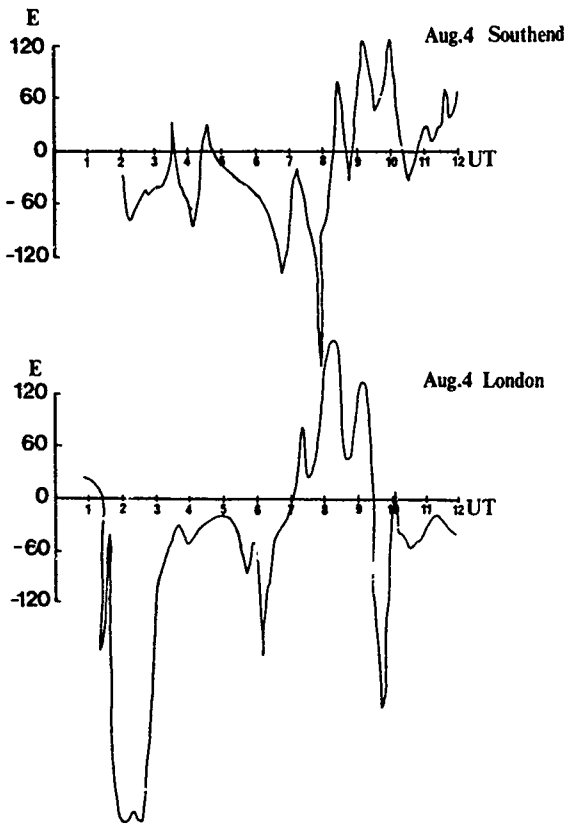


Fig. 8 A comparison of east-west convection measured simultaneously from Southend and from London

In summary the August convection near 60°Λ shows; (a) Average quiet convection is 40 m/sec westward with small north-south motion. (b) Scale sizes in the convection patterns are much greater than 500 km east-west.

There is only slight evidence in the scintillation data that convection or electric fields were playing an important role. Cross-correlations between measured convective speeds and scintillation indices from the various sites showed no significant correlation. Also, the east-west correlation in figure 3 shows a time shift of about 1 hour which would indicate westward drift of 140 m/sec. This is much greater than the measured average westward convection, 40 m/sec, and may be more a local time difference than an actual movement. Finally, the greater than 500 km scale size of the electric field would not be consistent with forming scintillation patches of smaller dimensions.

What Causes Patches of Scintillations?

A summary of the morphology of the August scintillation patches near 60°Λ is as follows:

1. Size; a few hundred kilometers east-west and smaller north-south.
2. Time scale; patches persist for several hours.
3. Movement; there is only slight evidence that patches move in a geographical coordinate system.

Most geophysical phenomenon do not have properties consistent with this morphology. As discussed above, the ionospheric electric field as measured by our technique does not have similar properties. Two geophysical phenomenon which do have comparable properties are localized weather systems and 'detached auroral patches'.

It has been shown (see for instance Hung et al., 1979 and references therein) that weather systems may produce gravity waves which are observable at F region heights. There does not seem to be any indication that weather systems can produce scintillation activity. The second phenomenon of auroral patches is therefore a more obvious possibility to associate with the scintillation patches.

The auroral patches observed by Moshupi et al. (1979) in the trough region from the ISIS 2 satellite would seem to have properties which are very similar to the patches of scintillations. Summarizing briefly the properties of the auroral patches they have;

1. Dimension; typically 100 km N-S by 200 km E-W.
2. Time scale; uncertain but probably of the order of an hour.
3. Movement; the patches co-rotate.
4. Geographical location; maximum occurrence frequency over central Canada and at about 62° invariant latitude.

It may be seen that the properties of the auroral patches are very similar to the scintillation patches. The auroral patches are a comparitively rare phenomenon however, only 22 having been seen on the ISIS photographs for over 1000 passes. Presumably the photographic data have some threshold before a patch would be detected. An auroral patch which could cause scintillations might usually be below this threshold. Moshupi et al. (1979) attribute the patches to detached plasma from the plasmaspheric bulge. This detached plasma is at a different temperature from the surrounding trough plasma. It is not clear whether it is the temperature difference or the auroral particle precipitation which would cause the scintillations. Fejer and Kelley (1980) discuss examples of irregularities associated with both heating and auroras.

CONCLUSION

In the region of the ionospheric trough during August, 1980 scintillation enhancements had an average patch size about 500 km east-west but much smaller in the north-south direction. Electric fields did not appear to have a significant effect in creating these patches of scintillation enhancements. A phenomenon called 'auroral patches' has many properties similar to those of the scintillation patches. This may indicate an association with the scintillation patches.

ACKNOWLEDGEMENTS

S. Frank, R.J. Haycock and M. Meighen assisted with the collection and analysis of this data. This research was supported by a grant from the Natural Sciences and Engineering Research Council of Canada.

REFERENCES

Fejer, B.G., and M.C. Kelley. Ionospheric Irregularities, Rev. Geophys. Space Sci. 18, 401-454, 1980.

Frihagen, J. Study of fine-scale ionospheric structure by satellite techniques, Radio Astronomical and Satellite Studies of the Atmosphere, North Holland, Amsterdam, 333-343, 1963.

Hung, R.J., T. Phan and R.E. Smith. Coupling of ionosphere and troposphere during the occurrence of isolated tornadoes on November 20, 1973, J. Geophys. Res., 84, 1261-1268, 1979.

JSSG (Joint Satellite Studies Group), On the latitude variation of scintillations of ionospheric origin in satellite signals, Planet. Space Sci., 16, 775-781, 1968.

Kent, G.S. High frequency fading observed on the 40 Mc/s wave radiated from artificial satellite 1957α, J. Atmos. Terr. Phys., 16, 10-20, 1959.

MacDougall, J.W., Distribution of the irregularities which produce ionospheric scintillations, J. Atmos. Terr. Phys., 43, 1981a.

MacDougall, J.W. Measurement of Ionospheric Convection by the long-line technique, J. Geophys. Res., 86, 1981b.

Moshupi. M.C., C.D. Anger, J.S. Murphree, D.D. Wallis, J.H. Whitteker andL.H. Brace, Characteristics of trough region auroral patches and detached arcs observed by ISIS 2, J. Geophys. Res., 84, 1333-1346, 1979.

Munro, G.H., Scintillation of radio signals from satellites, J. Geophys. Res., 68, 1851-1860, 1963.

Singleton, D.G., and G.J.E. Lynch, The scintillation of the radio transmissions from Explorer VII-1I, J. Atmos. Terr. Phys., 24, 363-374, 1962.

OCCURRENCE OF F LAYER IRREGULARITIES IN THE POLAR CAP

Allen Johnson
Avionics Laboratory
Wright Patterson Air Force Base, OH 45433

Jules Aarons, Jurgen Buchau, John P. Mullen, Edward Weber, and Herbert E. Whitney
Air Force Geophysics Laboratory
Hanscom Air Force Base, MA 01731

ABSTRACT

Over the past ten years a series of long term ground measurements and intensive airborne campaigns into the polar cap region have been undertaken to investigate F layer irregularities and their effect on earth-to-satellite communication links.

These measurements indicate that the occurrence of ionospheric scintillation fading in the polar cap is strongly correlated with solar flux density and season, but poorly correlated with magnetic activity and time-of-day.

The long term ground measurements taken during the low solar flux year of 1975 showed little scintillation activity. Similar measurements made during years of high solar flux density (1979-1980) showed a weak background level of scintillation and also scintillation with intensities greater than 20 dB peak-to-peak at 250 MHz occurring many hours per day. Maximum scintillation occurrence of these levels took place in the dark winter months of 1979-1980 with the summer period showing decreased activity. This behavior is similar to in-situ polar cap electric field measurements which show larger fluctuations in the winter than in the summer, an effect which may be related to E layer conductivity changes.

The F layer irregularity structures were identified in intensive campaigns in 1979. All-sky imaging photometer 6300 Å images show that intense polar cap scintillations are associated with sun-aligned F layer auroral arcs. The arcs drift predominately in the dawn to dusk direction, but reversals are also observed. Airborne ionosonde measurements also show this motion. The drift of these structures into the satellite-to-ground ray path results in dramatic increase in scintillation level, implying the existence of more intense ionospheric irregularities within these auroras.

Irregularity drift velocities were investigated using ground antennas/receivers spaced in a triangular pattern. Measurements of the time-difference-of-arrival of the irregularity at the three antennas provided drift velocities from 40 to 800 meters per second. The drift direction followed a complex diurnal pattern with rapid reversals of direction, indicating that the irregularities do not follow a simple noon-to-midnight flow pattern within the auroral arcs.

INTRODUCTION

Amplitude scintillation measurements were performed at Thule Air Base, Greenland, using a satellite beacon at ∿250 MHz to assess the diurnal, seasonal, geomagnetic and solar relationships of ionospheric irregularities in the polar cap.

The measurements were taken during two distinct periods: (1) April-October, 1975, a period of low solar flux and (2) November 1978-March 1980, a period of high solar activity.

Amplitude fluctuations converted to scintillation index (SI) by the method shown in Whitney et al. (1969) are summarized in both mean dB excursion, and percent occurrence of scintillation index above certain levels, i.e., 10 dB peak-to-peak, a level of activity in the strong scattering regime

indicative of intense ionospheric irregularities. The concept of this paper is to describe the scintillation levels observed with respect to possible geophysical forcing functions.

The observations from Thule were performed within an elevation angle range of 30° to 65° at azimuths towards the northwest and southeast. The resulting 350 km sub-ionospheric intersections are shown within the dotted lines in Fig. 1. This coverage is from ~82° to 90°N CG latitude. The results are presented in Universal Time (LT = UT-4)

DIURNAL AND SEASONAL VARIATION

A contour plot of the percent occurrence of scintillation index greater than 10 dB is shown in Figure 2. The plot was developed from hourly average values of the 15 minute SI for each month for low magnetic activity (Kp=0-3). Two patterns emerge: (1) Maximum occurrence of activity takes place in the months of little or no sunlight at F-region heights. Much lower scintillation occurrence takes place in the sunlit months. (2) The diurnal variation is weak, and apparent only during the winter months.

To investigate the seasonal variation, the data were ordered in three month periods (November, December, and January as winter, etc.) and the diurnal variation of mean SI for each season was plotted for both low magnetic activity (Kp=0-3), and for more active periods (Kp=3+-9) (Figure 3a-d).

The data show that the lowest values of mean SI (~6 dB) occur in summer, and the largest during fall and winter. In general, there is only a slight increase in mean SI with increased magnetic activity during summer, fall and winter.

SUNSPOT CYCLE

A series of measurements was taken between April and October 1975. During this period of low solar activity, there was an absence of strong scintillation activity to such an extent that we can plot only the occurrence of scintillation greater than 6 dB. Only one period of 15 minutes showed a single peak-to-peak fluctuation of 10 dB in the 1975 study. Figure 4 shows the contrast between the 1975 period when solar flux was low (10.7 cm flux was ~75) and the same months in 1979 when solar flux was high (~150-225). The occurrence levels for both low and high magnetic activity are shown in order to separate the components due to magnetic storms from those due to variations in the solar flux forcing function. It is clear that even with low magnetic activity the year of high solar flux shows a dramatic increase in the intensity and in the occurrence of ionospheric irregularities which produce scintillations on a satellite to ground path.

IRREGULARITY MOTIONS

To investigate the irregularity characteristics more fully an experiment was conducted using spaced antenna/receivers at Thule, Greenland during March 1980. The measurement involved calculation of the cross correlation function for three spaced antenna/receivers positioned in a triangular pattern with separation of approximately 300 meters. The velocity of the irregularity was determined by the difference in time of arrival of the fading pattern at each antenna.

Cross correlation values of from .48 to .98 were calculated with differences in time of arrival from 0.1 to 6.0 seconds. The corresponding irregularity velocities ranged from 40 to 800 meters/second with an average velocity of from 300 to 500 meters/second.

During these measurements, the maximum solar depression angle (at local midnight) was 11°, thus the altitude region above 120 kilometers was sunlit for the entire measurement period. A plot of the irregularity drift velocity in corrected geomagnetic latitude - local time shows a consistent portion of anti-sunward drift, Figure 5. This pattern is consistent with the expected ExB plasma drift (assuming that the ionospheric irregularities move with the background plasma) obtained from sunlit polar cap electric field measurements (Heppner, 1972).

An empirical two cell irregularity drift model proposed by Heppner (1977), postulated an anti-sunward drift from the noontime auroral oval over the magnetic pole to the midnight sector and a circulation back around the auroral oval to the noon sector. The observed irregularity drift supports the model. Consistent changes of drift direction were noted at 0800 and 2200 corrected geomagnetic time. The highest velocities were recorded during the noon and midnight periods.

AURORAL ARCS

Auroral arcs in the polar cap are approximately aligned with the noon-midnight magnetic meridian (Davis, 1962). These arcs generally drift in the dawn to dusk direction (Danielsen, 1969), however, reversals have been noted (Akasofu, 1972; Weber and Buchau, 1981). Recently Weber and Buchau (1981) described the orientation and motion of subvisual F-layer (λ=6300 Å OI) polar cap arcs. Kilometer-

size irregularities within the a. ›duced intense (saturated) amplitude scintillation at 250 MHz as the arcs drifted through a satellite to ground ray path Outside the arcs, however, scintillation frequently did not cease, but persisted at a lower level (SI ∿ 6 dB). This effect is also noted in the longer term data recorded on the ground at Thule during 1979-1980.

EVIDENCE OF TWO IRREGULARITY COMPONENTS

Results from the combined measurements point to two irregularity components in the polar cap; antisunward drifting irregularities which produce a background level of weak to moderate scintillation and intense irregularities within F layer polar cap arcs which produce more discrete (∿ 1 hr duration) intense scintillation events as the arcs drift through the ray path.

Scintillation activity is proportional to $<(\Delta N)^2>$, the mean square electron density deviation of the small scale irregularities responsible for fading, and the thickness of the irregularity layer If during years of high solar flux the ambient electron density is high, a small disturbance (20% change, for example) in electron density would produce a high value of ΔN and therefore (depending on layer thickness) intense scintillation activity. Correspondingly the same percentage perturbation during years when electron density is low would produce lower scintillation levels.

Monthly median values of foF2 at Thule during 1957 (a year of high solar flux) and 1963 (a year of low solar flux) support this hypothesis. The measured values are shown in Table 1.

There is a significant change for the month of December between years of high and low solar flux. The corresponding maximum electron density, $Ne_{(max)}$, at 00 LT changes from 4.4×10^5 to 4.9×10^4 el cm^{-3} between 1957 and 1963. The large variation in scintillation occurrence for the same season between years of high and low solar flux may thus be related to corresponding large changes in $Ne_{(max)}$. There is little change between December and June for 1957/58. Thus the seasonal variation of scintillation is not explained by the seasonal variation of foF2.

The seasonal variation of scintillation may, however, be related to E layer conductivity changes caused by the presence or absence of sunlight at ∿100 km. As proposed by Heppner (1972) to explain fluctuating E fields in the winter polar cap and more regular variations in the summer polar cap, the conducting E layer may tend to short circuit irregularities which can otherwise persist in the F layer. During the summer the short circuiting by the E layer could account for the low level of activity. This summer-winter asymmetry in the electric field has also been noted by Burke et al. (1979). In addition, polar cap electron density fluctuations measured by the AE-D satellite show higher fluctuation levels in the winter hemisphere as compared to the summer hemisphere during a 3-month period of low solar activity (Sunanda Basu, private communication).

In the auroral oval, both discrete and diffuse aurora as shown by Defense Meteorological Satellite Program images have been correlated with scintillations (Martin and Aarons, 1977). The evidence for this was also clear when specific days were examined in the time period January 1979 DMSP auroral images were compared with recordings of various satellite beacons observed from Goose Bay, Labrador (65° C.G.L.). The E and F layer irregularities in the auroral oval, which produce scintillations of UHF radio waves, are correlated with the optical auroral emissions detected by the DMSP imaging system. This is not the case for the polar cap scintillation activity; no corresponding polar cap optical emission was found from DMSP photographs.

As shown by associated measurements (Weber and Buchau, 1981) F layer auroral arcs with orientation along the noon-midnight direction are responsible for the intense radio wave fading. Using the DMSP sensitivity, determined by Eather (1979), the spectral characteristics of these auroras, < 5 R 4278 Å N_2^+ and 350 R 6300 Å (OI), (J.G. Moore, private communication, 1980) would not produce a detectible signal in the DMSP auroral images.

CONCLUSION

The existence of irregularities at F layer heights has been known from spread F measurements. However, we are now able to distinguish the enormous difference in intensity level between years of low and high solar flux and between months of sunlit and months of dark periods.

ACKNOWLEDGMENTS

The authors would like to thank Aage Paulsen and Fried Nielsen of the Danish Arctic Contractors for their assistance in taking the Thule data, the EnviroData Corporation for assistance in reducing the data, and, in addition, the AFWAL Test Team who assisted in the drift measurements at Thule and the 4950th Flight Test Wing who operated the aircraft used in the data collections. In particular the authors would like to thank Eileen MacKenzie for her work in the organization and analysis of the material. This research was supported in part by the Defense Nuclear Agency under Subtask Code I25AAXHX640, Work Unit 02.

REFERENCES

Akasofu, S.-I., Midday auroras and polar cap auroras, Geophy. Publ., 29, 73-85, 1972.

Burke, W.J., M.C. Kelley, R.C. Sagalyn, M. Smiddy and S.T. Lai, Polar cap electric field structures with a northward interplanetary magnetic field, Geophys. Res. Lett., 6, 21, 1979.

Danielsen, C., Auroral observations at Thule 1961-65, Danish Met. Inst., Geophys. Papers R-9, 1969.

Davis, T.N., The morphology of the auroral displays of 1957-1958. (2) Detailed analysis of Alaska data and analysis of high latitude data, J. Geophys. Res., 67, 75-110, 1962.

Eather, R.H., DMSP calibration, J. Geophys. Res., 84, 4134, 1979.

Heppner, J.P., Electric field variations during substorms: OGO-6 measurements, Planet. Space Sci., 20, 1475, 1972.

Heppner, J.P., Empirical models of high-latitude electric fields, J. Geophys. Res., 82, 1115, 1977.

Martin, E. and J. Aarons, F-layer scintillation and the aurora, J. Geophys. Res., 82, 2717, 1977.

Weber, E.J. and J. Buchau, Polar cap F-layer auroras, Geophys. Res. Letters, 1981 (in press).

Whitney, H.E., J. Aarons and C. Malik, A proposed index for measuring ionospheric scintillations, Planet. Space Sci., 11, 1069, 1969.

Table 1. Thule Monthly Median foF2

	High Solar Flux		Low Solar Flux
	December 1957		December 1963
00 LT	6.0 MHz	Winter	2.0 MHz
12 LT	7.5 MHz		3.0 MHz
	June 1958		June 1964
00 LT	5.6 MHz	Summer	4.0 MHz
12 LT	5.4 MHz		4.2 MHz

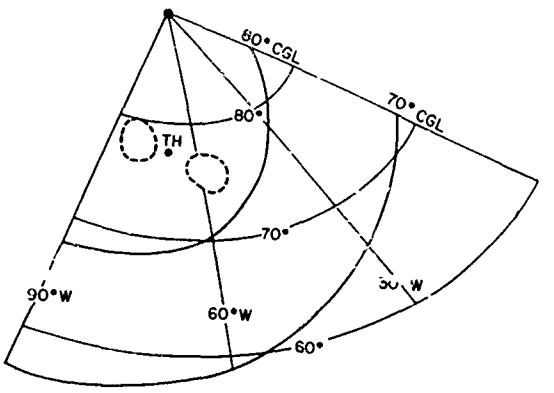


Fig. 1. Map of sky region around Thule with corrected geomagnetic and geographic latitudes shown. Observations are available in the subionospheric intersection (350 km) area encompassed by the dotted lines.

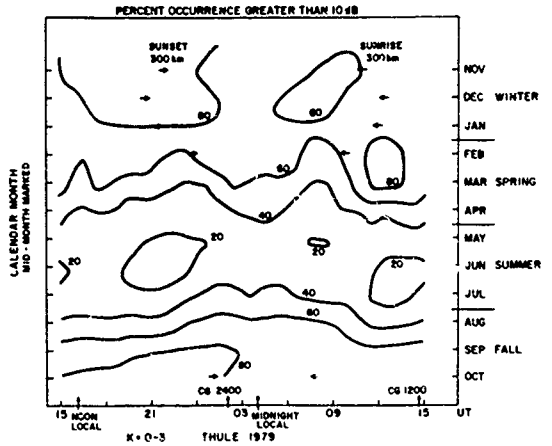


Fig. 2. Contour plot of diurnal pattern of monthly percent occurrence of scintillation greater than 10 dB for low magnetic activity (Kp = 0-3) periods. Observations were taken during March 1979 - February 1980.

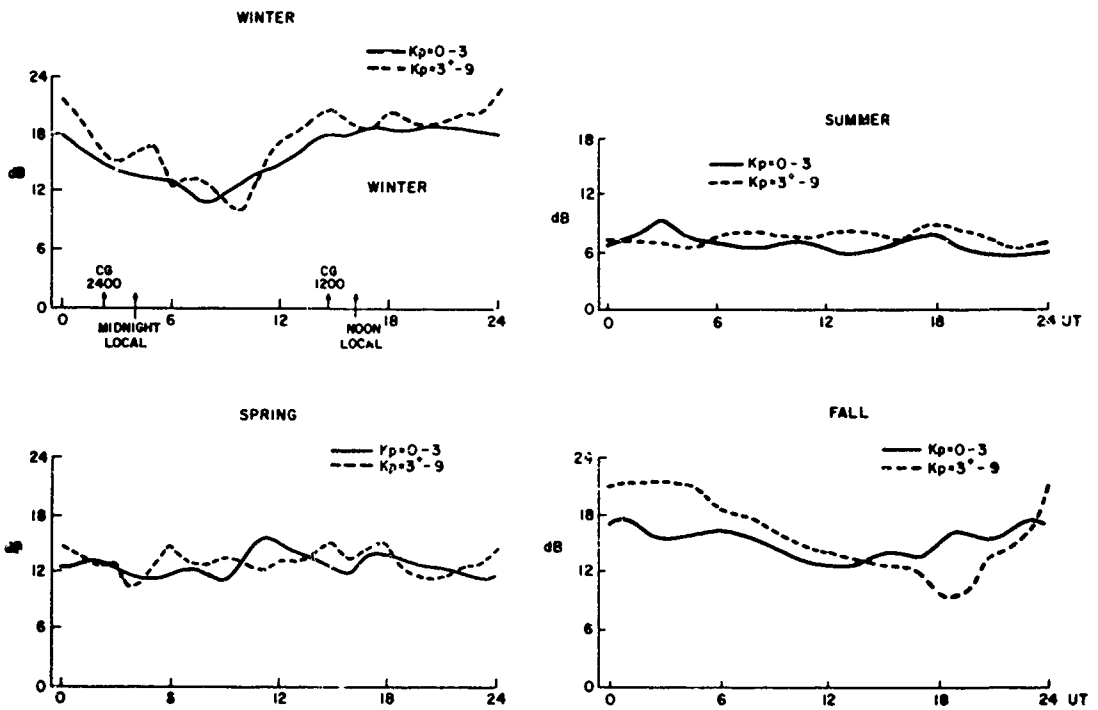


Fig. 3. Mean value of scintillation excursion observed at Thule for both quiet and disturbed magnetic activity periods. The pattern is given seasonally as: (a) Winter (Nov-Jan), (b) Spring (Feb-Apr), (c) Summer (May-Jul), (d) Fall (Aug-Oct)

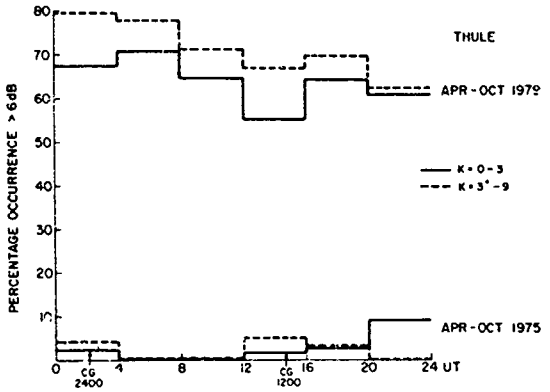


Fig. 4. Percentage occurrence of scintillation greater than 6 dB for low solar flux period April-October 1975 is contrasted with that for high solar flux period April-October 1979 for both quiet and disturbed magnetic conditions.

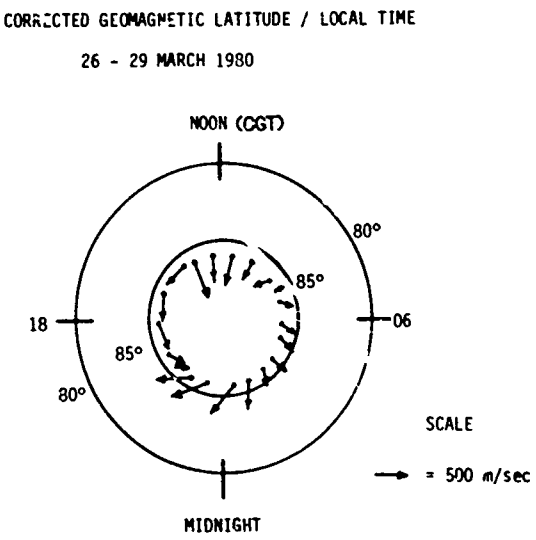


Fig. 5. Ionospheric irregularity flow in the arctic in corrected geomagnetic latitude/local time determined from a spaced receiver experiment at Thule in March 1980.

The Current Convective Instability and Its Relation to Diffuse Auroral Scintillation Causing F Region Irregularities

M. J. Keskinen,[+] P. K. Chaturvedi,[++] S. L. Ossakow,[+] and B. E. McDonald[+,*]

[+]Naval Research Laboratory, Washington, D.C. 20375

[++]Berkeley Research Associates, Springfield, VA 22150

* Present Address: Naval Oceanographic Research and Development Activity Bay St. Louis, Mississippi 39520

ABSTRACT

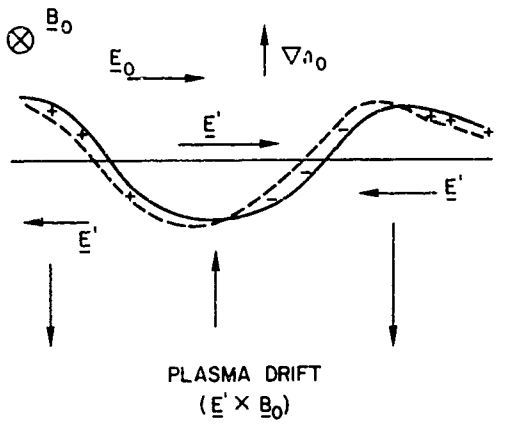
DNA Wideband satellite data have given evidence that F region irregularities causing high latitude scintillation enhancements are observed in regions with well defined north-south TEC gradients (as well as north-south gradients in plasma density) and magnetic field-aligned diffuse auroral particle precipitation. In addition these irregularities appear to be L-shell aligned localized sheet-like structures for wavelengths ~ 1km. Our present theoretical understanding indicates that the current convective instability (CCI), driven by weak magnetic field-aligned currents and a plasma density gradient, can directly result in long wavelength scintillation causing field aligned ($k_\parallel \ll k_\perp$, where k is wavenumber and $\parallel$ and $\perp$ denote parallel and perpendicular to the ambient geomagnetic field, respectively) F region ionospheric irregularities in the diffuse auroral region. Numerical simulations of the CCI show equatorward movement of plasma density enhancements, poleward movement of plasma depletions (holes), and a k^{-2} irregularity spatial power spectrum.

INTRODUCTION

Data from the DNA Wideband satellite have exhibited high latitude (auroral and subauroral) scintillation enhancements which appear to be associated with north-south gradients in total electron content [TEC; see Fremouw et al., 1977; Rino et al., 1978; Rino and Matthews, 1980; Rino and Owen, 1980]. Rino et al. [1978] have shown that a regularly occurring scintillation enhancement can be identified in the nighttime auroral zone data, in the region of diffuse auroral particle precipitation. It is also shown that this enhancement is due to sheet like F region ionospheric irregularities. These irregularities occur near where the TEC gradient points northward, i.e., increasing TEC as one goes northward. In addition, it appears that a d.c. electric field points either westward or northwest and that there is a very shallow plasma density gradient in altitude, i.e., the northward TEC gradient dominates [Vickrey et al., 1980]. Associated with the northward TEC gradient is a northward gradient in local F region plasma density [Rino and Owen, 1980; Vickrey et al., 1980] as measured by the Chatanika radar.

At first glance, with the dominant TEC gradient pointing northward (also local plasma density gradient), the ambient magnetic field, $\underline{B}_o$, pointing down and the d.c. electric field horizontal, this would appear to be a prime geometry for the usual F region $\underline{E} \times \underline{B}$ gradient drift instability. However, with the d.c. electric field pointing westward or northwest, the configuration is stable (see Fig. 1; the electric field would have to point eastward for instability). The saving feature, however, is the diffuse auroral precipitation (current) which acts to destabilize the above geometrical configuration [Ossakow and Chaturvedi, 1979] (see Fig. 1). The conditions of having a current along $\underline{B}_o$, a density gradient, ∇n_o, perpendicular to $\underline{B}_o$, and an electric field perpendicular to $\underline{B}_o$ and possibly ∇n_o can result in instability and these types of instabilities are generically called current convective instabilities [Kadomtsev, 1965].

USUAL F REGION $\underline{E} \times \underline{B}$ INSTABILITY PICTURE



F REGION $\underline{E} \times \underline{B}$ GEOMETRY WITH FIELD ALIGNED CURRENT (CURRENT CONVECTIVE INSTABILITY)

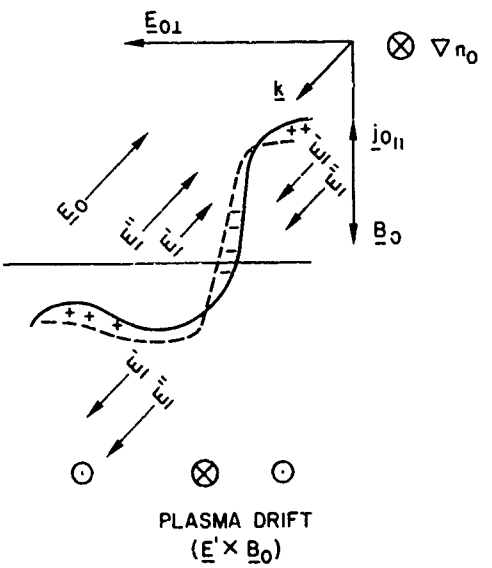


Fig. 1. Simplified physical picture of the $\underline{E} \times \underline{B}$ gradient drift instability and current convective instability.

This type of instability has been applied to the positive column of laboratory gas discharges [Lehnert, 1958; Hoh and Lehnert, 1960; Kadomtsev and Nedospasov, 1960; Kadomtsev, 1965], but to our knowledge has not been applied to the conditions in the F region ionosphere. A simple physical picture of the instability is as follows (see Fig. 1). First let us discuss the usual F region E x B gradient drift instability picture. In the upper part of Figure 1 the horizontal line represents an unperturbed contour of constant electron density, the background zero order electron density gradient points upward (northward), there is a horizontal d.c. electric field (eastward), and the ambient magnetic field, $\underline{B}_o$, is into the picture. Now let the density be perturbed by a small amplitude horizontal sinusoidal variation ($\underline{k}$ parallel to $\underline{E}_o$). The ions Pedersen drift to the right relative to electrons whose Pedersen drift is for all purposes negligible in the F region (the dashed sinusoid represents the electrons, the solid the ions). This gives rise to space charges (+ and -) which in turn cause small scale electric fields $\underline{E}'$, directed alternately to the left and right. In this geometry (eastward $\underline{E}_o$), the corresponding $\underline{E}' \times \underline{B}_o$ drifts will then carry the enhanced regions downward (southward) and the depleted regions upward (northward) so that they both will appear to grow in amplitude against the background density (convective instability). If $\underline{E}_o$ were to point westward, as in the diffuse auroral case, the perturbations would disappear, i.e., the situation would become stable (also if ∇n_o were reversed in Fig. 1 stability would result).

Now let us discuss the $\underline{E} \times \underline{B}$ geometry with a magnetic field aligned current as illustrated in the lower part of Figure 1. There is an ambient electric field, $\underline{E}_{o\perp}$, perpendicular to $\underline{B}_o$ (pointing down) and pointing to the left (westward), the density gradient, ∇n_o, is into the picture (northward), the current, $j_{o\parallel}$, is anti-parallel to $\underline{B}_o$. the horizontal line across the perturbation (where $\underline{k}$ is in the $\underline{E}_{o\perp} - \underline{B}_o$ plane and perpendicular to ∇n_o) represents the unperturbed constant electron density contour. In this picture without $j_{o\parallel}$ the system would be stable to the $\underline{E} \times \underline{B}$ instability (see previous paragraph). The projection on $\underline{k}$ of the ion Pedersen drift caused by $\underline{E}_{o\perp}$ results in a drift of the ions along $\underline{k}$ such that it would result in stability. However, the direction of $j_{o\parallel}$ implies that the relative drift between ions and electrons (in the frame where the electrons are at rest) is anti-parallel to $\underline{B}_o$. This motion projected on $\underline{k}$ results in establishing space charges which give rise to small scale electric fields $\underline{E}''$ as shown in Fig. 1 (note that the Pedersen motion due to $\underline{E}_{o\perp}$ results in small scale electric fields opposed to $\underline{E}''$). If the particle motion projected on $\underline{k}$ is dominated by the $j_{o\parallel}$ rather than $\underline{E}_{o\perp}$ then the space charges give rise to total small scale electric fields, $\underline{E}'$, as depicted in the figure. The corresponding $\underline{E}' \times \underline{B}_o$ drifts will then carry enhanced regions out (southward) of the figure and depleted regions into (northward) the figure so that they both will appear to grow in amplitude against the background density. From this picture the instability criterion is $|\underline{k} \cdot \underline{V}_d| > |\underline{k} \cdot (c\underline{E}_{o\perp}/B_o)(\nu_i/\Omega_i)|$, where $j_{o\parallel} = n_o e \underline{V}_d$ and ν_i and Ω_i are the ion-neutral collision frequency and ion gyrofrequency, respectively. This will be mathematically derived in the next section. The lower figure can be made similar to the upper figure by noting that $j_{o\parallel} = n_o e \underline{V}_d = \sigma_\parallel \underline{E}_{o\parallel}$ and the setting $\underline{E}_o = \underline{E}_{o\perp} + \underline{E}_{o\parallel}$. We note that in the above picture making $j_{o\parallel}$ in the direction of $\underline{B}_o$ or rotating $\underline{k}$ by 90° about $\underline{E}_{o\perp}$ or $\underline{B}_o$ results in stability. We also note that the above picture for instability is valid even for $\underline{E}_{o\perp} = 0$. If $\underline{E}_{o\perp}$ were reversed in direction the parallel current would enhance an already unstable situation.

This instability also has $k_\parallel \ll k_\perp$ (see next section) and so the irregularities generated will be field aligned. This current convective instability can directly result in long wavelength scintillation causing F region ionospheric irregularities in the diffuse auroral region.

THEORY

In this section we present a theoretical model [Ossakow and Chaturvedi, 1979] for a long wavelength fluid type plasma instability which may account for the scintillation causing diffuse auroral F region ionospheric irregularities. For our model we take the electron density gradient to be pointing northward (y), the ambient electric field $\underline{E}_o$ is in the westward direction (x) and the magnetic field points downward (z). In our simple model we have equated the TEC gradient with a gradient in plasma density. We assume that the horizontal (northward) electron density gradient is much sharper than the altitude density gradient, which we neglect. Our basic equations are

$$\frac{\partial n_\alpha}{\partial t} + \nabla \cdot (n_\alpha \underline{v}_\alpha) = 0 \tag{1}$$

$$\underline{V}_i = \frac{c}{B_o} \underline{E} \times \hat{z} - \frac{c}{B_o} \frac{\nu_i}{\Omega_i} \underline{E}_\perp + \frac{c}{B_o} \frac{\Omega_i}{\nu_i} \underline{E}_\parallel + \underline{V}_{io\parallel} \tag{2a}$$

$$\underline{V}_e = \frac{c}{B_o} \underline{E} \times \hat{z} - \frac{c}{B_o} \frac{\Omega_e}{\nu_e} \underline{E}_\parallel + \underline{V}_{eo\parallel} \tag{2b}$$

$$\nabla \cdot \underline{j} = 0, \quad \underline{j} = \Sigma n_\alpha q_\alpha \underline{V}_\alpha \tag{3}$$

where the subscript α is the species label (e is electron, i is ion), n is density, $\underline{V}$ is velocity, ν is collision frequency, q is charge, E is the electric field, B_o is the ambient magnetic field, $\hat{z} = \underline{B}_o/|B_o|$, $\Omega_\alpha = |q_\alpha| B_o/m_\alpha c$, $\underline{j}$ is the current and $\underline{V}_{\alpha o}$ represents the diffuse auroral particle precipitation velocity along $\underline{E}_o$ ($\|$ and $\perp$ denote parallel and perpendicular to $\underline{B}_o$) which results in a zero order current. In the momentum transfer eqn. (2), we have neglected inertial and temperature effects, and included electron collisions to first order for completeness. Equation (2) is valid for F region ionosphere altitudes ($\nu_\alpha/\Omega_\alpha \ll 1$). In (2a) ν_i is taken to mean ion-neutral collisions, whereas in (2b) ν_e is really electron-ion collisions. We have neglected the electron Pedersen drift compared with the ion Pedersen drift.

We assume quasi-neutrality so that $n_e \approx n_i \approx n$ and our final equations are the electron continuity equation, electron and ion momentum equations, and $\nabla \cdot \underline{j} = 0$. The equations are then linearized such that $n = n_o(y) + \tilde{n}$, $\underline{E} = E_o\hat{x} - \nabla\tilde{\phi}$, $\underline{V}_\alpha = \underline{V}_{\alpha o} + \tilde{\underline{V}}_\alpha$ with the perturbed quantities $\tilde{n}$, $\tilde{\underline{V}} \propto \exp i[k_x x + k_\| z - \omega t]$, where $\omega \equiv \omega_r + i\gamma$. We then obtain

$$\gamma = \frac{-\frac{1}{n_o}\frac{\partial n_o}{\partial y}\left[\frac{c}{B} E_o \frac{\nu_i}{\Omega_i} + V_d \frac{k_\|}{k_x}\right]}{\left[\frac{\Omega_i}{\nu_i} + \frac{\Omega_e}{\nu_e}\right]\frac{k_\|^2}{k_x^2} + \frac{\nu_i}{\Omega_i}} \tag{4}$$

where $V_d = V_{io\|} - V_{eo\|}$. We see that in (4) γ is independent of $|k|$ and only depends on the angle that k makes with $\underline{B}_o$. In the denominator of (4), the first term (in brackets) multiplying $(k_\|/k_x)^2$ comes from the parallel motion of the ions and electrons; whereas, the remainder of the denominator comes from the ion Pedersen motion. It should be noted that the instability is essentially unaffected by the current direction. Thus, downward currents work just as well as upward currents. In (4) if we set $k_\| = 0$ we obtain $\gamma = -(n_o^{-1}\partial n_o/\partial y)(cE_o/B_o)$ which is the usual result for the $\underline{E} \times \underline{B}$ gradient drift instability. For our geometry this shows γ is negative which implies stability. For instability in (4) $\gamma > 0$ which implies

$$\frac{c}{B_o} E_o \frac{\nu_i}{\Omega_i} + V_d \frac{k_\|}{k_x} < 0$$

This says that for instability, with the westward $\underline{E}_o$, we must have $V_d k_\|/k_x < 0$ and $|V_d k_\|| > (k_x cE_o/B_o)(\nu_i/\Omega_i)$, which is exactly the condition set forth on the basis of our physical picture presented in the first section. We can maximize this growth with respect to $\theta \equiv k_\|/k_x$. From (4) the growth rate maximizes, in general, for

$$\theta = -(cE_o/B_o V_d)(\nu_i/\Omega_i) \pm \{(cE_o/B_o V_d)^2(\nu_i/\Omega_i)^2 + (\nu_i/\Omega_i)[(\Omega_i/\nu_i) + (\Omega_e/\nu_e)]^{-1}\}^{1/2}$$

For altitudes $\sim$ 350-400 km corresponding to the observation altitudes [Rino et al., 1978] $\nu_i/\Omega_i \sim 10^{-4}$, $\nu_e/\Omega_e \sim 10^{-4}$ and $E_o \sim 10$ mV/m with $V_d \sim -500$ m/sec we find $\theta = 9.4 \times 10^{-5}$. For this case (4) gives a growth rate $\gamma_{max} \simeq 2.7 \times 10^{-3}\text{sec}^{-1}$ for $L \equiv n_o(\partial n_o/\partial y)^{-1} \sim 50$ km. Including pressure effects in the problem introduces diffusive damping in (4). A typical cross-field diffusion coefficient, $D_\perp$, is ~ 0.2 m^2/sec and a parallel diffusion coefficient, $D_\|$, is $\sim 10^8$m^2/sec. In the present study, these effects become important for perpendicular wavelengths, $\lambda_\perp$, $\lesssim 100$ m and parallel wavelengths, $\lambda_\|$, $\lesssim 1000$ km. However, typical scintillation causing perpendicular wavelengths are ~ 1 km and since $k_\|/k_\perp \sim 10^{-4}$ we are considering highly field aligned irregularities. Larger parallel currents, due to precipitation, will of course produce larger growth rates. However, too large a current, $V_d \gtrsim 1$ km/sec, would excite the collisional electrostatic ion cyclotron instability [Chaturvedi, 1976]. It may be noted that the linear theory of the current convective instability proposed here favors a wavevector perpendicular to the TEC gradient (as well as $\underline{B}_o$). However, in the nonlinear regime, mode coupling theory suggests [Chaturvedi and Ossakow, 1979] that the dominant wavevectors will lie parallel to the TEC gradient (and local F region plasma density gradient) thus accounting for the L-shell aligned nature of these irregularities [Rino et al., 1978; Rino and Owen, 1980]. Basically in this theory, the current convective instability saturates (stabilizes) nonlinearly by generating linearly damped harmonics (i.e., those along the northward plasma density gradient). These harmonics dominate over those between the northward and east-west directions. In addition, these northward modes produce a power spectrum for the irregularities $\propto k^{-2}$.

NUMERICAL SIMULATION RESULTS

Equations (1)-(3) can be rewritten [Keskinen et al., 1980]

$$\frac{\partial n}{\partial t} + \frac{c}{B}\hat{z} \times \nabla\varphi_1 \cdot \nabla n - \frac{c}{B}\nabla \cdot n\left(\frac{\nu_i}{\Omega_i}\nabla_\perp\varphi_1 + \hat{z}\frac{\Omega_i}{\nu_i}\frac{\partial\varphi_1}{\partial z}\right) = 0 \tag{5}$$

$$\nabla \cdot n\left[\nabla_\perp\varphi_1 + \hat{z}\left(\left(\frac{\Omega_i\Omega_e}{\nu_i\nu_e}\right)^2 + \left(\frac{\Omega_i}{\nu_i}\right)^2\right)\frac{\partial\varphi_1}{\partial z}\right] = \left(\underline{E}_o - \frac{\Omega_i}{\nu_i}\frac{B}{c}\underline{V}_d\right)\cdot\nabla n \tag{6}$$

where $\nabla\varphi = \underline{E} - \underline{E}_o$ with $\underline{E}(x,y,z,t)$ the total electric field. Equations (5) and (6) are simply a restatement of the ion-continuity equation written in a reference frame with velocity $\underline{V}_o = -E_o/B(\hat{y} - \nu_i/\Omega_i\,\hat{x})$ and quasi-neutrality $\nabla\cdot\underline{J} = 0$, respectively.

In the following simulations we take advantage of the fact that the fastest growing, most dangerous modes from linear theory are almost field-aligned ($k_\parallel/k_\perp \ll 1$). These modes are of most interest to us and, as a result, we solve equations (5) and (6) in a plane containing these modes which is nearly perpendicular to the magnetic field while fixing the value of $k_\parallel/k_\perp$ [Keskinen et al., 1980]. A similar approach has been adopted in numerical studies of drift-wave [Lee and Okuda, 1976] and trapped-particle [Matsuda and Okuda, 1976] instabilities in laboratory plasmas.

The simulation plane which is essentially horizontal at an altitude of 350 km with a north-south extent of 410 km and an east-west extent of 160 km is identical to the $x'y'$ plane as shown in Figure 2. The system of equations (5) and (6) was transformed to the $x'y'z'$ coordinate system by a simple rotation about the y-axis by the angle $\theta = k_\parallel/k_x \ll 1$. By solving the equations (5) and (6) in the $x'y'z'$ system a finite but small $k_\parallel$ is effectively introduced into the model. After neglecting the z'-dependence of all quantities, equations (5) and (6) were then cast into dimensionless form and initialized with the profile of the following type $n_o(y')/N_o = (1-A(1-\tanh(y'-y_o)/L))(1 + \epsilon(x',y'))$ where N_o, y_o, and L are constant with A = 5/11. This gives a total density maximum to minimum ratio of approximately 10. The function $\epsilon(x',y')$ has a root-mean-square value of 3% and is generated from a randomly phased Gaussian power spectrum. The computational mesh consisted of 258 grid points in the y'-direction (north-south) with 102 points in the x'-direction (east-west) so that $\Delta y = \Delta x = 1.6$ km. Periodic boundary conditions were imposed in the x'-direction with Neumann ($\partial/\partial y' = 0$) boundary conditions in the y'-direction. We now drop the prime notation for clarity.

Figures 3a-d show contour plots of $n(x,y)/N_o$ at t = 0, 900, 1400, 1900 sec. The following set of parameters were used: L = 50 km, y_o = 200 km, E_o = 10 mV/m, $\nu_i/\Omega_i = 10^{-4}$, $\nu_e/\Omega_e = 10^{-4}$. The value of $\theta = 9.4 \times 10^{-5}$ is held fixed and is chosen so as to maximize the linear growth rate (4). Figure 3a shows the initial configuration which includes the small random perturbation. Figure 3b at t = 900 sec ($\gamma_k t \simeq 5$) illustrates the linear stages of the simulation and shows unstable growth in the region where $\partial n_o/\partial y > 0$ as predicted by linear theory. Figure 3c exemplifies the early nonlinear regime where lower density plasma (depletions) are moving in the positive y-direction (poleward) while higher density plasma (enhancements) are convecting in the negative y-direction (equatorward). The approximate velocities of the depletions and enhancements are 270 m/sec and 30 m/sec, respectively. Finally, well-developed steepened enhancements and depletions (of over 90%) are seen in Fig. 3d at t = 1900 sec. This late-time configuration is reminiscent of the motion of depletions (bubbles) moving vertically in the equatorial F region [Scannapieco and Ossakow, 1976] and enhancements (striations) in ionospheric F region plasma clouds [Scannapieco et al., 1976]. The length scales in Figures 3a-d are distorted with the depletions and enhancements longer and narrower than is depicted. Similar linear and nonlinear development is observed when L = 10 km, but on a faster time scale.

Figures 4a-b give sample one-dimensional spatial power spectra at t = 1900 sec both in the x-direction (east-west) and in the y-direction (north-south). These spectra are obtained by first Fourier analyzing $\delta n(x,y)/N_o$ and integrating over the direction in $\underline{k}$-space corresponding to the north-south and east-west directions, respectively. For both cases, these power spectra are well-fitted with an inverse power law. Similar power law dependences were seen when L = 10 km.

The following physical picture of the evolution of current convective instability in the diffuse aurora is supported by these simulations. In the evening a westward electric field $\underline{E}_o$ begins to form which convects plasma in the auroral region equatorward. In regions where the northward gradient in total electron content becomes well-defined nearly field-aligned fluctuations ($k_\parallel/k_\perp \ll 1$) will grow unstable in regions where the field-aligned current velocities V_d, caused by precipitating particles, are such that $V_d(k_\parallel/k_\perp) > 0$ and $|V_d k_\parallel| > (k_\perp c E_o/B)(\nu_i/\Omega_i)$. In the plane

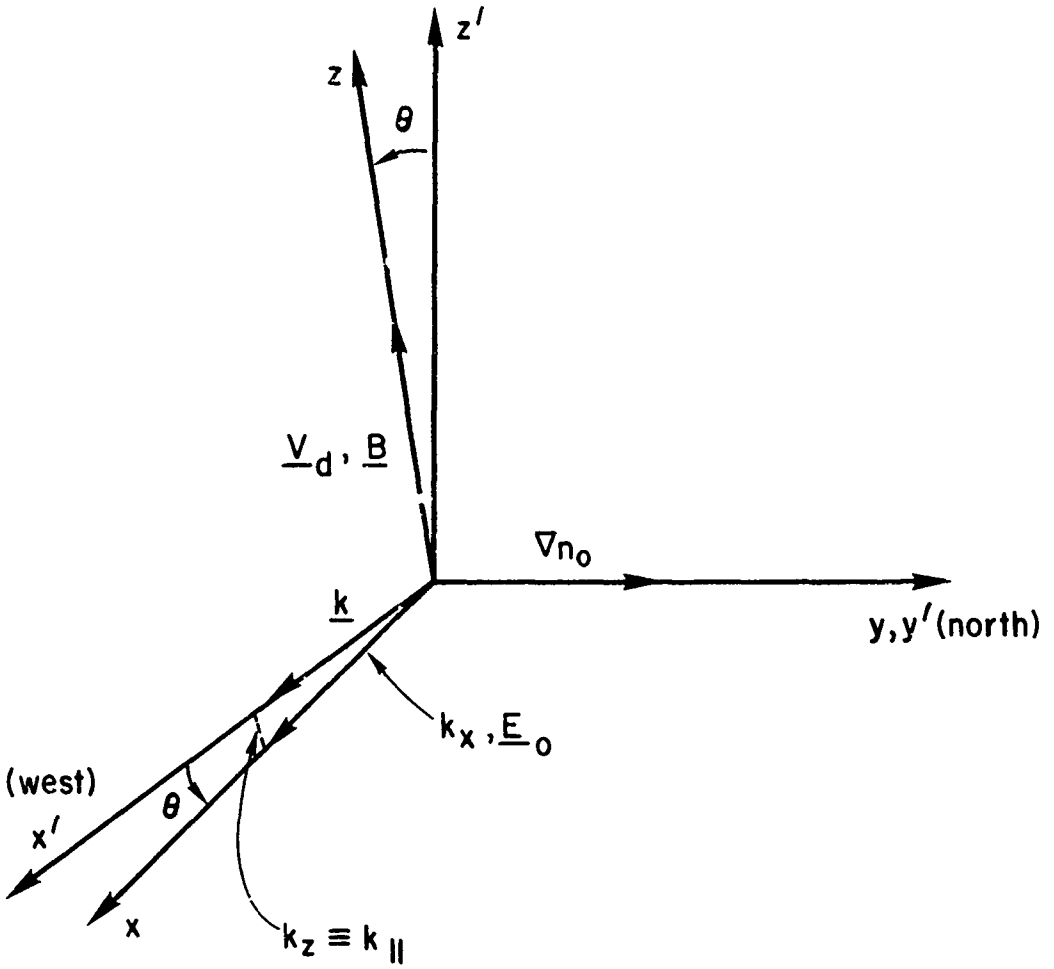


Fig. 2. Coordinate system used in simulations. The x'y' plane is the simulation plane. The x', x, z', z axes are coplanar.

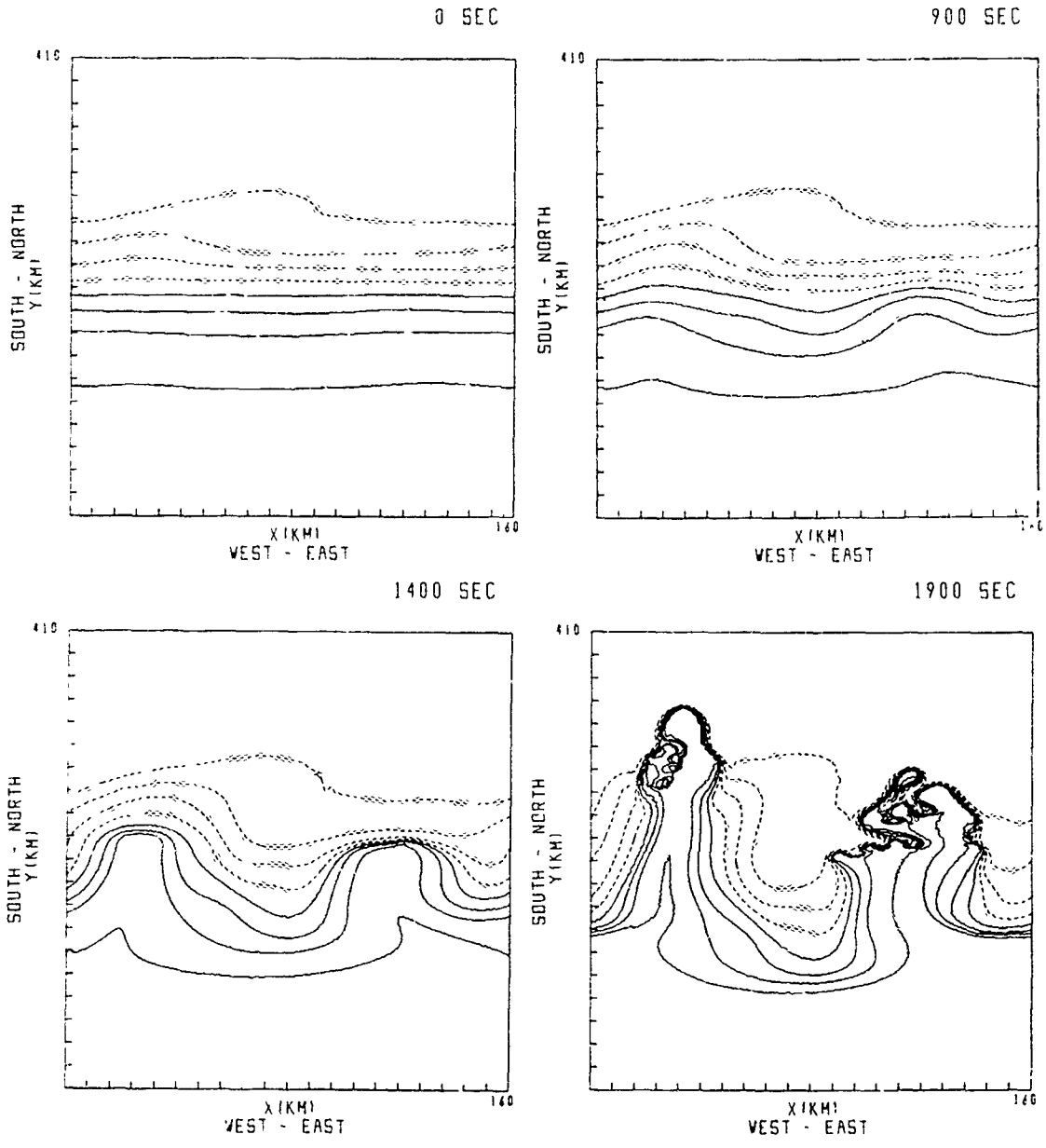
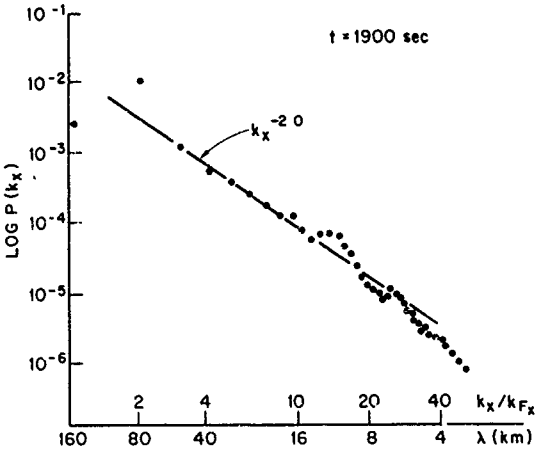
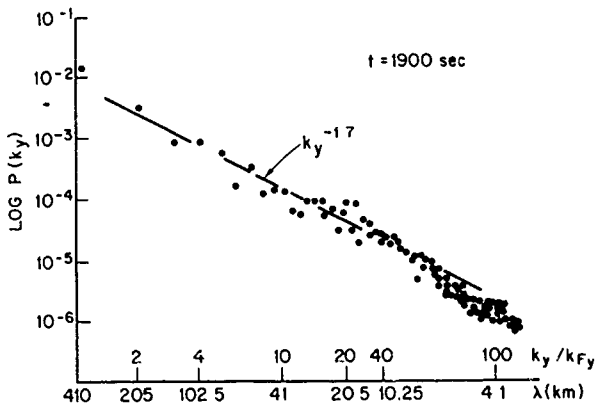


Fig. 3. Real space isodensity contour plots of $n(x,y)/N_o$ for L = 50 km at (a) t = 0 sec, (b) t = 900 sec, (c) t = 1400 sec, (d) t = 1900 sec. Eight contours are plotted in equal increments from 1 to 10 with the lower (higher) density contours denoted by solid (dashed) lines. The magnetic field $\underline{B}$ is directed into the page with the observer looking down the magnetic field lines toward the earth.



(a)



(b)

Fig. 4. Log-log plots of one dimensional (a) x power spectra $P(k_x)$ and (b) y power spectra $P(k_y)$ for L = 50 km at t = 1900 sec. $P(k_x)$ and $P(k_y)$ are obtained by averaging $|\delta n(k_x,k_y)/n_o|^2$ over k_y and k_x, respectively. In (a) $k_{Fx} = 2\pi/160$ km^{-1} while in (b) $k_{Fy} = 2\pi/410$ km^{-1}. The dots represent the numerical simulation results while the solid line is obtained from a least squares fit.

almost perpendicular to the magnetic field by an angle $\theta = k_{\parallel}/k_{\perp}$ plasma depletions and enhancements will move northward and southward, respectively, while steepening in the process.

SUMMARY

We have investigated analytically and numerically a simple plasma fluid model to account for the diffuse auroral scintillation causing F region ionospheric irregularities observed by the DNA Wideband satellite [Rino et al. 1978]. By taking account of the diffuse auroral particle precipitation (current) the stable $\underline{E} \times \underline{B}$ diffuse auroral geometry (corresponding to the observations) becomes destabilized by this parallel current. For a westward ambient d.c. electric field, $\underline{E}_o$, and a northward dominant electron density gradient, the relative drift velocity between ions and electrons parallel to B_o, V_d, must satisfy the condition $-\underline{k}_{\parallel} \cdot \underline{V}_d > \underline{k}_{\parallel} \cdot (c\underline{E}_o/B_o)\ (\nu_i/\Omega_i)$ for instability. The maximum growth rate for the instability is $\gamma \approx n_o^{-1}(\partial n_o/\partial y)\ V_d[1 + (\Omega_e\nu_i/\Omega_i\nu_e)]^{-1/2}/2$. The instability is mainly field aligned ($k_{\parallel} \ll k_{\perp}$). The instability is fluid-like in nature and so can directly account for the long wavelength diffuse auroral scintillation causing F region irregularities. In addition, nonlinear numerical simulations of this instability indicates equatorward movement (convection) of plasma enhancements and northward (poleward) movement of plasma depletions (holes). Furthermore, one-dimensional spatial power spectra of the irregularities in both the north-south and east-west directions are well described by inverse power laws $\propto k^{-2}$, for $\lambda \geq 4$ km (the present wavelength range over which the simulations have been performed).

In addition to investigating the nonlinear mode coupling theory [Chaturvedi and Ossakow, 1979] for the current convective instability, the linear theory has been extended to include other effects. For example, the inclusion of magnetic shear effects [Huba and Ossakow, 1980] reduces the linear growth rate somewhat (but does not stabilize the mode) and localizes the mode structure parallel to the northward plasma density gradient. Ion inertial effects, which extend the instability to higher F region altitudes, highly collisional (non-magnetized) ions (e.g., the E region), and electromagnetic effects (which can be important for long wavelengths) have recently been included [Chaturvedi and Ossakow, 1981] in the extended linear theory. The basic results found by Chaturvedi and Ossakow [1981] were: including ion inertia extends the instability to high altitudes with growth rates comparable to those achieved at the lower F region altitudes in the more collision dominated regime (as given in the present paper); in the case where the ions are highly collisional, e.g., the E region, the growth rates are comparable to F region growth rates, but much smaller than the $\underline{E} \times \underline{B}$ instability which would be operating at these altitudes; and electromagnetic effects are extremely weak for typical ionospheric situations. Based on all these results the current convective instability appears to be a relevant instability mechanism for high latitude scintillation causing irregularity studies.

ACKNOWLEDGEMENT

This work was supported by the Defense Nuclear Agency and the Office of Naval Research.

REFERENCES

Chaturvedi, P. K., Collisional ion cyclotron waves in the auroral ionosphere, J. Geophys. Res., 81, 6169, 1976.

Chaturvedi, P. K. and S. L. Ossakow, Nonlinear stabilization of the current convective instability in the diffuse aurora, Geophys. Res. Lett., 6, 957, 1979.

Chaturvedi, P. K. and S. L. Ossakow, The current convective instability as applied to the auroral ionosphere, J. Geophys. Res. (in press), 1981.

Fremouw, E. J., C. L. Rino, R. C. Livingston, and M. C. Cousins, A persistent subauroral scintillation enhancement observed in Alaska, Geophys. Res. Lett., 4, 539, 1977.

Hoh, F. C. and B. Lehnert, Diffusion processes in a plasma column in a longitudinal magnetic field, Phys. Fluids, 3, 600, 1960.

Huba, J. D. and S. L. Ossakow, Influence of magnetic shear on the current convective instability in the diffuse aurora, J. Geophys. Res., 85, 6874, 1980.

Kadomtsev, B. B., Plasma Turbulence, Academic Press, New York, pp. 11-15, 1965.

Kadomtsev, B. B. and A. V. Nedospasov, Instability of the positive column in a magnetic field and the "anomalous diffusion effect," J. Nucl. Energy, Part C, 1, 230, 1960.

Keskinen, M. J., S. L. Ossakow, and B. E. McDonald, Nonlinear evolution of diffuse auroral F region ionospheric irregularities, Geophys. Res. Lett., 7, 573, 1980.

Lee, W. W. and H. Okuda, Anomalous transport and stabilization of collisionless drift-wave instabilities, Phys. Rev. Lett., 36, 870, 1976.

Lehnert, B., Diffusion processes in the positive column in a longitudinal magnetic field, in Proceedings of the Second Geneva Conference on the Peaceful Uses of Atomic Energy, 32, 349, 1958.

Matsuda, Y. and H. Okuda, Simulation of dissipative trapped-electron instability in linear geometry, Phys. Rev. Lett., 36, 1976.

Ossakow, S. L. and P. K. Chaturvedi, Current convective instability in the diffuse aurora, Geophys. Res. Lett., 6, 332, 1979.

Rino, C. L., R. C. Livingston, and S. J. Matthews, Evidence for sheet-like auroral ionospheric irregularities, Geophys. Res. Lett., 5, 1039, 1978.

Rino, C. L. and J. Owen, The structure of localized nighttime auroral-zone scintillation enhancements, J. Geophys. Res., 85. 2941, 1980.

Rino, C. L. and S. J. Matthews, On the morphology of auroral-zone radiowave scintillation, J. Geophys. Res., 85, 4139, 1980.

Scannapieco, A. J. and S. L. Ossakow, Nonlinear equatorial spread F, Geophys. Res. Lett., 3, 451, 1976.

Scannapieco, A. J., S. L. Ossakow, S. R. Goldman, and J. M. Pierre, Plasma cloud late time striation spectra, J. Geophys. Res., 81, 6037, 1976.

Vickrey, J. F., C. L. Rino, and T. A. Potemra, Chatanika/Triad observations of unstable ionization enhancements in the auroral F-region, Geophys. Res. Lett., 7, 789, 1980.

OBSERVATIONS CONCERNING THE RELATIONSHIPS BETWEEN ELECTRON FLUX AND ELECTRIC FIELD VARIATIONS AND THERMAL PLASMA IRREGULARITIES IN THE POLAR CAP

William J. Burke
Air Force Geophysics Laboratory, Space Physics Division
Hanscom Air Force Base, Massachusetts 01731

ABSTRACT

Plasma density irregularities in the winter topside ionosphere over the polar cap can have deleterious effects on C^3I systems. However, the remoteness of the polar cap militates against investigations of the geophysical context in which the irregularities are found using standard, incoherent radar and sounding rocket techniques. The scientific package on S3-2 is uniquely capable of measuring some properties of the irregularities. The primary instruments for detecting irregularities are two oppositely directed, planar, ion sensors and a spherical thermal-electron probe. The surface normals of the ion sensor lie in the spin plane of the satellite, thus, limiting periods of useful information to data obtained in the direction of satellite motion. Environmental background information concerning the flux of keV electrons, electric fields, and Birkeland currents are also monitored. Available data come from two dusk to dawn north polar passes near the winter solstice of 1976. Save for the sign of B_z, solar wind and IMF parameters were similar. It is found that plasma density irregularities occur more frequently when the IMF had a northward rather than a southward component. Factors of 2 or greater variations in density over distance scales of $\leq$ 10 km were found at the edges of discrete polar cap arcs, in regions of strong shears of the convective electric fields where little or no precipitation was measured and along large-scale density gradients. Several possible local and remote sources of the irregularities are suggested

INTRODUCTION

Phase and amplitude scintillations of electromagnetic signals propagating through the night-time ionosphere frequently occur at equatorial, trough, auroral and polar-cap latitudes. The scintillations result from passage through density irregularities which alter phase speeds of waves. The largest contributions probably come from irregularities near the peak of the F layer. F layer irregularities may be observed remotely by means of incoherent backscatter radar or *in situ* by means of satellite-borne instrumentation. High-power radars are capable of detecting irregularities near the F peak, in restricted areas near widely spaced installations. At the present time no radar station is in a position to measure polar cap irregularities. Polar orbiting satellites have access to all geographic locations but mostly at altitudes well above the F-peak. Lacking regularly taken, simultaneous measurements of scintillations on the ground and density irregularities in the topside ionosphere it is frequently necessary for modellers to make the assumption that irregularities map along magnetic field lines. This is equivalent to assuming that $k_\parallel \ll k_\perp$ where $k_\parallel$ and $k_\perp$ are the components of the irregularity wave number spectrum parallel and perpendicular to the magnetic field.

The levels of our understanding of physical processes underlying topside irregularities vary according to latitude. There is a growing concensus that irregularities at equatorial latitudes, commonly referred to as spread-F, result from a collision dominated, Rayleigh-Taylor instability in the bottomside of the post-sunset F layer (Fejer and Kelley, 1980). Collocation of trough and auroral zone scintillations with latitudinal gradients in TEC (Fremouw et al., 1977) suggests that the irregularities result from combinations of E x B and current-convective instabilities (Keskinen et al., 1980). The morphology of polar-cap irregularities and scintillations is too little known at this time that any coherent theory, much less a concensus, concerning their origin has emerged. It is known that UHF scintillations measured at Thule are more pronounced when the polar ionosphere is in darkness than when it is sunlit (Aarons et al., 1981). Irregularities in F-region that cause some scintillations have been detected in the vicinity of a subvisual polar cap arc by Weber and Buchau (1981). It is hoped that a detailed analysis of the geophysical context in which polar-cap irregularities occur will suggest possible directions for investigating causative mechanisms.

This report presents measurements of topside ion density irregularities in the winter polar cap. Data were taken during two dusk to dawn, north polar cap passes of the USAF satellite S3-2. The

passes were selected as representative of magnetically quiet conditions when the interplanetary magnetic field (IMF) has southward or northward components. Our analysis of the density irregularities is facilitated by comparisons with simultaneously measured fluxes of energetic electrons, convective electric fields and field-aligned currents (FAC).

A brief overview of observed relationships between electric field and electron precipitation patterns in the polar cap and the orientation of the IMF is presented in the following section. We then describe the S3-2 scientific package. In the observations section measurements from two passes are presented. Irregularities are found to be widespread in the polar cap when the IMF has a northward component, but not when it has a southward component. The irregularities appear in regions of: (1) large-scale density gradients, (2) discrete polar cap arcs, and (3) fast varying electric fields. In the final section we discuss possible causal relationships between topside irregularities and polar cap precipitation, electric fields and FAC's.

POLAR-CAP ELECTRIC FIELDS, PRECIPITATION AND THE IMF

In this paper we make use of two coordinate systems. The first is the solar-magnetospheric system with its origin at the center of the earth; X positive toward the sun; Z positive toward magnetic north; and Y, which is positive toward magnetic dusk, completes the right hand system. IMF components are presented in this frame of reference. The second system, which is satellite-centered, is defined below.

Measurements from OGO 6 (Heppner, 1972) and S3-2 (Burke et al., 1979) show that the distribution of electric fields in the polar ionosphere is greatly affected by the orientation of the IMF. In the northern hemisphere the electric field tends to be strongest near the dawn (dusk) flank of the polar cap when the Y component of the IMF is positive (negative). The opposite polarity relationship maintains in the southern polar cap. Aside from small-scale structures, the cross polar cap electric field is directed from dawn to dusk (antisunward convection) when the IMF has a southward component. During periods of northward IMF extensive regions of dusk to dawn (sunward convection) electric fields are found in the central portion of the polar cap. Electric fields vary much more smoothly in the highly conducting summer than winter polar cap. This is especially true in the winter cap when the IMF is northwards. A geometrically intuitive explanation of the relationships between the IMF and high-latitude electric fields is given by Crooker (1979).

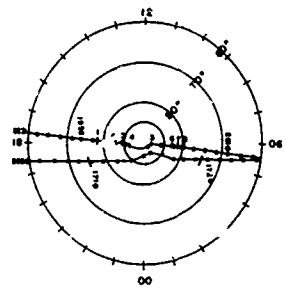


Fig. 1. S3-2 Trajectories in invariant latitude and magnetic local time during Rev 5056 and 5231

During magnetically quiet periods electron precipitation into the polar cap falls into two broad classes (Winningham and Heikkila, 1974). The first is a fairly uniform, low intensity flux of low energy (~ 100 eV) electrons called "polar rain". The second class is characterized by localized, more intense precipitation of higher energy electrons and is referred to as "polar showers". Ismail et al., (1977) have identified polar shower events with discrete polar-cap arcs which are normally found in periods of northward IMF. Electron fluxes into polar cap arcs vary in mean energy from a few keV down to ~0.5 keV (Weber and Buchau, 1981). They may be isotropic (Whalen et al., 1969) or highly field-aligned (Hoffman and Evans, 1968) in pitch angle. Burch et al. (1979) found that field-aligned fluxes, indicating electrostatic acceleration above the ionosphere, most likely apear in the northern polar cap when the IMF is in a northward, away sector ($B_x<0$; $B_z>0$).

INSTRUMENTATION

S3-2 was launched into polar orbit during December 1975 with an initial apogee, perigee and inclination of 1557 km, 240 km and 96.3° respectively. It was spin stabilized with a period of 21 sec. at the time of interest. The spin axis is nearly perpendicular to the orbital plane, in a cartwheel sense. The scientific package on S3-2 includes: (1) an electric field experiment consisting of pinning dipole, (2) a triaxial fluxgate magnetometer, (3) an energetic electron spectrometer, (4) thermal ion drift meter, and (5) a gridded, spherical electron probe. Data are presented using a satellite-centered coordinate system: X is positive in the direction of the satellite velocity; Z is positive toward local nadir and Y completes the right hand system. When the satellite trajectory lies in the dawn-dusk meridian, Y is positive in the antisunward direction (Smiddy et al., 1980).

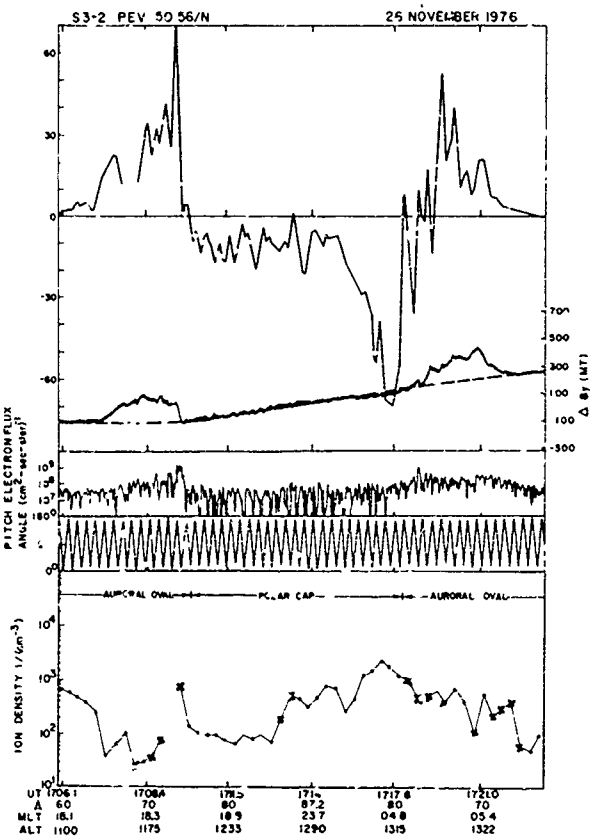


Fig. 2. Electric field components along and magnetic field component transverse to satellite velocity (top panel), directional flux and pitch angles of electrons with 80 eV $\leq$ E $\leq$ 17 keV (middle panel) and ion densities (bottom panel) measured during Rev 5056. Symbols X and • signify presence and absence of ion density irregularities.

The spinning dipole has a length of 29.72 m and measures the electric field component along satellite trajectory (E_X). Although potential differences between the ends of the dipole are sampled 32 times per second, data are normally presented as 5 second averages of E_X. The measurements can be used to infer values of E_X with a 1 second resolution. The magnetic field was measured 32 times per second with a 5 nT resolution. Data are presented as differences between measured and IGRF 75 values of the spin-axis component (ΔB_y). The electron spectrometer measures electrons with energies between 80 eV and 17 keV in 32 energy channels with a full spectrum compiled every second. The instrument's aperture is mounted in the spin plane of the satellite. Because of the instrument's small geometric factor (4.68×10^{-5} cm^2-sr) count rates are frequently at or below that of one per accumulation period (0.03 sec).

The ion drift-meter consists of two arrays of four passive thermal ion sensors, each with a circular aperture of 6.8 cm^2. The external configuration of the two arrays is identical. They are embedded in the surface of the vehicle at diametrically opposite positions. The internal geometry of sensors 1, 2, 3 and 4 is planar. To minimize their response to solar UV the internal geometry of sensors 5, 6, 7 and 8 is non planar (Wildman, 1977). Normal vectors to the apertures of sensors 2 (6) and 4 (8) lie in the spin plane at relative look angles of 40°. Sensors 1 (5) and 3 (7) are coplanar with 2 (6) but look out of the spin plane at angels of 40° toward and away from the spin axis, respectively. Here we only use measurements from times when sensors 2 and 6 face in the ram direction (T_R) to determine the total ion density once every 10.5 seconds.

The thermal electron sensor, which is mounted on a 1.2 m boom, consists of a solid collector of 2.54 cm radius surrounded by a concentric wire mesh grid of 3.30 cm radius. The instrument operates in a two-mode, repetitive cycle of 16 sec. duration. The grid is held at +1.5 V relative to satellite ground for 13.5 sec. In the second mode of operation the grid is swept from +8 to -8 V in 2.5 sec. The collector is held at +50V relative to the applied grid voltage. Thus, thermal electrons passing through the grid are accelerated to the collector but thermal ions are rejected. Electron densities and temperatures as well as satellite potential are determined using standard Langmuir theory.

In this report we limit ourselves to density irregularities that appear as rapid fluctuations in ion currents measured in intervals $T_R \pm 2$ sec. It is recognized that ion currents respond to changes in the convection velocity and the satellite potential as well as to the density. To distinguish between the three types of variation we compare the measurements of ion sensors 2 and 6 with those of the thermal electron probe. Where the ion sensor is sensitive to changes in the plasma flow velocity the spherical elecron probe is not; the ion sensor and the electron probe act in opposite senses to changes in the satellite potential. Thus, we call fluctuations in the ion current density irregularities only if both the ion and electron sensors vary in the same way.

OBSERVATIONS

Figure 1 shows the trajectories (invarient latitude and magnetic local time) of two north polar passes of S3-2 (Rev 5056 on 11 November and 5231 on 8 December, 1976) which were selected for analysis because both orbits were: (1) close to the winter solstice when irregularities are likely found in the polar cap (Aarons et al., 1981), (2) close to the dawn-dusk meridian passing to invarient lati-

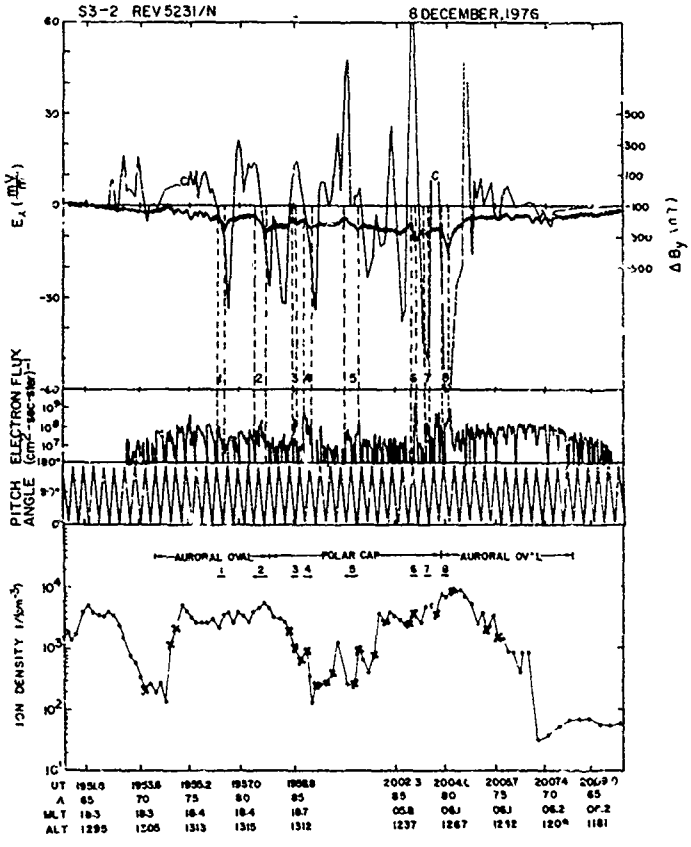


Fig. 3. S3-2 Measurements during Rev 5231 in the same format as Figure 2.

tudes (Λ) greater than 85°, (3) close to apogee over the magnetic pole thus, minimizing altitude variations in density. Background data listed in Table 1 shows that for both orbits, the level of magnetic activity was low, the solar wind velocity was moderate, the solar wind density was high and the IMF was in an away sector. The main difference was that the IMF had a southward component (-1.5 nT) during Rev 5056 and a northward component (+74 nT) during Rev 5231 (King, 1979). The symbols X and • indicate the progression of the satellite at one minute intervals during Rev 5056 and 5231, respectively. Lines cutting the Rev 5056 trajectory at Λ = 75° on the dawn and dusk sides approximate the boundary between the auroral oval and polar cap. The positions of eight high-latitude auroral arcs are marked on the Rev 5231 trajectory.

SS-2 Rev.	Date UT	Kp	Solar Wind V(km/s)	Solar Wind $n(cm^{-3})$	IMF B_x	IMF B_y	IMF B_z
5056	Nov.26,76 1715	2	330	15.7	-1.4	5.2	-1.5
5231	Dec.8,76 2000	1	439	10.1	-3.7	3.8	7.4

Table 1 Hourly Averaged Solar Wind & IMF

Observations from SS-2 Rev 5056 and 5231 are given in Figures 2 and 3, respectively. The format of both figures is the same. The top panel gives the dawn-dusk component of the electric field (E_x) and the Y component of the magnetic field perturbation (solid line). Positive values of E_x indicate a dusk to dawn component (equatorward convection). In the infinite current sheet approximation positive (negative) slopes in ΔB_y correspond to regions of FAC's into (out of) the ionosphere. The dashed line associated with the ΔB_y plot is a baseline that approximates the values of the transverse magnetic field component in the absence of FAC's. The middle two panels give the directional flux and pitch angles of electrons measured by the energetic electron detector. The bottom panel gives

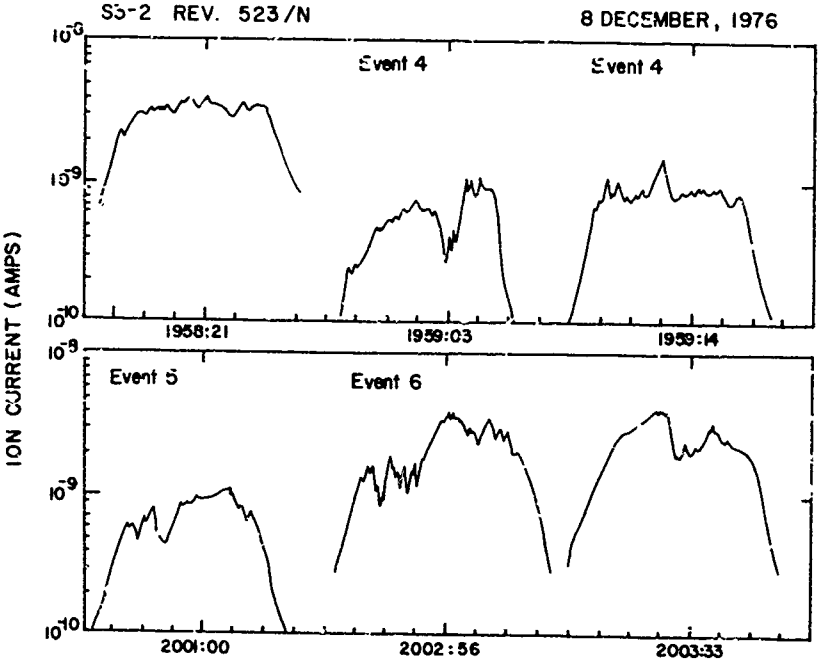


Fig. 4. Six selected examples of ion currents taken while ion drift meter sensors #2 or #6 faced in the ram direction.

the ion density (N_i) measured when sensors #2 or #6 of the drift meter faced the ram direction. During Rev 5056 only data from sensor 2 are available. Symbols X and • indicate the presence and absence or density irregularities. The data of Figures 2 and 3 which are equally spaced in UT are plotted as functions of UT, invariant latitude (Λ) magnetic local time and altitude.

Before discussing the data in Figures 2 and 3 it is useful to exemplify our criteria for determining whether the satellite passed through density irregularities. Six examples of ion current measurements taken in the polar cap during Rev 5231 are given in Figure 4. The tick marks are separated by one second; times when either sensor #2 or #6 faced the ram direction are listed for reference. Designations "Event 4, 5 or 6" indicate that the ion measurements were taken as the satellite passed close to a discrete polar cap arc (Burke 1981). We note that measurements taken near 1958:21 UT show much smoother variations than the other five. The irregularities can appear as density biteouts (1959:03 UT) or density spikes (1959:13 UT). Density measurements in Figure 2 and 3 are designated irregular if factors of 2 or greater density variations were found within the interval $T_R \pm 2$ sec.

Directing our attention to the high latitude data taken during the period of southward IMF (Figure 2) we note:

(1) E_x shows a clear evening side reversal in direction at 1710 UT (Λ = 76, MLT = 18.7) It maintains a low value across the dusk side of the cap, then decreases to a sharp minimum of -75 mV/m at 1717.8 UT (Λ = 80, MLT = 04.8). A strong dawn to dusk electric field on the dawnside of the polar cap is typical of $B_y > 0$ situations (Heppner, 1972). Between 1719.2 and 1719.3 UT, E_x undergoes a series of reversals before assuming an equatorward orientation in the dawnside auroral oval.

(2) The ΔB_y plot shows distinct Region 1/Region 2 signatures (Iijima and Potemra, 1976) in the dusk and dawn sides of the auroral oval. Across the polar cap ΔB_y stays close to its baseline value. It departs from the baseline by 25 nT as $|E_x|$ increases from ~10 to 75 mV/m at 1718 UT. Using equation (7) of Smiddy et al., (1980) we estimate the height integrated Pedersen conductivity of the darkened polar ionosphere to be 0.3 mho.

(3) Electron precipitation across the polar cap is that of a uniform polar rain.

(4) The ion density near 1250 km altitude varies from ~30 cm^3 in the duskside trough to a maximum of 2×10^3 at 1718 UT. The coincidence of maximum N_i and $|E_x|$ suggests that some of the polar cap, topside plasma has been convected from a remote production region such as the dayside cusp (Knudsen et al., 1977). Two examples of density irregularities are found in the central polar cap in the region of a steep density gradient. It is not clear whether the irregularities collocated with rapid variations in the direction of E_x lie in the polar cap or in the high-latitude reaches of the dawnside auroral oval. If they lie outside the cap only 2 of 23 N_i measurements in the polar cap show irregularities.

The relationship between the electric field, magnetic field and electron measurements shown in Figure 3, and discrete polar cap arcs is discussed elsewhere (Burke, 1981). During this period of northward IMF the 5 scond average values of E_x are highly variable, with a maximum strenghth of 120 mV/m on the dawnside of the polar cap. The ΔB_y measurements show no sign of standard Region 1/ Region 2 FAC systems. Small-scale variations in ΔB_y are found. In Figure 3 we have marked eight events which are characterized by negative slopes in ΔB_y and localized, enhanced precipitation. These events have the characteristics of discrete arcs. To determine which arcs lie in the polar cap we used the first and last detection of polar rain as marking the boundary between open and closed field lines (Heelis et al., 1980). The evening and morning side boundaries of the polar cap arc set at 82° and 81.6°, respectively. Thus, events 3 through 7 lie within the polar cap. It is possible that event 8, which lies poleward of a significant reversal in E_x, also lies in the polar cap.

On the evening (morning) side ion densities range from a low of 150(30)cm^{-3} in the trough to a maximum of 5×10^3 (10^4)cm^{-3} at the poleward boundary of the auroral oval. As in the case of Rev 5056 we find a density trough on the evening side of the polar cap. However, background density levels during Rev 5231 are higher by a factor of 5 or more. Within the polar cap irregularities were observed in 14 of the 32 density measurements. Eleven of the irregurities are found in conjunction with a large scale density gradient in N_i and/or in the vicinity of a polar cap arcs. The remaining three irregularities are found at the polar cap trough minimum ($\leq$ 2000 UT) with only polar rain precipitation. They are also found in a region of very rapidly varying electric field. Raw E_x measurements (not shown) from this period show multiple reversals in 1 second averaged data.

To investigate possible relatinships between density irregularities and polar cap arcs we have plotted data taken near event 6 in an expanded scale in Figure 5. The top panel contains plots of E_x, ΔB_y and the electron directional flux. Values of E_x are one, rather than five second averages. The bottom panels contains six spectra taken across the event. Dashed lines indicate one-count background levels of flux. The UT at the beginnings of spectra accumulations and the pitch angels half way through one second accumulation periods are indicated to the right of the lower panel.

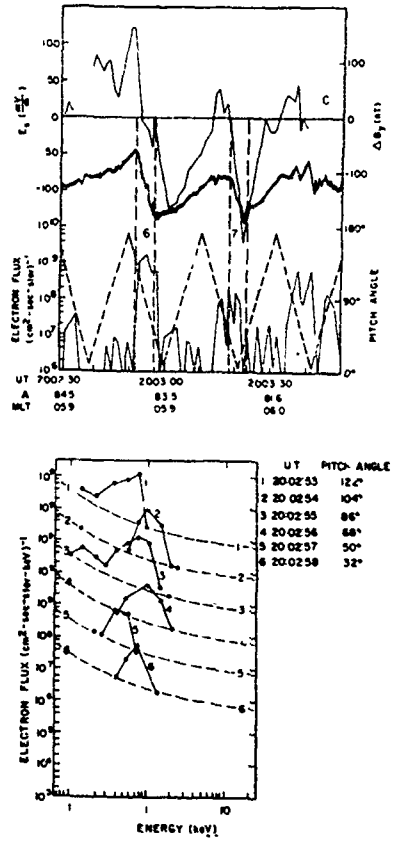


Fig. 5. Expanded Plots of E_x, ΔB_y and electron fluxes measured during Rev 5231 near the time of event 6 (top panel) and six consecutive electron spectra taken across this event.

The event is marked by: (1) a strong shear in the electric field. E_x changed from 120 mV/m dusk to dawn to 120 mV/m dawn to dusk in 10 seconds. (2) a field-aligned current out of the ionosphere of ~ 2 $\mu A/m^2$. Such a FAC demands a net flux of electrons into the ionosphere of $2.5 \times 10^9 (cm^2 sec\text{-}ster)^{-1}$ (3) an abrupt increase in flux to $2 \times 10^9 (cm^2 sec\text{-}ster)^{-1}$ and return to background that coincides with the satellite passing through the upward FAC sheets. The relatively uniform flux measured at different pitch angles across the event suggest that the flux is fairly isotropic over the down coming hemisphere. (4) an average electron energy of ~ 1 keV. The directional energy flux is in the range of 2-3 ergs/cm^2sec-ster, and is sufficient to produce visible emissions.

Examining the ion current measurements marked "event 6" in Figure 4 we note the presence of density irregularities between 2002:53 and 2002:55 UT. Between 2002:55 and 2002:56 N_i increased from 2.1 to $5.2 \times 10^3 cm^{-3}$. From the bottom panel of Figure 3 the densities on both sides of event 6 (2002:45 and 2003:03) were 2×10^5, and no irregularities were observed. We conclude that the density peak and irregularities found above the arc are probably by products of the arc formation mechanism. Possible causal relationship are discussed below.

SUMMARY AND DISCUSSION

In the previous section we presented a broad set of measurements form two dusk to dawn passes of S3-2 across the northern, winter polar cap. The orbits were chosen for analysis because: (1) they followed similar invariant latitude/magnetic local time trajectories, (2) they occurred when the polar cap was in darkness and irregularities are apt to be found (Aarons et al., 1981), and (3) they occurred when, except for the polarity of B_z, solar wind and IMF conditions were similar.

While the IMF had a southward component (Rev 5056) the convective electric field pattern was well defined. When the IMF was northward E_x was highly irregular. As expected for $B_y>0$ the largest electric field were found in both cases, near the dawn flank of the polar cap (Crooker, 1979). The background ion density level was a factor of five or more larger during Rev 5231 than 5056. We lack evidence to determine whether this due to a UT effect (Sojka, et al., 1979), a generally higher level of ionospheric production from the polar cap precipitation when the IMF is northward or to geomagnetic conditions preceeding the chosen orbits. In both cases distinct density troughs were found on the dusk side of the polar cap and density maxima on the dawnside. That the large-scale density maximum coincides with the region of strongest convective electric fields suggests that a large fraction of the plasma was convected from the dayside cusp to the place of observation (Knudsen et al., 1977).

Uniform, polar rain precipitation was measured across the polar cap during Rev 5056. The polar rain was interrupted by at least five, intense precipitation events during Rev 5231 that could produce visual or subvisual arcs. Because of the low level of ionospheric conductivity at the time of Rev 5056 only weak FAC's were found in the polar cap. Ionization created in the vicinity of polar cap arcs was sufficient to support > 1 $\mu A/m^2$ FAC's.

Density irregularities were almost entirely absent from the central polar cap when the IMF had a southward component. The two examples of polar cap irregularities occurred along a large-scale density gradient. Several examples of irregularities were also found in a region of topside electric field variations near the boundary of the polar cap and the dawnside auroal oval. It is not clear to which region these irregularities should be assigned.

By way of contrast, density irregularities were quite widespread in the polar cap when the IMF was northward. The irregularities, which appeared as plasma biteouts and enhancements,were found in the vicinity of large-scale density gradients, polar cap arcs and highly irregulr electric fields. If some of the biteouts crossed by S3-2 in ≤ 1 sec map to the peak of the F layer they would have scale lengths of ~ 5 km and would be capable of producing UHF scintillations.

A detailed theoretical analysis of polar cap irregularities is well beyond the scope of this observational report. Depending on circumstances several mechanisms do suggest themselves.

(1) Irregularities have been found near large-scale density gradients at the equatorward edge of the auroral oval. These have been attriuted to current-convective instabilities (Keskinen et al., 1980). In the absence of large-scale FAC's in the winter polar cap this mechanism near large scale density gradients in the polar cap. However, if the convective electric field had a small antisunward component (not measured by S3-2) near the morning side gradient, the plasma might be subject to an ExB instability (Ossakow and Chataverdi, 1979). A sunward electric field component would be required to render the density gradient between the auroral oval and evening side polar cap trough unstable. Verification of these possibilities awaits future space missions such as Dynamics Explorer.

(2) Irregularities found near polar cap arcs result from a host of competing processes. Enhanced precipitation creates localized strips of enhanced ionization. Because of the low energies of the precipitating electrons a substantial fraction of the ionization is created or is transported to F layer altitudes. Effects of enhanced ionization are evident in data associated with event 6. However, electric fields and FAC's associated with the arcs tend to produce F region depletions (Schunk et al., 1975, Block and Falthammer, 1968), as well as mixing of low and high density flux tubes via ExB and current-convective instabilities (Ossakow and Chataverdi, 1979).

(3) Irregularities collocated with rapidly varying electric fields may result in part from a Kelvin-Helmholtz type instability.

(4) It is possible that some of the irregularites were produced as a result of spatial and temporal variations in soft precipitation in the region of the cusp (cf Figure 46 of Fejer and Kelley, 1980). They could then be convected to their place of observation in the central polar cap. For an antisunward convection speed of 1 km/sec ($E_x = -40$ mV/m) it takes $\leq 1/2$ hour to move from the cusp to the central polar cap.

ACKNOWLEDGEMENT

The author gratefully recognizes the various contributions of F.J. Rich, M. Smiddy, B. Shuman, R.P. Vancour and P.J.L. Wildman of AFGL in providing S3-2 data for this study. Special thanks is extended to Laurie Cassidy of Regis College for drafting figures and to Karen Leccese of AFGL for typing the text.

REFERENCES

Aarons, J., J.P. Mullen, H.E. Whitney, A.L. Johnson and F.J. Weber, UHF scintillation activity over polar latitudes, Geophys. Res. Lett. (to be published) 1981.

Block, L.P. and C. -G. Falthammer, Effects of field-aligned currents on the structure of the Ionosphere, J. Geophys. Res., 73, 4807, 1968.

Burch, J.L., S.A. Fields and R.A. Heelis, Polar cap electron acceleration regions, J. Geophys. Res., 84, 5863, 1979.

Burke, W.J., M.C. Kelley, R.C. Sagalyn, M. Smiddy and S.T. Lai, Polar cap electric field structures with a northward interplanetary magnetic field, Geophys. Res. Lett., 6, 21, 1979.

Burke, W.J., Electric fields, Birkeland currents and electron precipitation in the vicinity of discrete auroral arcs, in Physics of Auroral Arc Formation, ed by, S. -I. Akasofu, American Geophysical Union, Washington, DC (to be published) 1981.

Crooker, N.U., Antiparallel merging, the half-wave rectifier response of the magnetosphere, and convection, Proceedings of Magnetospheric Boundary Layers Conference, Alpbach, 1979.

Fejer, B.G. and M.C. Kelley, Ionospheric irregularities, Rev. Geophys. Sp. Sci., 18, 401, 1980

Fremouw, E.J., C.L. Rino, R.C. Livingston and M.C. Cousins, A persistent subauroral scintillation enhancement observed in Alaska, Geophys. Res. Lett., 4, 539, 1977.

Heelis, R.A., J.D. Winningham, W. B. Hanson and J. L. Burch, The relationship between high-latitude convection reversals and the energetic paticle morphology observed by Atmospheric Explorer, J. Geophys. Res., 85, 3315, 1980.

Heppner, J.P., Polar cap electric field distributions related to the interplanetary field direction, J. Geophys. Res., 77, 4877, 1972.

Hoffman, R.A. and D.S. Evans, Field-aligned electron bursts at high latitudes observed by OGO 4, J. Geophys. Res., 73, 6201, 1968.

Iijima, T., and T.A. Potemra, The amplitude distribution of field-aligned currents at northern high latitudes observed by TRIAD, J. Geophys. Res., 81, 2165, 1976.

Ismail, S., D.D. Wallis, and L.L. Cogger, Characteristics of polar cap sun-aligned arcs, J. Geophys. Res., 82, 4741, 1977.

Keshinen, M.J., S.L. Ossakow and B.E. McDonald, Nonlinear evolution of diffuse auroral F region ionospheric irregularities, Geophys. Res. Lett., 7, 573, 1980.

King, J.H., Interplanetary Medium Data Book Suplement 1 1975-1978, NSSDC/WDC-A-R&G 79-08, Goddard Space Flight Center, Greenbelt, MD 20771, 1979.

Knudsen, W.C., P.M. Banks, J.D. Winningham and D.M. Klumpar, Numerical model of the convecting F_2 ionosphere at high latitudes, J. Geophys. Res., 82, 4784, 1977.

Ossakow, S.L. and P.K. Chaturvedi, Current convective instability in the difuse aurora, Geophys. Res. Lett, 83, 2085, 1979.

Schunk, R.W., W.J. Raitt and P.M. Banks, Effects of electric fields on the daytime high latitude E and F regions, J. Geophys. Res., 80, 3121, 1975.

Smiddy, M., W.J. Burke, M.C. Kelley, N.A. Saflekos, M.S. Gussenhoven, D.A. Hardy and F.J. Rich, Effects of high latitude conductivity on observed convection electric fields and Birkeland currents, J. Geophys. Res., 85, 6811, 1980.

Sojka, J.J., W.J. Raitt and R.W. Schunk, Effect of displaced geomagnetic and geographic poles on high-latitude plasma convection and ionospheric depletions, J. Geophys. Res., 84, 5943, 1979.

Weber, E.J. and J. Buchau, Polar cap F-layer auroras, Geophys. Res. Lett. (to be published) 1981.

Whalen., B.A., J.R. Miller and I.B. McDiarmid, Sounding rocket observations of particle precipitation in a polar-cap electron auroral, J. Geophys. Res., 76, 6847, 1971.

Wildman, P.J.L., A low energy ion sensor for space measurements with reduced photo-sensitivity, Sp. Sci. Inst., 3, 363, 1977.

Winningham, J.D. and W.J. Heikkila, Polar cap auroral electron fluxes observed with ISIS I, J. Geophys. Res., 79, 949, 1974.

The Night-Time Auroral-E Layer: Particle Production, Latitudinal and Longitudinal Structure and Dynamics

James A. Whalen
Ionospheric Physics Branch
Air Force Geophysics Laboratory
Hanscom AFB, MA 01731

and

James R. Sharber*
Department of Physics and Space Sciences
Florida Institute of Technology
Melbourne, FL 32901

It has been known for some time that auroral-E, a thick layer of ionization found at E-layer heights at auroral latitudes, is associated with the continuous (diffuse) aurora (Whalen et al, 1971; Whalen et al., 1977; Weber et al., 1977; and Winningham et al., 1978). In the first section of this paper we report a detailed study of the spectral shapes and relative total energies of the electrons and protons which produce it and the correlation of the spectral features with its ionospheric signatures, the critical frequency f_0Ea and the virtual height h' Ea. In the second section we describe the morphology of the continuous aurora/auroral E layer based on statistical studies. In the third section we describe the structure and dynamics of this aurora in its entire latitudinal and local night time extent for a period of 8 hours duration.

PARTICLE SOURCE OF AURORAL E

The particle study was made possible by the analysis of data taken as Isis 2, a polar satellite orbiting at 1400 km, passed over the AFGL Flying Ionospheric Observatory as it made ionospheric soundings of the auroral-E layer in the midnight sector. Figure 1, reproduced from Whalen et al. (1977) shows the aircraft (heavy curve) and Isis 2 (dotted line) trajectories plotted in corrected geomagnetic (CG) latitude and longitude. The particle measurements reported here were taken during the satellite pass labeled 0610, the universal time at which it crossed the aircraft flight path. The satellite traversed the auroral precipitation region in about five minutes while the plane flew from 75° CG latitude to 67° CG latitude between 0517 and 0700 UT.

Particle Precipitation Over the Continuous (Diffuse) Aurora

The particle observations were made by the Isis 2 soft particle spectrometer (SPS), a swept voltage, divergent plate energy analyzer which provides a differential spectrum for electrons and positive ions over the energy range 10 eV to 15 keV each second. The measurements taken during the 0610 UT pass (Orbit 3203) of 09 December 1971 are displayed in the energy/time spectrogram format of Figure 2. Electrons and positive ions (labeled "protons" in the figure) are shown in the top two panels of the figure. (The SPS measures energy per unit charge and thus does not distinguish positive ion species. The terms positive ion and proton will be used interchangably in this paper.) Each vertical line in the spectrograms represents a differential energy spectrum (erg/cm^{2}s sr) measured over the range ~10 eV to ~15 keV. The ordinate scale is the logarithm of energy. Intensity changes during a spectral sweep (one vertical line) are indicated by gray shading; lighter shades correspond to greater intensity. The abscissas are universal time minutes along the bottom of the data frame, and magnetic local time (MLT) and invariant latitude (INV. LAT.) along the top at each minute mark. The third panel from the top shows particle pitch angle θ p, the repetitive scan which is obtained as the satellite spins with a period of 22s. For easy reference, seven spin cycles are labeled above the spectrograms. The lowest panel, labeled DOWNWARD E FLUX, shows the result of integrating the particle energy flux (erg/cm^{2}s) over the ~50° loss cone for electrons, E and protons, P.

* Currently an NRC/AFSC Senior Postdoctoral Associate at the Air Force Geophysics Laboratory, Hanscom AFB, MA 01731

It is clear from Figure 2 that between invariant latitudes of 73° and 66° the electron spectra are broadly peaked and show very little structure. This is characteristic of the continuous aurora (Meng, 1976; Doehr et al. 1976; and Lui et al., 1977) and is confirmed in this case by the lack of any discrete arcs observed at this time by the airborne all-sky camera (Whalen et al., 1977). The electron average energy peaks during spin cycle 4 (Λ=69.6°) and decreases both equatorward and poleward of the peak. The rotation of the satellite is apparent by the regular (and darker) gaps in the data as the analyzer aperture looks earthward (centered on θ_L=180°), but we also note that during spin cycles 3, 5, 6, and 7, another gap appears, particularly at high energies, as the aperature looks upward along the field line. This is the unfilled loss cone for downcoming electrons. The fact that isotropy does not always exist means that in any modeling effort using particle data from satellites to calculate energy deposition into the ionosphere, the pitch angle distribution within the loss cone should be determined.

Positive ions (protons) accompany the electrons but do not occur in significant numbers below 67°. Their average energy increases with decreasing latitude down to spin cycle 5(Λ=68.7°). Except for the return fluxes from below the satellite, the positive ions show isotropy at all latitudes above 68.7°. This is a general characteristic of positive ion precipitation into the continuous auroral oval.

Particle fluxes are actually observed as far poleward as 78.5°. The electrons show structure suggestive of faint polar cap arcs. Their energies decrease gradually to about 30 eV at 78.5°. The positive ion energies decrease to about 100 eV at 78.3° but they show almost no structure with latitude. Both components of this low-energy, high-latitude plasma appear to merge into the electrons and ions of the continuous oval. They are typical of very quiet times (Sharber and Premaze, 1980), are the quiet time boundary layer of the plasma sheet, the BPS of Winningham et al. (1975), and presumably bear the same relation to polar cap arcs as the continuous auroral plasma does to oval arcs. The continuous aurora, between ~66° and ~73°, is therefore the low altitude extension of central plasma sheet, or the CPS of Winningham et al. (1975). Poleward of 78.3° a very faint, uniform electron "polar rain" (Winningham and Heikkila, 1974) is observed.

In order to produce the latitudinal profile of precipitating particle fluxes, shown in the lowest panel of Figure 2, the spectra have been integrated first in energy over the range of measurement, then over the 50° loss cone (at 1400 km). It should be noted here, that in order to calculate the energy flux reaching the ionosphere (i.e., the loss height), the flux shown this panel must be multiplied by $1/\sin^2 50°$ or 1.70. The precipitating electrons have a peak flux at 69.6° (cycle 4) invariant latitude; the protons have peak at 68.7° (cycle 5). The ratio of proton to electron precipitating energy flux is ~0.1 at 69.6° and ~0.2 at 68.7°. This difference in latitude of the peak flux is a persistent observation. The equatorward displacement of the protons is characteristic of the evening--midnight magnetic local time sector.

Other Isis 2 passes show that in this sector it is not unusual during very quiet times to see the precipitating energy flux of the proton population exceed that of the electrons near the equatorward edge of the particle precipitation region. Thus in general, the continuous aurora is caused by keV electrons and protons with the relative contribution from each constituent varying as a function of latitude.

Spectra

Electron spectra characterizing the continuous aurora may generally be represented by a Maxwellian distribution above a few hundred eV of the form $j(E)=(j_o e/\alpha)Ee^{-E/\alpha}$, where j(E) is differential number flux (#/cm^{2}s sr eV) at energy E, α is the characteristic energy, and $j_o \equiv j(\alpha)$. However, α and j_o usually vary with pitch angle as shown in Figure 3, where two spectra from the second spin scan of cycle 5 are plotted. The spectra measured at pitch angles of 38° and 9° are well-represented by the Maxwellians at energies above about 600 eV. The labels on the figure refer to the Maxwellian curves. The energy flux E, the number flux N, and the number density n are derivable from the parameters j_o and α which define each distribution. We note that the energy flux at 38° exceeds that at 9° by about a factor of two. In order to compare precipitating energy flux with the sounder measurements of $(fcEa)^4$, all spectra measured during a pitch angle scan must be integrated over the loss cone.

An example of the positive ion spectrum, taken from data within the loss cone during cycle 5 (Λ=68.7°) is shown in Figure 4. The notation " θpLCA" (loss cone average) means that spectral sweeps within the loss cone have been averaged to obtain the spectrum. The procedure is justified by the isotropy of the positive ion flux. We note that above a keV the spectrum is Maxwellian with α = 8,300 eV, much higher than that of the electrons, but that the total energy flux under the spectrum is only 0.14 erg/cm^{2}s sr. This is the most energetic of all the ion spectra obtained over cycles 1 through 5. In fact the characteristic energy α and energy flux E decrease to values of 2,300 eV and 4.6×10^{-3} erg/cm^{2}s sr during cycle 1 (Λ=72.5°). This is a general characteristic of the positive ion component of the continuous aurora. The straight line of negative slope is the

instrumental 2-count sensitivity level. Since the positive ions are isotropic, their precipitated energy flux (at ionospheric heights) is obtained by multiplying the spectral energy flux by π. Using the Maxwellian curve of Figure 4, the flux of positive ions precipitated at ionospheric heights is 0.44 erg/cm^{2}s.

Maxwellian Distribution of Precipitating Electrons

Since in general the electron flux over the continuous aurora is not isotropic, flux values j_i measured at different pitch angles in the scan must be integrated over the loss cone in order to obtain the flux j precipitating to ionospheric heights. We have used the expression

$$j = 1.75\ \Sigma\, j_i\ (\theta i, \theta z)\cos\theta_i \sqrt{\sin^2\theta_z - \cos^2\theta_i} \qquad (1)$$

where θ_z is the angle between the satellite spin axis and the local magnetic field line, and θ_i is the pitch angle. The equation has been derived for the Isis 2 SPS geometry and spin period and takes into account the detector angular response and the angular spacing and overlap in pitch angle space of successive spectral measurements.

Applying this expression to the individual differential elements of each spectrum results in a composite differential (in energy) number spectrum (#/cm^{2}s eV) already summed over angles within the loss cone. It is a single spectrum representating the precipitating electrons. An example shown in Figure 5 is from the second pitch angle scan of cycle 5 (scan 5-2). The interesting feature of this spectrum is that it is a Maxwellian distribution. Above 600 eV the parameters α =690 eV and j_o =1.3x10^5/cm^{2}s eV completely specify the precipitating distribution. In a later section the characteristic energies obtained from each spin scan will be compared with the virtual heights measured from the airborne sounder. As was the case for the individual spectra of Figure 3, the low-energy part of the spectrum deviates significantly from the Maxwellian. These electrons, however, contribute very little (<5%) to the total energy as shown by plotting the data in integral form in Figure 6.

Latitudinal Profile of Energy Flux and Characteristic Energy

Using the summing technique described in the previous section, the energy flux of precipitating electrons measured by Isis 2 has been calculated for each spin cycle over the continuous aurora. The values are plotted against invariant latitude in Figure 7. Cycle 1 is on the right, cycle 7 on the left. Note that the precipitating energy fluxes calculated at 1400 km have been evaluated at the loss height by multiplying by 1/sin^{2}50°. The measurements are therefore of electron energy flux (erg/cm^{2}s) in the ionosphere. A Gaussian curve of 3.2° FWHM has been fitted to the points from Cycles 2, 3, 5, 6, and 7; but it is clear that the high value of 2.5 erg/cm^{2}s obtained during Cycle 4 departs considerably from the Gaussian fit. This was the only cycle over which electron isotropy was observed. Although the Gaussian shape appears to be a regular feature of the energy deposition associated with the continuous aurora, it, and the narrower curve through the peak flux value, are used here primarily for later comparison with the latitudinal distribution of $(foEa)^4$.

The characteristic energies obtained from the precipitating Maxwellian distributions are plotted in Figure 8. The points represent a determination of α from the two spin scans of each cycle. The data points have been fitted with another Gaussian curve, one which has a 3.5° FWHM.

Comparison with the Airborne Sounder

During the flight of 09 December 1971, ionospheric soundings were made from the aircraft once each minute with the Granger, Model 3905-1 sounder. Each measurement required 20s to sweep through the range 2 to 8 MHz. Further details concerning the measurement of plasma critical frequency foEa and virtual height h'Ea during this flight may be found in Whalen et al. (1977).

The plasma critical frequency is associated with the maximum electron density n_e(max) of the E-layer by the well-known relation

$$n_e(\max) = 1.24 \times 10^4\ (f_oEa)^2 \qquad (2)$$

The ion recombination rate at E-layer maximum is proportional to n_e^2(max), hence to $(f_oEa)^4$. Since equilibrium can be assumed, the rates of ion production and recombination are equal, and $(f_oEa)^4$ is proportional to the ion production rate at n_e(max). This ion production rate is proportional to the total energy flux deposited by the precipitating particles. This explanation, originally put forth by Oxholt (1955) and further supported by Whalen et al. (1971) and Jones (1974), provides the basis for the present comparison between the Isis 2 measurement of incoming particle energy and $(f_oEa)^4$.

The sounder measurements of $(f_oEa)^4$ are plotted vs. CG latitude in Figure 9. The profile was obtained as the aircraft's CG latitude decreased from 75° at 0517 UT to 67° at 0700 UT (refer again to Figure 1); the auroral-E layer was observed between ~74° and ~68°. The equatorward part of the distribution has been fitted with the same curves used to fit the particle energy flux data of Fig-

ure 7 - the Gaussian of 3.2° FWHM and the sharper peak. In fitting these curves, an effort has been made to include on the Gaussian the two points measured between latitudes 70° and 71° since these were measured coincidently. Points equatorward of 70° fit the distributions well, although it is clear that the peak in $(foEa)^4$ is displaced toward lower latitudes by about half a degree. One half hour had elapsed between the satellite measurement of peak flux and the aircraft measurement; and as a later section of this report will show, during this period the continuous auroral form moved slowly equatorward in the CG latitude. The higher value of $(foEa)^4$ above about 71° is apparently due to a larger flux there at the earlier time of the aircraft measurement.

In comparing the Gaussian portions of the particle and sounder data (Figures 7 and 9) we find that the calibration factor is 0.013 erg/cm^{2}s MHz^4. This is the multiplication factor for converting sounder data to precipitating electron energy flux.

Virtual height, h'Ea, is expected to correlate with the energy of particles incident on the atmosphere from above since the depth to which they penetrate and produce maximum ionization depends on energy. We have already demonstrated that the incident spectra of the continuous aurora are Maxwellian distributions. The characteristic energy α is the energy of particles at the peak of such a distribution and provides a natural parameter to compare with virtual height.

We begin by using the virtual height as an estimate of the altitude of maximum ionization production. Preliminary study indicates that this assumption is reasonable since both virtual height and the altitude of peak ionization production occur at a height which is roughly 0.8 the height of the electron density maximum. It is further supported by its success in inferring kinetic energies of electrons incident on the ionosphere (Whalen et al., 1971).

An ionization production curve for Maxwellian electron spectra incident isotropically on the auroral atmosphere has been constructed using data from Jones and Rees (1973), Roble and Rees (1977) and recent calculations by Strickland (private communication). Figure 10 shows the altitude of peak ionization production plotted again Maxwellian characteristic energy α. Figure 11 shows a latitude plot of energies α (h'Ea) obtained from the sounder virtual height measurements and the curve of Figure 10. Over these points has been graphed the same 3.5° FWHM Gaussian used to fit the satellite characteristic energies of Figure 8. The approximate fit of the high latitude points to the same curve displaced by about one degree suggests a temporal increase between 0510 UT and 0610 UT.

Comparing magnitudes of the distributions of Figures 8 and 11 suggests that the characteristic energy obtained from the sounder measurements only slightly overestimates the characteristic energy of the incident electron spectrum. To within less than 10% the sounder measurements of virtual height therefore provide the characteristic energy of the incident electron spectrum.

MORPHOLOGY OF THE CONTINUOUS AURORA

Gaussian Latitudinal Distribution

The energy flux of the continuous aurora typically has a Gaussian latitudinal distribution which is observed directly in terms of precipitating particles, and via their effects on the atmosphere by photometric and ionospheric means. Six examples of these measurements are shown in Figure 12 in a plot of energy flux versus CG latitude near midnight. The triangles are particle (electron) measurements; open circles, photometric; closed circles, ionospheric (foE^4).

The Gaussian nature of the distributions, which is independent of the type of measurement, are dependent on 3 parameters: the maximum value of energy flux, Q max, the latitude of the maximum, Λ_{max}, and the Gaussian scale parameter, σ which describes the latitudinal width of the distribution.

The distribution at highest latitude is the only one that is clearly asymmetric: its poleward edge can be fit by a Gaussian of $\sigma \sim .68°$, however the equatorward edge is truncated much more abruptly - if Gaussian, $\sigma \sim 0.34°$. This distribution records an extremely quiet condition and the asymmetry may be related to that.

Latitude-Local Time Distributions

A collection of 14 Gaussian distributions is shown in a plot of CG latitude vs. CG local time in Figure 13. Each point marks the location of a measurement of Q max, i.e. Λ max; the bars denote 1/2 Q max points. Here the x's are derived from ionospheric soundings (foE^4); the squares, from precipitating electrons; and the circles, from photometric measurements (open, visible; closed, VUV).

The solid vertical lines extend to the half maxima points of measured latitude distributions which are Gaussian. The dashed vertical lines represent regions of the distribution which were not measured - their extent assumes that the unmeasured portion is symmetric with the measured part.

The solid curve and the two dashed satellite curves represent the results of a statistical study of ground-based E-layer ionospheric soundings which appear to be consistent in latitude and local time with the individual instantaneous measurements.

The statistical study of the auroral E-layer using an ensemble of ground-based ionospheric sounders was performed by Besprozvannaya and Shchuka (1976) who studied a 4 month winter period during the IGY. These workers derived contours in CG latitude-CG LT of frequency of occurrence of the auroral E layer and of the median fbEs. In Figure 13 the solid line is the locus of maximum frequency of occurrence of auroral E in CG latitude and CG local time for Kp of 0 and 1 in 1958. To a good approximation this maximum frequency of occurrence is described by two arcs of a circle: for the local time sector which extends from 1600 CGT through midnight to 0800 CGT the maximum frequency of occurrence can be well described by the arc of a circle of radius 19° centered about a pole from the offset location at 85.8° CGL and 00:00 CGT. The remaining 8 hours of CGT, through the day sector, are best described by the arc of a circle of radius 17° CGL centered at 90° CGL. The maxima both of the frequency of occurrence and of the median fbEs have the same shape and are taken to be identical within the uncertainties. The dashed lines are half maxima points of the median fbEs[4].

Offset Pole Reference Frame

The indication that the auroral E layer/continuous aurora is organized by a frame of reference, the origin of which is offset from the CG pole, is an important result. A number of other workers have found other auroral phenomena to be ordered in such a system beginning with Starkov (1969) who found that the low latitude boundaries of the Feldstein auroral ovals were approximately circles centered about a pole located at 84.9° CG latitude, and 00.8h CGT the radius of which increased approximately linearly with the magnetic index Q.

Holzworth and Meng (1975) and Meng et al. (1976) found certain quiet auroral arcs detected by the DMSP satellites to be circles centered at a pole whose mean position was the same that indicated by the aforementioned ionospheric study. Other workers have also found that similar (but not identical) offset poles order precipitating particle data.

STRUCTURE AND DYNAMICS OF THE ENTIRE NIGHTTIME CONTINUOUS AURORA

Instantaneous Latitude and Local Time Distributions

The foregoing results that foEa[4] measures particle energy flux in the continuous aurora, that the latitudinal distribution of this energy flux is characteristically Gaussian, and that this aurora is centered about an offset pole, provide the impetus for the description of the entire pattern of this aurora: the latitudinal profile, the longitudinal distribution of the profile and the dynamics of this entire pattern. The period described is that of the case study 9 Dec 1971 and the measurements are of foEa from the 5 ionospheric sounders shown in Figure 1: Godhavn (GO), Narssarssuaq (NQ), Churchill (CH), College (CO) and the aircraft (A/C). Simultaneous measurements, the locations of which span 12° of CG latitude and nearly 10 hours of CG local time, are ordered by the transformation of their locations into the coordinate system of the offset pole located at 85.8° CG latitutde and 0Ch CG time.

The latitudinal variation with UT of the 4 ground stations is shown in Figure 14 which is a rectangular plot of the latitude with respect to the offset pole, Γ versus UT. In the offset reference frame the location of a fixed point in the earth's surface changes in an approximately sinusoidal manner as a function of UT. A station is at maximum latitude at CG midnight, (CG latitude + 4.2°) and minimum at noon, (CG latitude - 4.2°). The trajectories are shown dashed in the day sector since the statistical results indicate that coordinate system which organizes the continuous aurora between 08 and 16 CGT may not be the offset pole. The trajectories differ in phase as a result of their longitudinal differences, and in baseline latitude as a result of their latitudinal differences.

As a particular example the positions in Γ of the stations at 0545 UT are given by Figure 14 as follows: GO, 80.1°; NQ, 71.4°, CH 74.4°, CO, 65.7°; A/C (not shown), 76.6°. The value of foEa[4] measured at each station is plotted at its corresponding Γ latitude in Figure 15. Three of the measurements (NQ, CH and A/C) have been fit by a Gaussian distribution of 3.2° FWHM, which is the same as that determined from the ISIS-2 electrons at 0610 UT, and is typical of other cases. No measurement is available at CO at this time, because foEa is below threshold (i.e. <1 MHz). Note that the measurement from GO does not fall on the Gaussian distribution.

Another case is shown in Figure 15 for 0900 UT corresponding to quite different conditions (GO, 76.7°; CH, 73.8°; CO, 68.6°; NQ, 68°), yet fit by the same Gaussian. Again the highest latitude point, does not fall on the Gaussian curve. Both this and the 0545 UT time illustrate the general condition in which the high latitude region is non-Gaussian.

Both times presented illustrate the situation in which three or more measurements can uniquely define three parameters: the Gaussian scale parameter, maximum energy flux, Qmax, and latitude of the maximum Γ max. At both times the Gaussian FWHM=3.2°; however at 0545 UT Qmax=179 MHz and Γ max=72.9° while at 0900 UT Qmax=28 MHz^4 and Γ max=71.5°. Two measurements, on the other hand, can give Qmax and Γ max only if the Gaussian scale parameter is assumed. These 2 and 3 measurement determinations will be identified separately. If only one measurement exists, it supplies a lower limit of Qmax for the Gaussian, but no information about the latitude Γ max. This case exists at 1045 UT.

The above procedure has been applied to the measurements from the 5 sounders at 15 minute intervals from 05 UT to 13 UT on 9 Dec 1971. The resulting values of Γ max and Qmax are plotted in Figure 16 together with the Auroral Electrojet Index. In 23 of these intervals measurements from 3 or more stations are available and so represent cases in which it can be established that a Gaussian latitudinal profile exists and its scale parameter measured. In each of these cases the Gaussian has been found to be 3.2° FWHM. These 23 cases are denoted by the larger black dots in the figure. In 9 of these cases measurements from only 2 stations are available. In these cases the Gaussian scale FWHM of 3.2° has been assumed, and as a result Qmax and Γ max can be determined. An overall test of the consistency of the measurements of the Gaussian scale parameter will be undertaken in the next section as well as a discussion of the merits of assuming its value where no measurement is possible.

In Figure 16 the Auroral Electrojet Index indicates that two isolated substorms occurred, the first commencing near 03 UT, the second, near 09 UT. During the first substorm the latitudinal distribution could not be resolved since measurements were available from only one sounder at the appropriate latitude. Thus the measurements of Q max and Γ max beginning at 0500 UT record the aftermath of the first substorm, the period between the two substorms, and the entire history of the second substorm.

Q max decreases between 0500 and 0900 UT by a factor of about 5 with some minor fluctuations. There is an increase which peaks near 0715 UT which may be related to the perturbation in AE during that period. Γ max decreases by about 3° during this same period, which is to say that the radius of the circle centered at the offset pole which defines the location of Q max increases by 3°.

The flags which appear on two of the points between 06 UT and 07 UT are measurements which exceed the values of Q max determined for the Gaussian distribution and which are denoted by the black dots. Apparently there was a departure from the Gaussian distribution observed near local midnight during this period. This departure was also observed in the ISIS-2 precipitating electron energy fluxes during this period.

After 0900 UT Q max increases by a factor of 25 during the second substorm and subsequently decreases by about the same amount following the substorm. Γ max fluctuates during this substorm but only about ± 1° overall. The lack of large changes in Γ max during the substorm is an interesting result of this study. In addition the changes which occur in Q max and Γ max appear not be be related to one another.

During the large factor of 9 increase in Qmax between 0915 and 0930 UT, Γ max remains constant. On the other hand during the major decrease in Γ max by 2° between 0955 and 1015 UT, Qmax decreases by only a factor of 0.2. The decrease in Qmax following the second substorm is much more rapid than that which occurred following the first, decreasing by a factor of 10 within the hour between 11:15 and 12:15, then decreasing more slowly thereafter. Γ max increases by ~ 2.5° between 11:15 and 1300 UT.

Consistency of Gaussian Function

The Gaussian distribution has been determined in each of the 23 cases when 3 or more stations provided simultaneous measurements. These are independent determinations which find the Gaussian to be constant 3.2° FWHM throughout the period considered. It is possible to describe the consistency of these indpendent measurements in the following way: each of the values of Q measured by the individual sounders have been normalized to the value of Q max determined for that case and the Γ max determined for that case redefined as the origin, i.e. Q/Qmax has been plotted versus $\Gamma - \Gamma$ max for each measured value of Q.

These have been further differentiated as to whether or not they occurred during substorm periods (8 cases) or non-substorm periods (15 cases). Figure 17 is the totality of all the measured values plotted in this way for the 15 non-substorm cases. The test for consistency in the data as a whole is the degree to which the measurements are in agreement with one another. The agreement is quite good - as a whole the measurements conform quite well to the same 3.2 FWHM Gaussian defined in the individual cases and do so down to a factor of about 30 below the peak.

A similar plot for the substorm cases also appears in Figure 17. For the high latitude half of the distribution agreement of the measurements with the 3.2° FWHM Gaussian is about as good as in the non-substorm case. For the low latitude, however, agreement exists down to about the half maximum. Near $\Gamma - \Gamma_{max} \sim -2.5°$ the measurements fall below the Gaussian by about a factor of 100, an asymmetry which apparently occurs only during the substorm.

As a whole and particularly during the non-substorm cases the constancy of the Gaussian distribution is established by the measurements at individual times and in combination. The rationale for assuming that the same Gaussian exists in the 9 cases where it could not be determined is that there is no departure from the value in the 70% of the cases where measured; so it is reasonable to assume the same value in the 30% of the cases where the Gaussian is not measured.

SUMMARY

We have described the particles which produce the continuous (diffuse) aurora for relatively quiet conditions in the midnight sector. Electron spectra are Maxwellian at energies above a few hundred eV. The characteristic energy decreases with latitude poleward and equatorward of the center of the distribution. Except near the peak, the angular distribution of electrons is not isotropic; α increases with pitch angle. Integration over the loss cone results in Maxwellian spectrum with α near 1 keV near the center. Precipitating positive ions (protons) are present at the same latitudes as the electrons, although the latitude at which they peak is ~1° lower. They are isotropic and are approximately Maxwellian with α's decreasing poleward of the peak ionosphere ($\alpha \sim 8$ keV). At this latitude, the positive ions accounted for at least 20% of incident particle energy. Sounder measurements of virtual height and critical frequency provide quantitative measures of the electron spectral energy and total energy flux respectively. The nature of the incident electrons and positive ions suggests that the continuous aurora and therefore the auroral-E layer represents the low-altitude termination of the central plasma sheet.

The latitudinal distribution of the auroral-E/continuous aurora energy flux is found by a variety of measurements to be characteristically Gaussian. Longitudinally this distribution is ordered in a frame of reference, which has as its origin, a pole which is offset from the CG pole, based on statistical studies. These latitudinal and longitudinal properties are observed and quantitatively determined instaneously by means of simultaneous measurements of foEs by 5 ionospheric sounders arrayed across 10 hours of CG local time and 12° of CG latitude. Each set of simultaneous measurements, when their locations are transformed to the reference frame of the offset pole at 85.8° CG latitude and 00h CG local time, defines a Gaussian distribution in energy flux of 3.2° FWHM, implying that iso-energy flux contours are circles centered at the offset pole. Such a set of measurements has been analyzed at each 15 minute interval for a continuous period of 8 hours during which the Gaussian shape remains constant but the maximum energy flux, Qmax, and the latitude of the maximum in the offset reference frame undergo considerable variation. As a result of conditions during this period which range from quiet to substorm, Qmax varies by a factor of 28 overall (between ~0.25 and ~7 ergs/cm^{2}sec). Γ_{max} (hence the radius of the circle which defines Qmax) changes by nearly 5° overall due to effects which may be both UT and substorm dependent. Two regions depart from the Gaussian distribution: the first, which appears during quiet conditions, is at latitudes 4° and more poleward of the maximum where energy flux is greater than the Gaussian; the second, which appears during substorm conditions, is at latitudes 2.5° and more equatorward of the maximum where energy flux is less than the Gaussian. Except for these two departures, the dynamic history of the entire nighttime pattern of the continuous aurora/auroral E layer throughout the 8 hour period is described by the two parameters, Qmax and Γ_{max}, of the single Gaussian distribution.

ACKNOWLEDGEMENTS

We especially wish to thank Mrs. R.A. Wagner for her efforts in reading the ionograms. The research was supported in part by Air Force Grant AFOSR 78-3625 and NSF Grant ATM 77-15257. J.R.S. gratefully acknowledges the support of an NRC/AFSC Senior Postdoctoral Research Associateship.

REFERENCES

Benprozvanaya, A.S. and T.I. Shchuka, Distribution of anomalous ionization in the high-latitude E region according to ground based sounding, Geomag. and Aeron., 16, 430, 1976.

Deehr, Charles Sterling, J. David Winningham, Fumihiko Yasuhara, and Syun-Ichi Akasofu, Simultaneous observations of discrete and diffuse auroras by the ISIS2 satellite and airborne instruments, J. Geophys. Res., 81, 5527, 1976.

Holzworth, R.H. and C.-I. Meng, Mathematical representation of the auroral oval, Geophys. Res. Lett., 2, 377, 1975.

Jones, Alister Vallance, Aurora, D. Reidel Publishing Co , Boston, 1974.

Jones, R.A. and M.H. Rees, Time dependent studies of the aurora, 1, Ion density and composition, Planet. Space Sci., 21, 537, 1973.

Lui, A.T.Y., D. VanKatesan, C.D. Anger, S.-I. Akasofu, W.J. Heikkila, J.D. Winningham, and J.R. Burrows, Simultaneous observations of particle precipitations and auroral emissions by the ISIS-2 satellite in the 19-24 MLT sector, J. Geophys. Rev., 82, 2210, 1977.

Meng, C-I., R.H. Holzworth, and S.-I. Akasofu, Auroral circle-deliniating the poleward boundary of the quiet auroral belt, J. Geophys. Res., 82, 164, 1977.

Meng, C.-I., Simultaneous observations of low-energy electron precipitation and optical auroral arcs in the evening sector by the DMSP-32 satellite, J. Geophys. Res., 81, 2771, 1976.

Omholt, A., The auroral E-layer ionization and the auroral luminosity, J. Atmosph. Terr. Phys., 7, 73, 1955.

Pike, C.P., J.A. Whalen, and J. Buchau, A 12 hour case studyof auroral phenomena in the midnight sector: F-layer and 6300 Å measurements, J. Geophys. Res., 82, 3547, 1977.

Roble, R.G. and M.H. Rees, Time-dependent studies of the aurora: Effects of particle precipitation on the dynamic morphology of ionospheric and atomspheric properties, Planet. Space Sci., 25, 991, 1977.

Sharber, James R. and Drazen M. Premate, Low-energy ion penetration into the earth's polar regions, abstract, E.O.S., 61, 342, 1980.

Weber, E.J., J.A. Whalen, R.A. Wagner, and J. Buchau, A 12 hour case study of auroral phenomena in the midnight sector: Electrojet and precipitating particle characteristics, J. Geophys. Res., 82, 3557, 1977.

Whalen, James A., Jurgen Buchau. and Rosemarie A. Wagner, Airborne ionospheric and optical measurements of noontime aurora, J. Atmosph. Terr. Phys., 33, 661, 1971

Whalen, J.A., R.A. Wagner and J. Buchau, A 12 hour case study of auroral phenomena in the midnight sector: Oval, polar cap, and continuous auroras, J. Geophys. Res., 82, 3529, 1977.

Winningham, J.D. and W.J. Heikkila, Polar cap auroral electron fluxes observed with Isis 1, J. Geophys. Res., 79, 949, 1974.

Winningham, J.D., F. Yasuhara, S.-I. Akasofu, and W.J. Heikkila, The latitudinal morphology of 10-eV to 10-keV electron fluxes during magnetically quiet and distrubed times in the 2100-0300 MLT sector, J. Geophys. Res., 72, 3148, 1975.

Winningham, J.D., C.D. Anger, G.G. Shepherd, E.J. Weber, and R.A. Wagner, A case study of the aurora, high-latitude ionosphere, and particle precipitation during near-st ady state conditions, J. Geophys. Res., 83, 5717, 1978.

FIGURE CAPTIONS

Figure 1 Aircraft (heavy line) and Isis 2 (dotted) trajectories on 09 Dec 1971. The Isis 2 pass used in this paper is labeled 0610, the universal time at which it crossed the aircraft flight path. The ionospheric sounder stations are labeled in bold letters.

Figure 2 Isis 2 SPS electron and positive ion data (Orbit 3203) beginning at 0605:04 on 09 Dec 71. The continuous (diffuse) aurora extended from about ~67° to 73° INV. LAT. Seven spin cycles are labeled above the spectrogram for reference.

Figure 3 Electron differential number spectra characteristic of the continuous aurora. The spectra taken during spin cycle 5 are Maxwellian. The legends refer to the Maxwellian solid curves.

Figure 4 Positive ion (proton) spectrum measured during cycle 5. Six spectral measurements within the loss cone were averaged to obtain the spectrum. Quantities in the legend refer to the Maxwellian curve.

Figure 5 Spectrum of precipitating electrons at 1400 km obtained by integrating individual spectral elements over the loss cone.

Figure 6 Integral spectrum for the data of in Figure 5. Only 5% of the electron energy is contained below 600 eV.

Figure 7 Isis 2 measurements of precipitated energy flux at ionospheric levels vs. invariant latitude. The lower solid curve is a Gaussian of 3.2° FWHM.

Figure 8 Isis 2 measurements of Maxwellian characteristic energy α vs. invariant latitude. The solid curve is a Gaussian of 3.5° FWHM.

Figure 9 Aircraft sounder measurements of $(foEa)^4$ fitted with the solid curves of Figure 7, the 3.2° FWHM Gaussian plus the sharper peak.

Figure 10 Altitude of peak ionization production rate for electrons with a Maxwellian spectrum of characteristic energy α at isotropic incidence. Points are from Jones and Rees (1973), Roble and Rees (1977), and D. Strickland (private communication).

Figure 11 Latitudinal distribution of α (h'Ea) obtained from sounder measurements of h'Ea and the ionization production curve of Figure 10. The points are fitted with the same Gaussian of 3.5° FWHM used to fit the particle characteristic energies of Figure 8.

Figure 12 Gaussian latitudinal distributions of precipitated energy flux in the continuous aurora near midnight. Triangles are precipitating particles (electrons). Open circles are derived from photometric measurements, closed circles, from ionospheric soundings.

Figure 13 Local time distribution of: (a) instantaneous measurements of Gaussian latitudinal distributions of precipitated energy flux in the continuous aurora (points), and (b) statistical determination of maximum occurrence of auroral E (solid line).

Figure 14 Latitudinal variation with UT of the ground-based ionospheric sounders in the offset pole frame of reference.

Figure 15 Energy flux Q derived from auroral E critical frequencies plotted versus Γ, latitude in the offset pole reference frame. The simultaneous measurements by the ionospheric sounders conform to the Gaussian distribution in this reference frame even though the location of the sounders are separated by as much as 10 hours of CG local time. AT 0545 UT the measurements are from (in order of increasing latitude) NQ, CH, A/C, GO. At 0900 UT, NQ, CO, CH, and GO.

Figure 16 Variation with UT of the Gaussian latitudinal distribution of energy flux in the continuous aurora measured by the array of ionospheric sounders as defined in the coordinates of the offset pole frame of reference. Top: Γmax, latitude of the maximum energy flux, Center: Qmax, maximum energy flux, and Bottom: Auroral electrojet index.

Figure 17 Totality of measurements on 9 Dec 71 where 3 or more stations permit the determination of Gaussian scale parameter. Individual distributions have been normalized and shifted to common origin to define the latitudinal distribution from the data as a whole. Gaussians of 3.2° FWHM are the solid curves. Two conditions are plotted separately. Left: Non-Substorm (15 cases), Right: Substorm (8 cases).

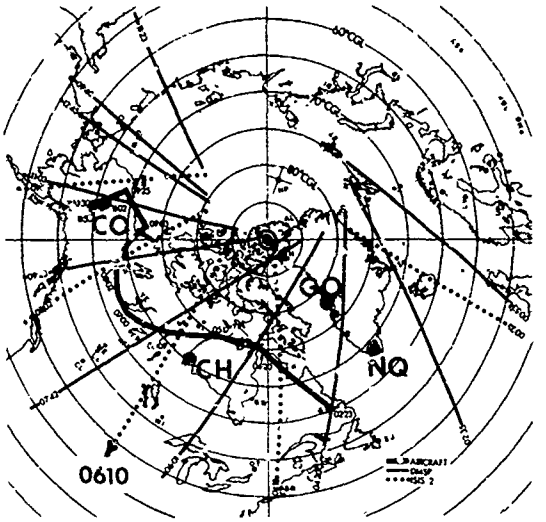


Figure 1

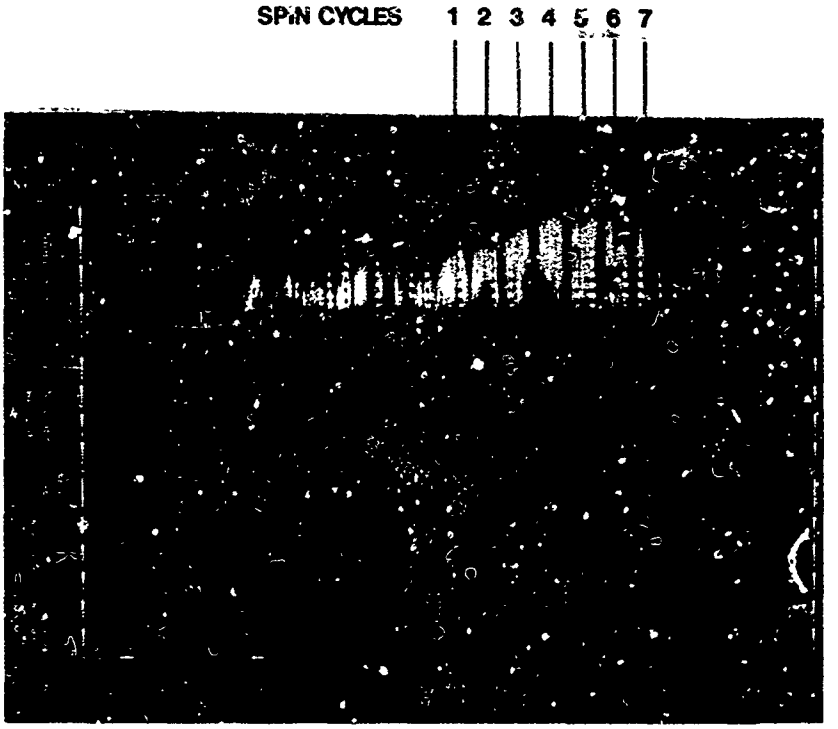


Figure 2

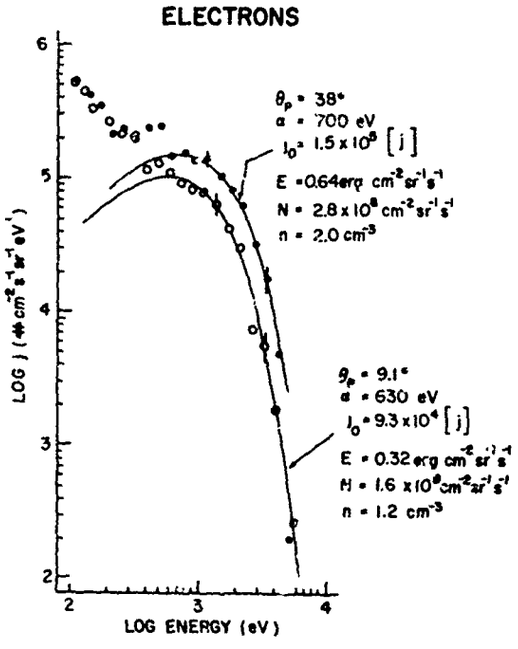


Figure 3

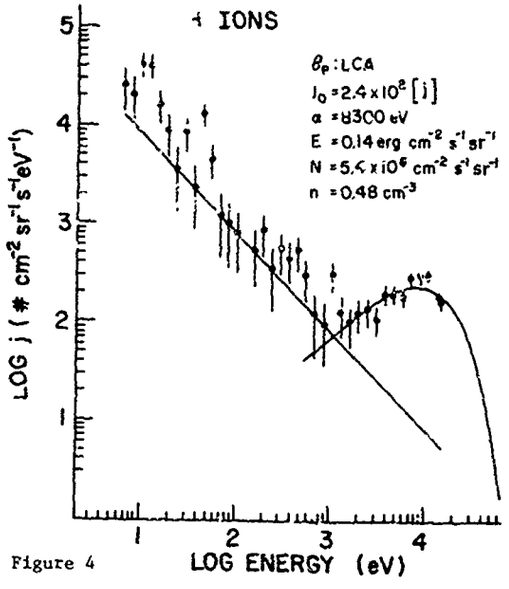


Figure 4

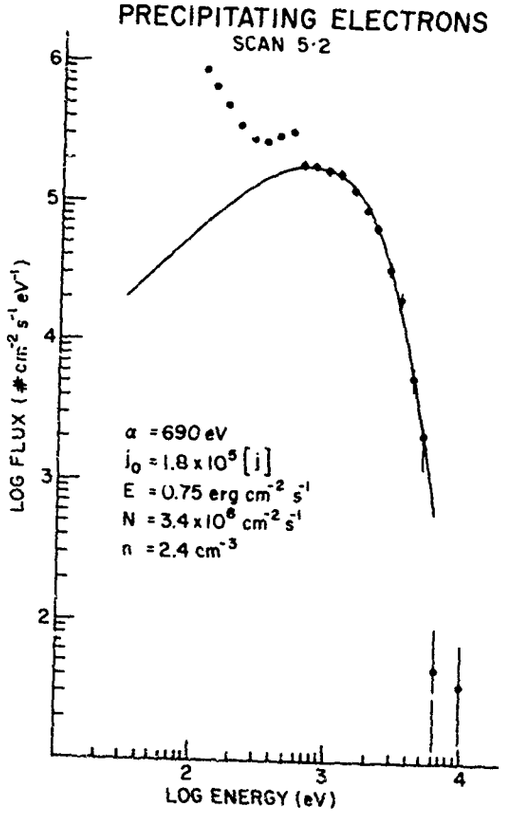


Figure 5

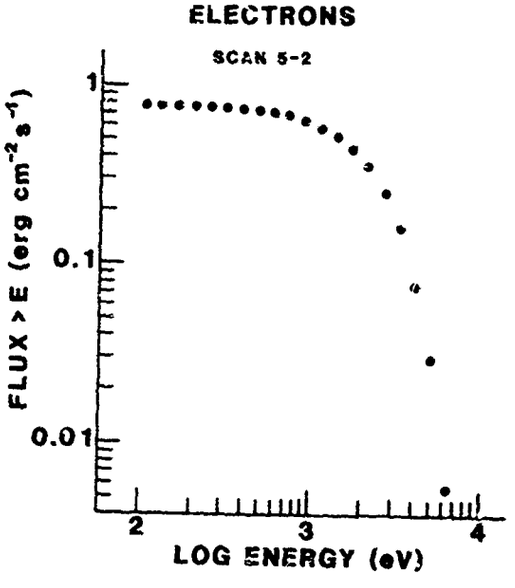


Figure 6

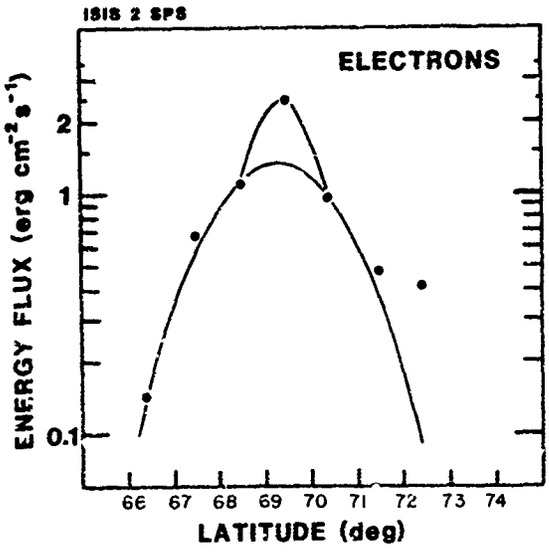


Figure 7

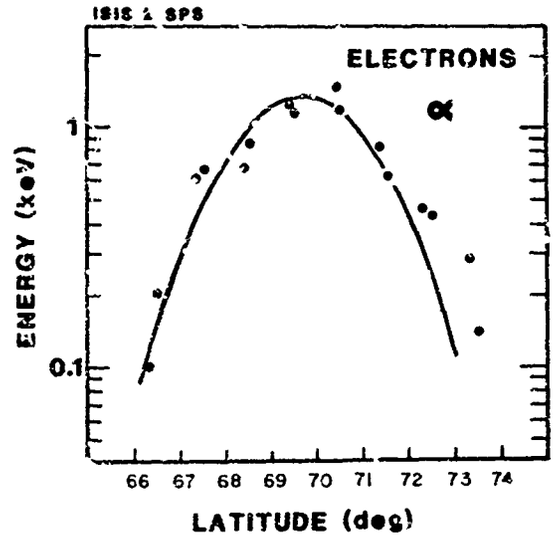


Figure 8

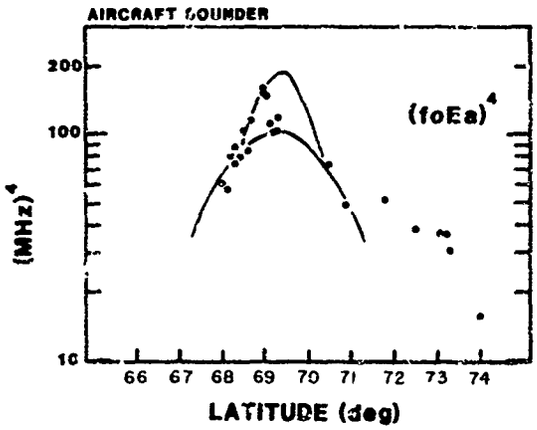


Figure 9

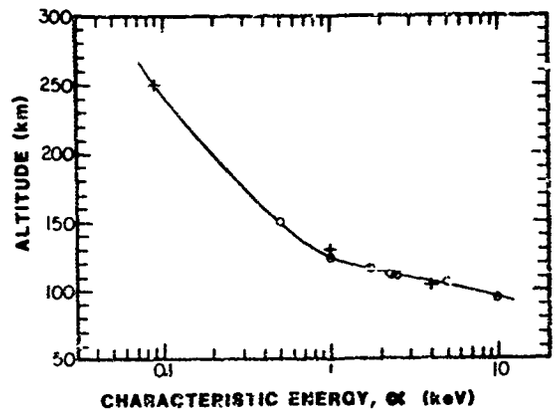


Figure 10

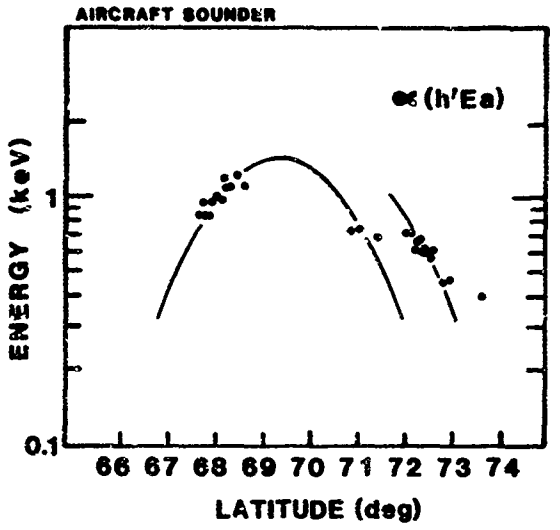


Figure 11

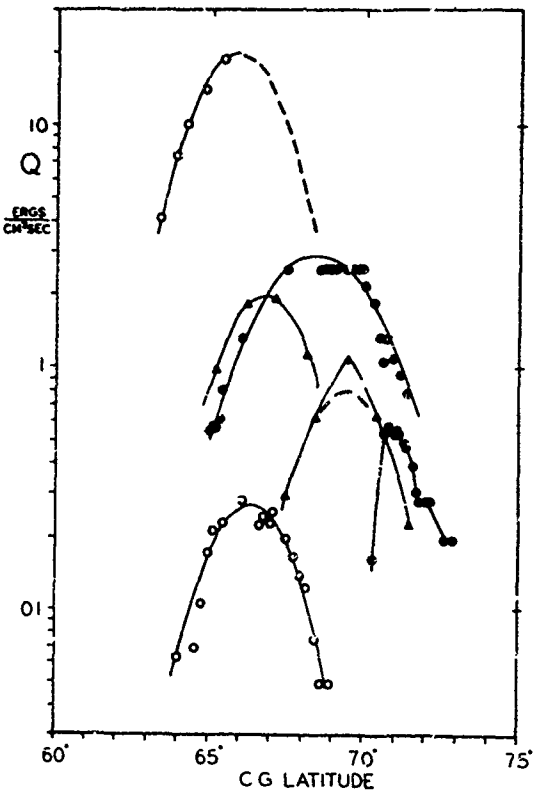


Figure 12

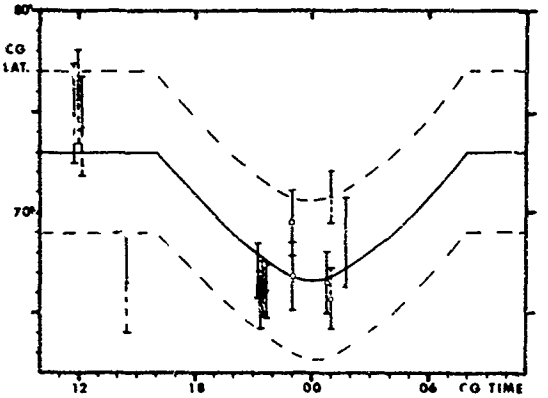


Figure 13

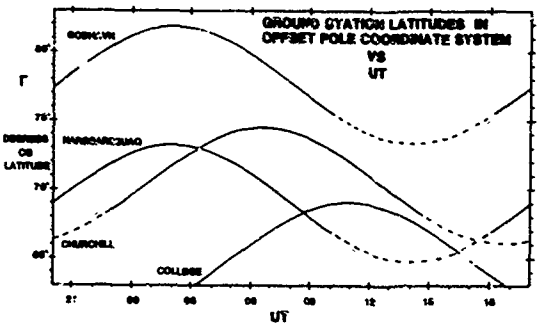


Figure 14

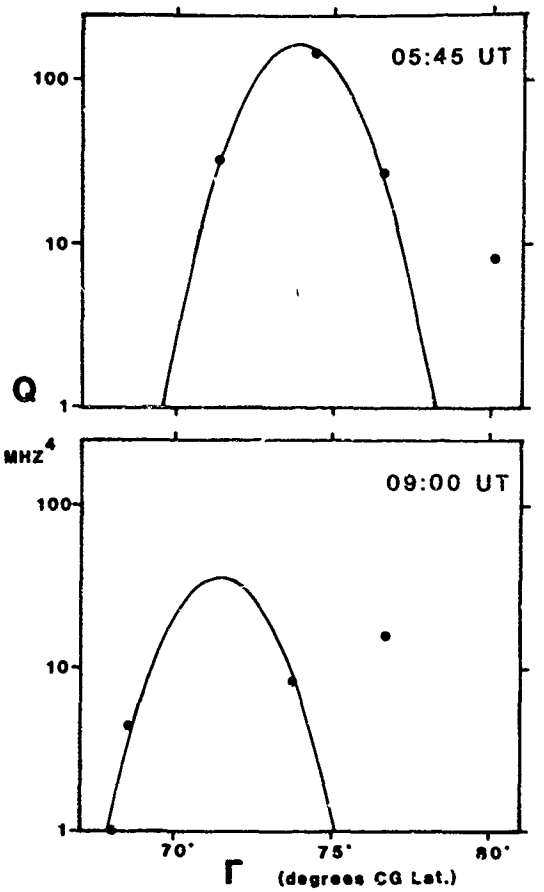


Figure 15

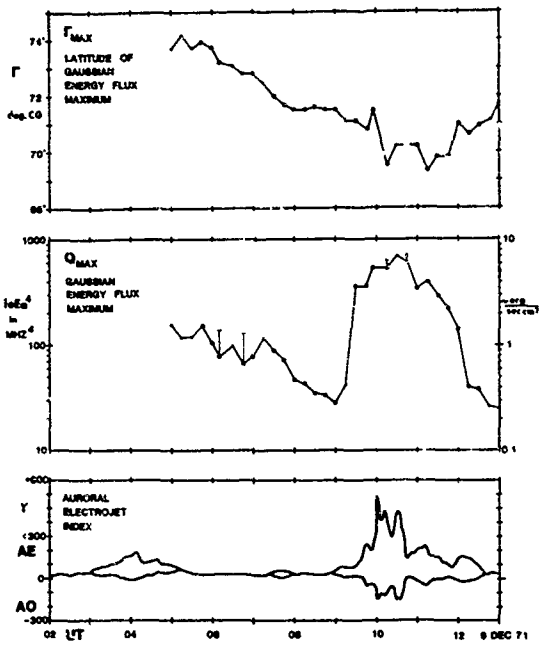


Figure 16

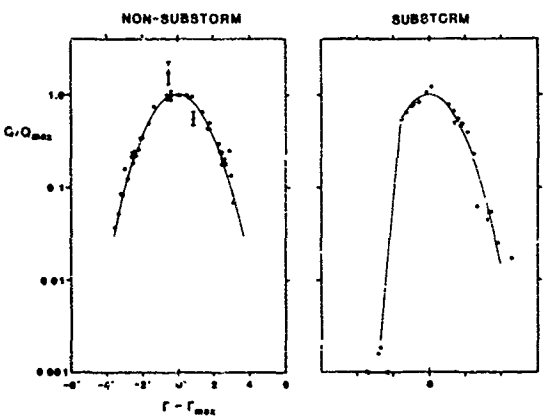


Figure 17

THE STP/S3-4 SATELLITE EXPERIMENT: HIGH LATITUDE LARGE SCALE DENSITY IRREGULARITIES

P. Rodriguez, M. Singh[+], E.P. Szuszczewicz, D.N. Walker and J.C. Holmes
E.O. Hulburt Center for Space Research
Naval Research Laboratory
Washington, DC 20375

[+] Sachs/Freeman Associates, Inc.
Bowie, MD 20715
On leave from Punjabi University, Patiala, India

ABSTRACT

Large scale density irregularities in the nighttime auroral zone F-region are routinely detected by a pair of pulsed plasma probes on the S3-4 satellite. The absolute density variations can be as large as an order of magnitude and in the case of a quiet diffuse aurora, the irregularities appear to be consistent with the sheet-like structures that have been postulated to explain high latitude scintillation enhancements. In a more dynamic situation, which we believe to be a surging aurora, the density variations may be associated with a density gradient- and/or current-driven plasma instability. We have made preliminary analyses of the density fluctuations and FFT power spectra for evidence of characteristic scale sizes and power law dependence. Scale sizes from 10-300 km are clearly evident in the irregularities; power law fits to the spectra have spectral indices in the range -1.5 to -1.7. For the diffuse aurora our results suggest support for a recent numerical study of the nonlinear evolution of the current convective instability.

INTRODUCTION

The polar regions of the earth's ionosphere are known to exhibit a wide range of plasma kinetic interactions that are very dynamic and complex. The plasma kinetic processes have interaction scale sizes that range from the order of auroral longitudinal dimensions down to particle gyroradii. Although the phenomenology and morphology of auroral processes are well known, the physical mechanisms are not well understood. Plasma density irregularities in the auroral F-region have been detected through ground-based scintillation observations (Aarons, 1973; Fremouw et al., 1977; Buchau et al., 1978; Rino et al., 1978), satellite measurements (McClure and Hanson, 1973; Dyson et al., 1974; Sagalyn et al., 1974; Phelps and Sagalyn, 1976; Clark and Raitt, 1976; Weber and Buchau, 1981), and rocket observations (Kelley et al., 1980). Several physical mechanisms have been suggested to explain the high latitude irregularities. A recent review (Fejer and Kelley, 1980) lists three general sources of the irregularities: 1) particle production (especially by low energy precipitating electrons), 2) electrostatic turbulence, and 3) plasma instabilities. Except in extremely simple situations, it is likely that all three mechanisms will contribute to the spectrum of irregularities.

In this report we will examine two cases of high latitude measurements of electron density irregularities obtained in the nightside auroral F-region with the pulsed plasma probe experiment on the STP/S3-4 satellite. These two cases suggest the importance of plasma instabilities associated with particle precipitation, i.e, field-aligned currents, and illustrate the variety of density irregularities encountered at high latitudes. The first observation we discuss is that associated with a quiescent aurora, in which the structure of density irregularities is relatively well defined on several scale lengths, from 10 km to about 300 km. These scale lengths are significant for some recent theoretical studies on the applicability of the current convective instability in diffuse auroras. The second observation we discuss was obtained during a more dynamic state of the aurora and serves to illustrate some of the complexity associated with possible gradient-

and/or current-driven plasma instabilities in discrete auroras.

MEASUREMENTS AND ANALYSIS

Plasma Density Measurements

The polar-orbiting STP/S3-4 satellite is instrumented with a pair of pulsed-plasma-probes that allow high resolution measurements of electron density N_e, temperature T_e, density fluctuation power spectra $P_{N_e}(k)$, plasma potential V_∞ and variations in the mean ion mass $\delta\bar{M}_i/\bar{M}_i$ [Szuszczewicz et al., 1981]. We briefly review the experimental operation only for the particular data mode of interest here. In the low-data-rate mode, the double probe system sampled electron- and ion- saturation currents simultaneously at 100 samples s^{-1}. Under conditions of fixed-payload potential and stable velocity-field and magnetic-field aspect (see e.g., Szuszczewicz and Takacs, 1979), the density N_e and density fluctuation ΔN_e are directly proportional to the probe saturation current I_e and its fluctuations ΔI_e. The STP/S3-4 satellite was in a sun synchronous orbit with an inclination of 96.5° and crossed the equatorial plane at local times of 2230 and 1030. The altitude ranged from 160 to 260 km. In the discussion below of irregularity scale lengths, we will use an average spacecraft velocity of 8 km s^{-1} to convert time variations to spatial variations, using $\Delta L = v\Delta t$, where ΔL is the spatial length, v is the spacecraft velocity, and Δt is the sampled time interval.

Quiescent Aurora

About five minutes of saturation currents, I_e and I_i, obtained on orbit 330 are shown in Figure 1. These measurements were obtained in the night side auroral F-region over the south pole at an altitude of 260 km with the spacecraft moving from the south polar cap region toward the nightside equator (left to right in Figure 1). The probe currents clearly show that irregularities were detected on both electron and ion probes. The detailed tracking of one probe by the other indicates that the probe system was working properly and that the irregularities were due primarily to variations in the ambient electron and ion densities and not the result of secondary effects such as changing spacecraft potential, aspect sensitivity, etc. In this preliminary report, only the electron density irregularities will be discussed, and we simply note the correlative behavior of the ion density irregularities.

The large depletion in electron density (almost an order of magnitude) centered at 1840 UT corresponds to a diffuse aurora. The geomagnetic Kp index for the period of this measurement was 1+, indicating relatively quiescent auroral conditions. The ion density shows a similar depletion. Both the electron and ion depletions extend over a time interval of about 25 s, corresponding to a distance of about 200 km along the approximately south-north direction of spacecraft motion. This aurora was also detected by Huffman et al. (1980) with the vacuum ultraviolet (VUV) experiment on board the same spacecraft. A detailed comparison of the density measurements in Figure 1 with the VUV measurements (Fig. 15 in Huffman et al.) indicates that a good correlation exists. A few discrete arcs detected by the VUV measurement in the time interval 1839-1840 UT appear to be correlated with some of the density structure. Equatorward (to the right) of the large depletion, the electron density shows a relative enhancement with a scale length of about 300 km.

A more detailed plot of the electron density depletion of Figure 1 and its contiguous domains is shown in Figure 2. (The expanded interval is indicated by the arrows above the electron current trace in Figure 1.) In this plot every data point sampled at 100 s^{-1} is shown and the density scale on the left of the figure is approximate. Within the depletion at least three well-defined density variations with nearly periodic behavior are indicated. Assuming the auroral structure to be relatively stable, the width of the three density structures along the spacecraft trajectory is about 20 km and the large scale depletion has a width of about 200 km. On either side of the large scale depletion the density irregularities generally appear to have high and low frequency components with the low frequency components having scale widths comparable to the widths of the density structures within the depletion, i.e., about 20 km.

The Fast Fourier Transform (FFT) power spectrum of the density irregularities in each of three intervals are shown in the insets to Figure 2. The data time interval over which the FFT was calculated is indicated below each spectrum. The Nyquist frequency for each spectrum is 6.25 Hz, corresponding to a minimum scale size of about 0.6 km. Each spectrum is normalized to the maximum power calculated from the FFT. The straight line drawn through each spectrum is the least-squares power law $P=P_o f^n$, where P/P_o is the normalized power, f is the frequency, and n is the spectral index. The spectra are calculated without smoothing, except for the removal of the lowest frequency (DC) component by applying a linear detrend.

The first power spectrum (on the left) suggests a k-dependence of $k^{-1.48}$, where $k = 2\pi/\lambda$ is the wave number and λ is the wavelength of the spatial irregularity. The value of the index is consistent with previous reports (Dyson et al., 1974; Sagalyn et al., 1974; Phelps and Sagalyn, 1976; Kelley et al., 1980) on high latitude irregularities where scale sizes from less than 1 km up to 200 km have been considered. In the second power spectrum, taken within the large scale depletion, the spectral index implies a $k^{-1.58}$ variation in spatial irregularity. The low frequency spectral range 0.1-0.3 Hz corresponds to the well-defined structures of about 20 km half-width in the density-vs-time plot. The spectral power in the 20-km structures inside the large density depletion is about three times greater than the power in similar scale size structures outside the depletion, a fact not obvious in the power spectra because of the normalization. The third power spectrum shows a power law dependence with a spectral index of -1.69.

An examination of the density-vs-time plot of Figure 2 shows that approximately 20-km scale size structures are clearly evident both inside and outside the large depletion. A possible source of these 20-km density irregularities is suggested by recent theoretical studies on the current convective instability in diffuse auroras (Ossakow and Chaturvedi, 1979; Keskinen et al., 1980; Huba and Ossakow, 1980). The current convective instability is proposed as a mechanism for generating large scale ($\sim$ 10-50 km) density irregularities in the diffuse auroral F-region. The current convective instability may thus provide an explanation for the sheet-like density irregularities that have been detected by Fremouw et al. (1977) and Rino et al. (1978) in their analysis of enhanced scintillations at auroral latitudes. In a numerical analysis, Keskinen et al., (1980) considered the nonlinear evolution of the current convective instability with an initial density gradient scale size of 50 km. They obtained a spectrum of irregularities with a spectral index of -1.7 in the north-south direction. This scale size and spectral index are similar to the scale lengths and to the spectral indices shown in Figure 2. Thus we suggest that the observations provide experimental support for the current convective instability as a possible mechanism for the density irregularities associated with the diffuse aurora.

Dynamic Aurora

The plasma probe measurements of Figure 3 illustrate the case of a dynamic aurora observed on orbit 244 in the southern night-side auroral F-region. The K_p index for this period was 3+, indicating that moderately active geomagnetic perturbations were occurring. As in Figure 1, the electron and ion saturation currents were sampled in the low-data-rate mode (100 s^{-1}). The two currents track each other very well so that we are confident of proper measurements directly proportional to the ambient plasma density. The outstanding features of the density measurements in Figure 3 are the sharp transition in density occurring at about 1915:15 UT and the large peak-to-peak variations between 1916 and 1917 UT. The sharp transition in density is more apparent in the electron current than in the ion current because of noise interference occurring in the ion probe before about 1915:10 UT. A detailed comparison of the two currents shows that the transition itself is unaffected by the noise and the two currents track each other almost exactly within the transition.

A detailed plot of the electron current in the time interval indicated by the arrows in Figure 3 is shown in Figure 4 with an approximate density scale to the left. The sharp transition begins at about 1915:17 UT and consists of large irregular oscillations over the transition width of about 1.5 s. The density jump associated with the transition is about a factor of 4. Immediately behind the transition there are at least four distinct oscillations with an average period

about 2 s. These oscillations seem to decay exponentially, as indicated by the dashed line (on a log scale, an exponential curve will appear as a straight line). The VUV observations on the same spacecraft indicate that auroral emissions were detected at the same time the density transition occurred (R.E. Huffman, private communication) suggesting that the density jump is associated with a discrete auroral arc. The observation of a sharp transition suggests that a moving boundary has been detected and we conclude that the density jump is associated with the surge of an auroral arc. The surge direction cannot actually be determined from our observations; generally the movement is westward and poleward (Akasofu, 1968). The sharp transition in density implies that gradient-driven plasma instabilities are likely to occur; with our interpretation of a surging auroral arc, the decaying oscillations behind the density jump would represent spatial oscillations left behind by the movement of the source of instability, the surge front. The electron density increase by a factor of 4 across the surge front probably results from the enhanced ionization due to precipitating particles. Similar ionization enhancements of the F-layer associated with discrete auroral arcs have been observed with ionospheric sounders (Pike et al., 1977). The surge front very likely has field-aligned currents associated with it (Iijima and Potemra, 1978), so that current-driven plasma instabilities are also likely to be present and may contribute to the density oscillations. It is possible that the large and irregular variations seen at later times (after 1915:55 UT) also may have had their origin in the surge of the auroral arc. If the irregular density structures behind the surge front are quasi-stationary then a wide range of scale lengths is evident, from about 4 km to 70 km and longer. In contrast, the region ahead of the surge front has a distinctly different appearance, with only low amplitude, high frequency variations. This region ahead of the surge front may be a relatively quiescent ionosphere about to be perturbed by the surging auroral arc. The different character of the irregularities on either side of the density jump that we have identified as a surging auroral arc is additional evidence that the density jump is a dynamic transition between two states of the F-region ionosphere.

The power spectra for three time intervals are also shown in Figure 4. Just ahead of the density jump the spectrum of the irregularities shows that the slope of the least-squares power law is relatively flat, suggesting that thermal fluctuations dominate the spectrum and the ionosphere is in a quiescent state. The spectrum immediately behind the density jump shows a power law with steeper slope, indicating the presence of enhanced power at low frequencies. The well-defined decaying oscillations behind the density jump appear in the spectrum as the component centered on a frequency of about 0.3 Hz. Under the assumption that they are spatial variations, these oscillations have a scale length of about 10 km. The spectral index $n = -1.66$ for the second spectrum in Figure 4 is nearly identical to one of the spectra in Figure 2. This near-equality suggests that similar transfer processes in spectral energy may be involved. (We note, however, that a specific spectral index does not uniquely define an instability process.) The third spectrum in Figure 4 corresponds to the region of large peak-to-peak variations. The spectrum clearly shows the predominance of structure in the range 0.1-1.0 Hz (40-4 km). The spectral index is near -2, possibly due to the presence of steep edges in the irregularities.

CONCLUDING COMMENTS

The aurora-related F-region density irregularities discussed in this report illustrate the variety of plasma kinetic processes that are involved in high latitude ionospheric effects. Our study of the density irregularities is in a preliminary stage, and we expect that new and significant results will be forthcoming as the study advances. It is already obvious from an initial survey of available satellite data that aurora-related effects are detected outside the average auroral boundaries, i.e., both at higher and lower latitudes. Since the high latitude region is directly associated with magnetosphere-ionosphere coupling processes, we anticipate that the analysis of the ionospheric density irregularities in regions such as the polar cusp will address problems of global importance.

ACKNOWLEDGMENT

The authors wish to thank the Air Force Space Test Program, the Naval Space Systems Activity and the Office of Naval Research for support throughout this program. We are also indebted to Messrs. Swinney and Kegley for their diligent efforts in all phases of fabrication, testing and integration. This work was

supported within the Ionospheric and Stratospheric Task Area (0949-0) and conducted within the Ionospheric Diagnostics Section of the Space Science Division.

REFERENCES

Aarons, J., "A descriptive model of F-layer high-latitude irregularities as shown by scintillation observations", J. Geophys. Res., 78, 7441 (1973).

Akasofu, S.-I., "Polar and magnetospheric substorms", Springer, New York (1968).

Buchau, J., J. Aarons, J.P. Mullen, E.J. Weber, J.A. Whalen, H.E. Whitney and E.E. Crampton, "Amplitude scintillation studies in the polar region on 250 MHz", in Effects of the Ionosphere on Space and Terrestrial Systems, edited by J.M. Goodman, U.S. Government Printing Office, Washington, DC (1978).

Clark, D.H., and W.J. Rait, "The global morphology of irregularities in the topside ionosphere as measured by the total ion current probe on ESRO-4", Planet. Space Sci., 24, 873 (1976).

Dyson, P.L., J.P. McClure and W.B. Hanson, "In situ measurements of the spectral characteristics of F-region ionospheric irregularities", J. Geophys. Res., 79, 1497 (1974).

Fejer, B.G. and M.C. Kelley, "Ionospheric irregularities", Rev. Geophys. Sp. Phys., 18, 401 (1980).

Fremouw, E.J., C.L. Rino, R.C. Livingston and M.C. Cousins, "A persistent subauroral scintillation enhancement observed in Alaska", Geophys. Res. Lett. 4, 539 (1977).

Holmes, J.C. and E.P. Szuszczewicz, "A versatile plasma probe", Rev. Sci. Instr., 46, 592 (1975).

Huba, J.D. and S.L. Ossakow, "Influence of magnetic shear on the current convective instability in the diffuse aurora", J. Geophys. Res., 85, 6874 (1980).

Huffman, R.E., F.J. LeBlanc, J.C. Larrabee and D.E. Paulsen, "Satellite vacuum ultraviolet airglow and auroral observations", J. Geophys. Res. 85, 2201 (1980).

Iijima, T., and T. A. Potemra, "Large-scale characteristics of field-aligned currents associated with substorms", J. Geophys. Res., 83, 599 (1978).

Kelley, M.C., K.D. Baker, J.C. Ulwick, C.L. Rino and M.J. Baron, "Simultaneous rocket probe, scintillation, and incoherent scatter radar observations of irregularities in the auroral zone ionosphere", Radio Sci., 15, 491 (1980).

Keskinen, M.J., S.L. Ossakow and B.E. McDonald, "Nonlinear evolution of diffuse auroral F-region ionospheric irregularities", Geophys. Res. Lett., 7, 573 (1980).

McClure, J.P. and W.B. Hanson, "A catalog of ionospheric F-region irregularity behavior based on OGO 6 retarding potential analyzer data", J. Geophys. Res., 78, 7431 (1973).

Ossakow, S.L. and P.K. Chaturvedi, "Current convective instability in the diffuse aurora", Geophys. Res. Lett., 6, 332 (1979).

Phelps, A.D.R., and R.C. Sagalyn, "Plasma density irregularities in the high-latitude topside ionosphere", J. Geophys. Res., 81, 515 (1976).

Pike, C.P., J.A. Whalen and J. Buchau, "A 12-hour case study of auroral phenomena in the midnight sector: F-layer and 6300-A measurements", J. Geophys. Res., 82, 3547 (1977).

Rino, C.L., R.C. Livingston and S.J. Mathews, "Evidence for sheet-like auroral ionospheric irregularities", Geophys. Res. Lett., 5, 1039 (1978).

Sagalyn, R.C., M. Smiddy, and M. Ahmed, "High-latitude irregularities in the topside ionosphere based on ISIS 1 thermal ion probe" J. Geophys. Res., 79, 4252 (1974).

Szuszczewicz, E.P. and P. Z. Takacs, "Magnetosheath effects on cylindrical Langmuir probes", Phys. Fluids, 22, 2424 (1979).

Szuszczewicz, E.P. and J.C. Holmes, "The STP/S3-4 ionospheric irregularities satellite experiment', J. Geophys. Res., (submitted, 1981).

Weber, E.J. and J. Buchau, "Polar cap F-layer auroras", Geophys. Res. Lett., 8, 125 (1981).

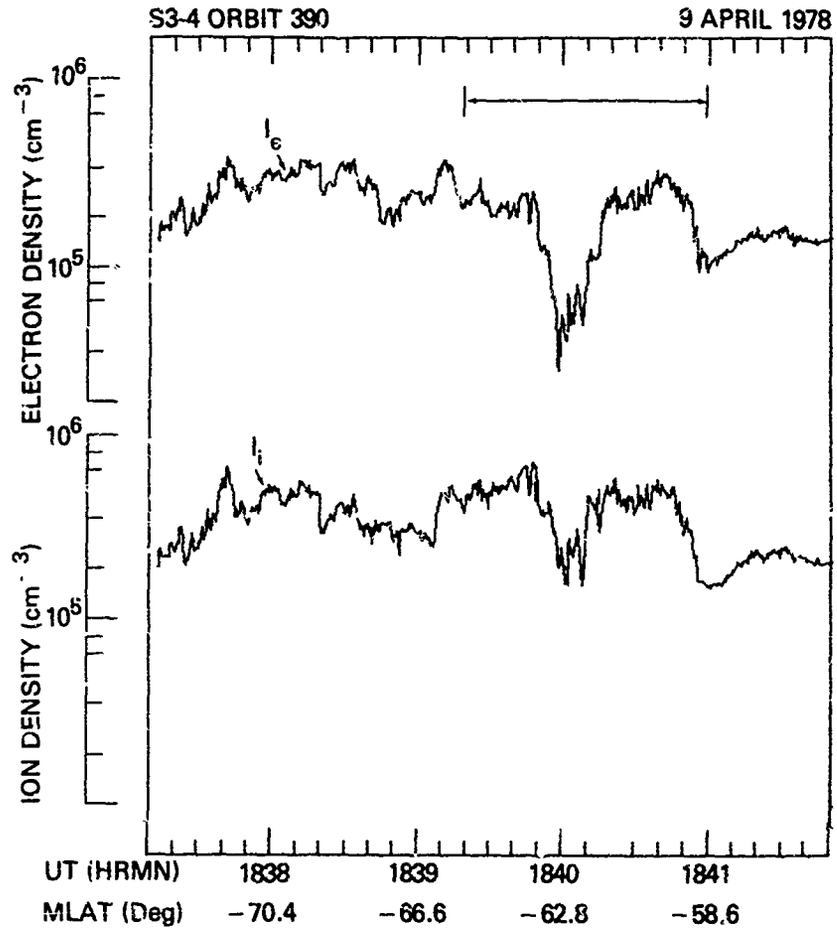


FIGURE 1. The electron and ion currents sampled in the nightside auroral F-region during a pass through a quiescent diffuse aurora.

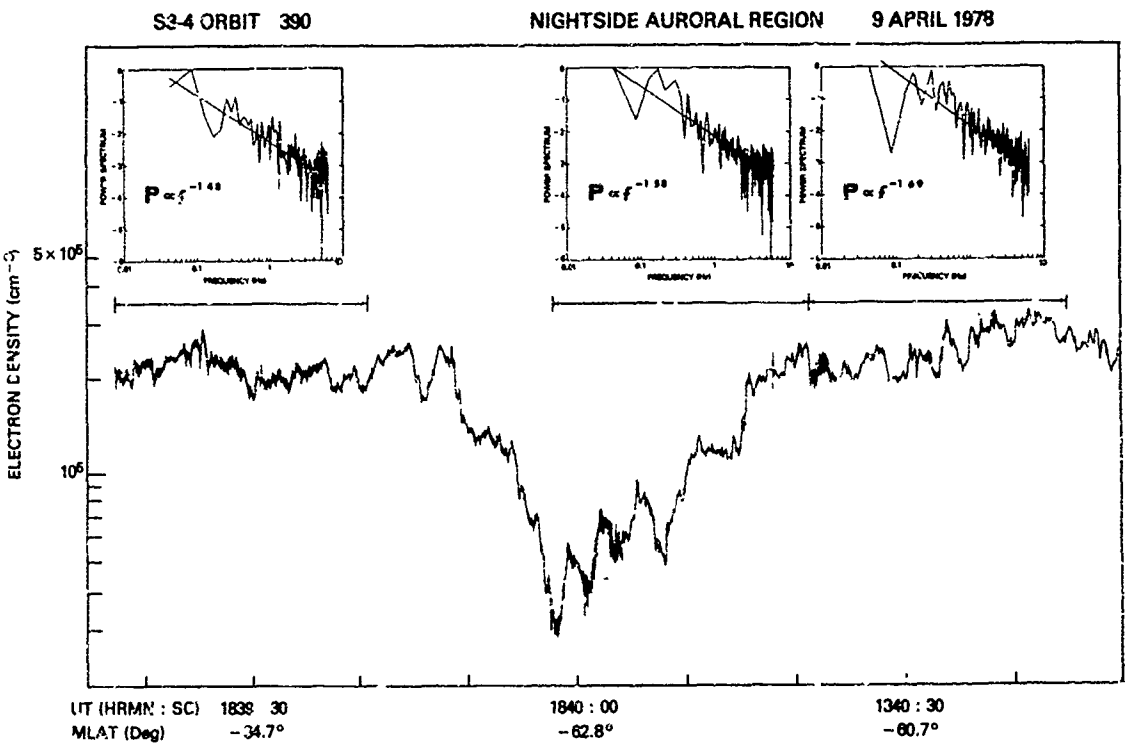


FIGURE 2. An expanded time-scale plot of electron density irregularities in the diffuse aurora of Figure 1. The power spectra indicate the frequency dependence of the least-squares power law. Frequency f is directly proportional to wave-number k, assuming the irregularities are spatial structures.

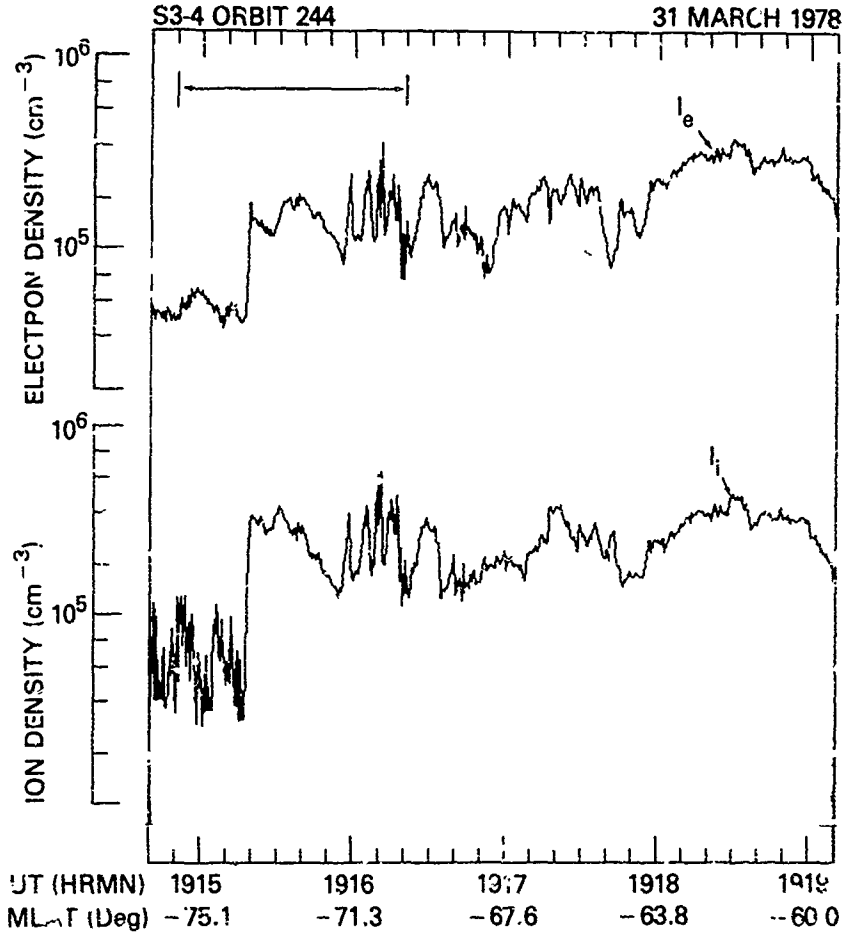


FIGURE 3. The electron and ion currents sampled in the nightside auroral F-region for a dynamic state of the aurora.

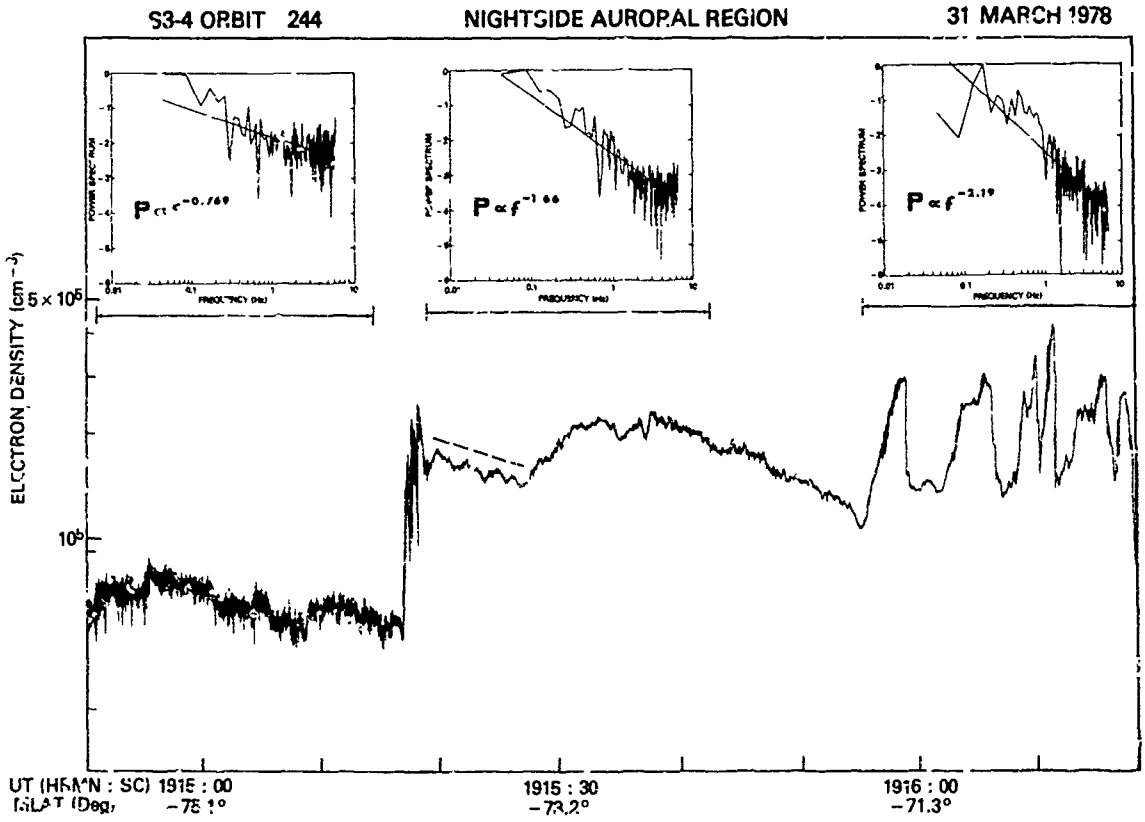


FIGURE 4 An expanded time-scale plot of the electron density for the aurora of Figure 3. The sharp density jump is identified as the surge of a discrete auroral arc. An exponential decay for the well-defined oscillations behind the surge front is suggested by the dashed line.

THE ROLE OF IN-SITU MEASUREMENTS IN SCINTILLATION MODELLING

Sunanda Basu
Santimay Basu
Emmanuel College
Boston, MA 02115

W.B. Hanson
University of Texas at Dallas
Richardson, TX 75080

ABSTRACT

Satellite borne in-situ measurements of irregularity amplitude have been very useful in providing the general features of the global F-region irregularity morphology. Thus, the conflicting equatorial scintillation observations obtained from greatly separated ground stations could be organized in the framework of a longitudinal variation of irregularity occurrence. Similarly, for the current study, high inclination satellite data have been utilized to delineate the essential morphological features of the polar cap irregularity environment. The lack of a significant diurnal and magnetic control of the irregularity morphology within the low solar flux northern winter polar cap ($\Lambda > 80°N$) distinguishes this region from the auroral oval regime. Scintillation results from a polar cap station during the winter season in years of high and low solar flux can be explained on the basis of in-situ observations of irregularity amplitude coupled with the variation of ambient density. A polar orbiting communication system sensitive to phase perturbations may observe large differences in the phase to amplitude scintillation ratio as it traverses through the auroral oval (with its possible magnetic E-W sheet-like irregularity anisotropy) and into the polar cap with its sun-aligned arc system. The current status of scintillation modelling is such that information on the gross variations of the electron density deviation and anisotropy seem to be more important than subtle changes of irregularity spectrum and layer thickness.

1. INTRODUCTION

Ground-based measurements over two decades have established the broad morphological features of three major scintillation regions, two covering the auroral ovals and polar caps and the third one approximately centered on the magnetic equator (Aarons, 1975). However, the detailed morphology has remained unknown because of some obvious handicaps, namely, the limited number of ground stations, the unavailability of suitable sources of VHF-UHF transmissions at certain longitudes and the existence of large ocean surfaces. Satellites carrying out in-situ observations of irregularity parameters present a viable alternative for mapping the irregularity morphology at both high and low latitudes. While only the high latitude morphology was studied by Dyson (1969) and Bagalyn et al. (1974), Clark and Raitt (1976) studied global irregularity morphology over an eighteen-month period. Basu et al. (1976) and Basu and Basu (1980) confined their attention to equatorial morphology only. However their major objective was to convert the observed irregularity morphology into estimated phase and amplitude scintillation models using other realistic model parameters. The resulting models showed in a rather dramatic fashion the longitudinal control of equatorial scintillations during the solstices. Apparently conflicting ground-based scintillation data from far-flung stations (Taur, 1973; Aarons, 1977) could readily be fitted into the synoptic picture of the world-wide equatorial scintillation occurrence pattern that emerged from the in-situ data.

The availability of data from the AE-D satellite with its inclination of 90° provided an ideal platform for the extension of such modelling efforts to the high latitude environment. Since little scintillation data is available from polar cap stations (Aarons et al., 1981) the in-situ model would be able to provide much needed information for global scintillation modelling efforts such as undertaken by Fremouw et al. (1980). Unfortunately, most of the AE-D data base is limited to a 3-month period around the northern winter solstice during 1975-76. Thus the model will be appropriate for northern winter under sunspot minimum conditions. The paucity of the data also prevents us from

studying the seasonal-cum-longitudinal behavior of auroral scintillations discussed earlier by Basu (1975) and Basu and Aarons (1980).

2. MODEL DEVELOPMENT

The ion-drift-meter data obtained from the AE-D satellite have been used for the modelling effort. The drift meter provides the ion density (or electron density for charge neutrality) every 1/16 sec (~500 m) along the orbital track as a function of position and magnetic local time of the satellite as shown by the solid line in Figure 1 which represents a typical AE-D pass through the noontime cusp, polar cap and nighttime auroral oval regions. From these samplings the rms irregularity amplitude $\Delta N/N$ is computed over 3 secs of data (approximately 20 km of path length perpendicular to the magnetic field) at intervals of 8 secs as shown by the open circles. The first object of the modelling process is to determine the average irregularity amplitude as a function of magnetic activity, invariant latitude, and magnetic local time. The irregularity amplitude may be a function of altitude so that separate groupings of the data were made as a function of satellite altitude as well.

The next step is to convert the observed morphology into an estimated model of phase and amplitude scintillations. Scintillations are, of course, proportional to the integrated electron density deviation along the ray path (Rufenach, 1975). In order to estimate this a model of the ambient density N and irregularity layer thickness L are necessary. We further assumed that for propagation studies where only the total integrated effect along the ray path is considered, it is probably quite realistic to assume that the irregularity layer thickness is the same as the slab thickness of the ionosphere under similar geophysical conditions. The ITS-78 model (Llewellyn and Bent, 1973) with suitable updating for sunspot cycle and season was used to determine N and L values as a function of position and local time. Combining the observed rms $\Delta N/N$ values with N derived above the morphology of ΔN was obtained.

In the framework of diffraction theory under the limit of weak scattering, the irregularity morphology can be translated into models of phase and amplitude scintillations. The equations for the determination of σ_ϕ, the phase variance, and S_4, the second central moment of intensity are taken from Rino (1979). Rino (1979) derived the equations for a variable three-dimensional spectral index p of the irregularities. However, the equations take on a simplified form if p is assumed to be 4 which was the most commonly observed spectral index by in-situ techniques (Dyson et al., ; Phelps and Sagalyn 1976). With p=4, these equations are

$$\langle\sigma_\phi\rangle^2 = \frac{1}{2^5\pi^3} (r_e\lambda)^2 \, (L \sec\theta) \, G \, C_s \, (v_{eff}\tau)^2$$

$$S_4 = \frac{1}{2} (r_e\lambda)^2 \, (L \sec\theta) \, C_s \, \left(\frac{\lambda z \sec\theta}{4\pi}\right) F \qquad \text{.....} \qquad (1)$$

where

- r_e - classical radius of the electron (2.8×10^{-15} m)
- λ - radio wavelength
- L - irregularity layer thickness
- θ - ionospheric zenith angle
- C_s - strength of turbulence $= 2^3\pi\langle\Delta N\rangle^2$
- Λ_o - outer scale (~20 km, determined by data interval)
- G,F - geometrical parameters for field aligned irregularities
- z - reduced distance of center of irregularity layer
- v_{eff} - effective scan velocity
- τ - detrend interval

It is important to note in the above equations that in a power law environment with a large outer scale, the data segment over which the rms $\Delta N/N$ is computed effectively sets an outer scale for modelling purposes. The strength of turbulence C_s will thus also depend on the length of the data segment used. Further, while the S_4 index can be uniquely determined in terms of irregularity and anisotropy parameters, the phase scintillation also depends on the detrend interval over which the phase variance is computed and the effective scan velocity of the ray path across the iso-correlcids of electron density. For this modelling effort we considered a detrend interval of 10 secs to make the results compatible with those of the Wideband satellite (Fremouw et al., 1978). For the determination of the effective scan velocity we considered a simple orbital geometry shown in Figure 2. The satellite beacon was assumed to be orbiting in the magnetic meridian plane at an altitude of 1000 km. The projected velocity at F region heights was considered to be 3 km sec^{-1} in the magnetic N-S direction with no component of velocity in the magnetic E-W direction. The modelling was done for an overhead geometry. Three kinds of irregularity anisotropies were considered. In the righttime auroral oval magnetic L-shell aligned E-W sheets were used (Martin and Aarons, 1977; Rino et al., 1978) while in the daytime sub-auroral region field-aligned rods were used

(Fremouw et al., 1980) in keeping with earlier experimental observations. However, no data exists on anisotropy characteristics either in the daytime auroral oval or in the entire polar cap region. Considering the nature of the arc system (Shepherd, 1979), we made the assumption that in the daytime auroral oval the irregularity anisotropy is also in the form of E-W sheets whereas in the polar cap the sun-aligned arc system (Lassen, 1979; Weber and Buchau, 1981) could be associated with N-S sheets. In Section 4, we shall present some results of phase and amplitude scintillation modelling based on such composite irregularity characteristics.

3. HIGH-LATITUDE IRREGULARITY MORPHOLOGY

In Figure 3a, we show the average irregularity morphology observed in the Northern Hemisphere as a function of invariant latitude and magnetic local time for quiet magnetic conditions (Kp < 3.5) obtained between October 22 - December 31, 1975. The irregularity map was obtained by sorting the AE-D data of rms $\Delta N/N$ (expressed as a percentage) in 5° wide invariant latitude and 4 hour magnetic local time boxes and determining the average value of $\Delta N/N$. The 80°-90° invariant latitude range was combined into one box. For Figure 3a, the data was confined to the time periods when AE-D was between 200-400 km while Figure 3b shows the morphology when AE-D was above 400 km altitude. The highest altitude attained by AE-D was approximately 1400 km so that Figure 3b represents the altitude range between 400-1400 km.

The most striking difference between Figures 3a and b is the difference in irregularity amplitudes in the nighttime auroral oval. In this region (65-80°Λ latitude) the high altitude data shows much larger amplitudes while in the polar cap region of open field lines (>80° Λ) the irregularity amplitudes seem to be independent of altitude. The diurnal variation within the polar cap is quite small with the largest irregularities being observed at midday and smallest at midnight. The sharp irregularity boundary at 75°Λ at the equatorward edge of the cusp in the noon sector is another noteworthy feature of these two maps. In addition, the high altitude data shows an asymmetry between the morning and afternoon sectors of the cusp with the morning sector irregularities being larger. Figure 3b is remarkably similar to the map of small scale transverse magnetic disturbances determined by the Triad satellite at 800 km altitude (Saflekos et al., 1978) if we keep in mind the differences in the way the data were sorted in each study.

The irregularity morphology observed during disturbed magnetic conditions (Kp > 3.5) is shown in Figure 3c. Because of the smaller amount of data, an altitude separation was not feasible. The nighttime auroral oval shows a pronounced broadening. However irregularity amplitudes do not attain the large values observed in Figure 3b probably because of both altitudinal and latitudinal smearing. There is a definite equatorward motion of both the nighttime auroral oval and the dayside cusp.

4. MODEL OF PHASE AND AMPLITUDE SCINTILLATIONS

As mentioned in Section 2, the observed morphology of rms $\Delta N/N$ in the high latitude region can be converted to equivalent phase and amplitude scintillation models by using a model of maximum F-region densities to obtain the magnitude of ΔN. Figures 4A and B show maps of electron density deviation ΔN generated by combining the ITS-78 model of densities for the sunspot conditions pertaining to November and December, 1975 with the irregularity amplitudes obtained in the AE-D study presented in Section 3. Figure 4a was determined by combining all the data used in the preparation of Figures 3a and b with the density model while Figure 4b was obtained from Figure 3c and the density model.

The point to note from Figure 4a is the persistence (even in ΔN) of most of the features found in $\Delta N/N$. Thus the polar cap still shows a relatively small diurnal variation with the maximum electron density deviation being observed in the daytime. The cusp is a region of high ΔN while the nightside auroral oval expectedly shows much larger electron density deviations than observed at the same latitudes in the daytime.

Following the procedure outlined in Section 2, Figure 4a is converted into models of σ_ϕ and S_4 by using a composite of three different types of irregularity anisotropy in various sectors of the high latitude region. Only the low Kp data set is used for scintillation models because of the better data coverage. The assumed form of the irregularity anisotropy in the different MLT and invariant latitude sectors are as follows. E-W sheets (10:10:1 with larger dimensions along the field and in the magnetic E-W direction) have been assumed over the nighttime and daytime auroral ovals (60°-80°Λ lat between 18 MLT to 06 MLT and 75°-80°Λ lat between 06 MLT to 18 MLT). Over the polar cap also, E-W sheets have been assumed except for the noontime and midnight sectors (10-14 MLT and 22-02 MLT) where N-S sheets (10:10:1 with larger dimensions along the field and in the magnetic N-S direction) are considered. In the sub-auroral region (60°-75°Λ lat between 06-18 MLT) irregularities in the form of field-aligned rods (with 10:1 ratio) have been assumed. As outlined in Section 2 the irregularity layer thickness has been derived from the slab thickness obtained from ITS-78. These values usually range between 200 and 250 km.

Figure 5a shows the resultant model for amplitude scintillation index (S_4) at 137 MHz under overhead propagation conditions. The non-uniform scaling of ΔN values shown in Figure 4a to S_4 index in Figure 5a arises from variations in the form of anisotropy and magnetic dip. The model developed for sunspot minimum conditions yields values of $S_4 < 0.5$ which thus conform to the weak scatter limit.

The model of rms phase deviation, shown in Figure 5b, corresponds to the reception of 137 MHz signal on the ground from a beacon satellite orbiting in the magnetic meridian plane, with a scan velocity of 3 km/sec at F region heights. The detrend interval for rms phase deviation has been chosen to be 10 sec. It should be noted that in the presence of irregularity anisotropy the rms phase deviation is not only dictated by the electron density deviation (ΔN) but is markedly controlled by the scan velocity with respect to the irregularity axes. In the present model, the scan velocity of the satellite has been considered to be very large compared to the irregularity drift. However, in the general case, the scan velocity of the ray path relative to the irregularity motion needs to be considered. From Figure 5b, the marked effect of the anisotropy with respect to the satellite motion may be noted. For the assumed alignment of the orbital track with the magnetic meridian, the rms phase deviation (σ_ϕ) is accentuated in regions where the anisotropy is in the form of E-W sheets. For example, the high levels of σ_ϕ in the daytime cusp region arises not only from high ΔN values in this region but is due to the N-S orbital motion in a region where the irregularity anisotropy is assumed to be in the form of E-W sheets. This is again demonstrated in the polar cap where the σ_ϕ values are deemphasized in the noon and midnight sectors where N-S sheets are considered and are accentuated in the afternoon and post-midnight sectors where E-W sheets have been assumed to exist. From an observational standpoint, the form of the anisotropy can thus be obtained from the ratio of σ_ϕ and S_4 (Fremouw, 1980). For a satellite orbiting in the magnetic meridian, this ratio (σ_ϕ/S_4) is largest where E-W sheets occur, intermediate where the anisotropy is in the form of magnetic field aligned rods and is minimum where N-S sheets exist.

In Figures 6a and 6b, models for S_4 and σ_ϕ at 250 MHz are shown for identical forms of irregularity anisotropy and constraints of satellite orbit. As is to be expected, the general patterns are similar to those of Figures 5a and 5b with the values of σ_ϕ and S_4 reduced at the higher frequency in accordance with the frequency scaling laws appropriate for a 3-dimensional irregularity spectral index of 4.

5. COMPARISON WITH SCINTILLATION DATA

To determine the effectiveness of the in-situ modelling process presented in the last section, it is important to compare the model with actual ground-based scintillation data. A suitable data base for the comparison with Figure 5a was available at the Air Force Geophysics Laboratory. This was the Narssarssuaq ATS-3 scintillation data obtained during the winter of 1975-76 for which the ionospheric intersection point was 64°Λ. This data formed part of the high latitude morphology published earlier (Basu and Aarons, 1980) and is shown in Figure 7a for quiet magnetic conditions. Little diurnal variation is observed with maximum S4 observed being .22 near midnight. Now if the geometrical factor for ATS-3 observations at 20° elevation is computed for E-W sheets with axial ratios of 10:10:1 and compared with the overhead geometrical factor at the same location for the same anisotropy then a factor of 2.6 enhancement is obtained at the low elevation angle point. Thus the nighttime values shown in Figure 7a when scaled down by a factor of 2.6 are found to be quite consistent with the modeled values close to 0.1. Similar arguments are true for the daytime values.

Rino and Matthews (1980) in their Figure 10 have shown the median S_4 values at 137 MHz as a function of magnetic latitude for the winter of 1976-77 obtained with the Wideband satellite in the nighttime auroral region. Our model was done for the winter of 1975-76 but the sunspot numbers during the two succeeding winters were quite similar. There is good agreement between the model and scintillation observations if one disregards the geometrical enhancement region of the W deband satellite data. The phase scintillation is more difficult to compare because of the idealized orbit chosen for the study.

In the polar cap little data is available for comparison. The only available data during quiet susnpot years is at 250 MHz from Thule, Greenland for the period August-October, 1975. The subionospheric latitude coverage is 82°-90° CG latitude. Even though the geometrical factors are not accurately known for this data set (Aarons et al., 1981) and the periods do not match exactly with the model, we exhibit the S_4 indices obtained at Thule in Figure 7b to show that the modelled values of 0.15 to 0.2 in the polar cap are at least of the right order of magnitude as the data.

6. CONCLUSIONS

The satellite in-situ irregularity measurements provide a direct measurement of electron density deviation (ΔN) parameter which can be used to derive models for amplitude and phase scintillations. In view of the insufficient coverage of ground scintillation observations caused by either the absence of suitable ground locations or satellites, the usefulness of in-situ probing with unlimited latitude and longitude coverage cannot be overemphasized. The evaluations made in the previous

section show that scintillation models based on the quantitative measure of electron density deviation (ΔN) by satellites at high latitudes provide realistic estimates. They also provide a synoptic picture over the northern polar cap which is almost impossible to achieve with ground-based measurements. We have thus been able to provide models of both equatorial and high latitude amplitude and phase scintillations from in-situ data.

It should, however, be mentioned that our current efforts are based on satellites whose primary function was not concerned with irregularity measurements at F region heights for scintillation modelling. As such, the restrictions imposed on satellite altitude, time of transit, etc., limited this data base. A dedicated satellite performing such measurements at F-region altitudes with suitable orbital characteristics will be an ideal vehicle for the development of a world-wide model of phase and amplitude scintillations.

ACKNOWLEDGMENTS

We wish to thank J. Aarons and H.E. Whitney for making the scintillation data available to us and for many helpful discussions. We thank J. Noonan and N. Scotti for assistance with computer programming.

The work at Emmanuel College was partially supported by NASA Grant NSG 5419 and AFGL Contract F19628-81-K-0011.

REFERENCES

Aarons, J., AFCRL Report No. TR-75-0135. Air Force Cambridge Research Laboratories, Hanscom AFB, MA 01731, 1975.

Aarons, J., Antennas Propagat., AP-25, 729, 1977.

Aarons, J., J.P. Mullen, H.E. Whitney, A.L. Johnson and E.J. Weber, Geophys. Res. Lett., 8, 1981 (in press).

Basu, S., J. Geophys. Res., 80, 4725, 1975.

Basu, S. and J. Aarons, Radio Sci., 15, 59, 1980.

Basu, S. and S. Basu, Cospar Symposium Series, 8, 189, 1980.

Basu, S., S. Basu and B.K. Khan, Radio Sci., 11, 821, 1976.

Clark, D.H. and W.J. Raitt, Planet. Space Sci., 24, 873, 1976.

Dyson, P.L., J. Geophys. Res., 74, 6291, 1969.

Dyson, P.L., J.P. McClure and W.B. Hanson, J. Geophys. Res., 79, 1497, 1974.

Fremouw, E.J., J. Atmos. Terr. Phys., 42, 775, 1980.

Fremouw, E.J., R.L. Leadabrand, R.C. Livingston, M.D. Cousins, C.L. Rino, B.C. Fair and R.A. Long, Radio Sci., 13, 167, 1978.

Fremouw, E.J., J.M. Lansinger and D.A. Miller, Report No. PD-NW-81-237R, Physical Dynamics Inc., Bellevue, WA 98009, 1980.

Lassen, K., Geophys. Res. Lett., 6, 777, 1979.

Llewellyn, S.K. and R.B. Bent, Report No. AFCRL-TR-73-0657, Air Force Cambridge Research Laboratories, Hanscom AFB, MA 01731, 1973.

Martin, E. and J. Aarons, J. Geophys. Res., 82, 2717, 1977.

Phelps, A.D.R. and R.C. Sagalyn, J. Geophys. Res., 81, 515, 1976.

Rufenach, C.L., Radio Sci., 10, 155, 1975.

Rino, C.L., Radio Sci., 14, 1135, 1979.

Rino, C.L., R.C. Livingston and S.J. Matthews, Geophys. Res. Lett., 5, 1039, 1978.

Shepherd, G., Rev. Geophys Space Phys., 17, 2017, 1979.

Rino, C.L. and S.J. Matthews, J. Geophys. Res., 85, 4139, 1980.

Saflekos, N.A., T.A. Potemra and T. Iijima, J. Geophys. Res., 83, 1493, 1978.

Sagalyn, R.C., M. Smiddy and M. Ahmed, J. Geophys. Res., 79, 4252, 1974.

Weber, E.J. and J. Buchau, Geophys. Res. Lett., 8, 125, 1981.

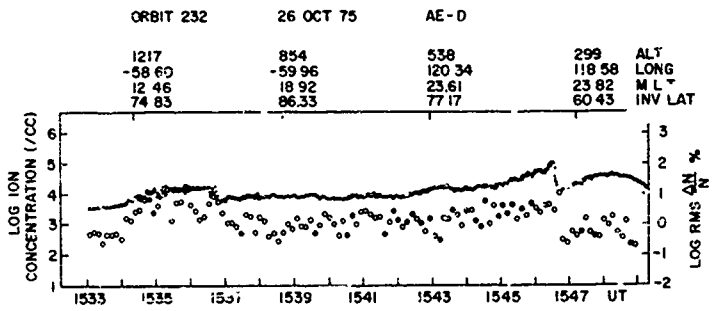
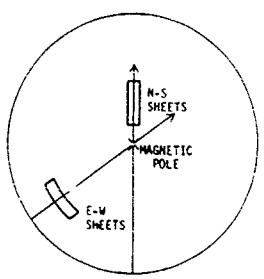


Figure 1. Ion concentration and irregularity amplitude (open circles) obtained by the ion drift meter on AE-D satellite. The altitude (ALT), longitude (LONG) and invariant latitude (INV. LAT.) of the satellite as a function of magnetic local time (MLT) and universal time (UT) are indicated on the diagram.



OVERHEAD SCINTILLATION MODEL FOR SPECIAL SATELLITE ORBIT AND IRREGULARITY CHARACTERISTICS

SATELLITE HEIGHT 1000 km
SATELLITE LOCATION OVERHEAD
ORBIT ALIGNED WITH MAGNETIC MERIDIAN
EFFECTIVE SCAN VELOCITY 3 km/sec
DETREND INTERVAL 10 sec
IRREGULARITY GEOMETRY
1 FIELD ALIGNED ROD
2 E-W SHEETS
3. N-S SHEETS

Figure 2. The idealized orbital geometry of the satellite and the irregularity anisotropy considered in the model are illustrated.

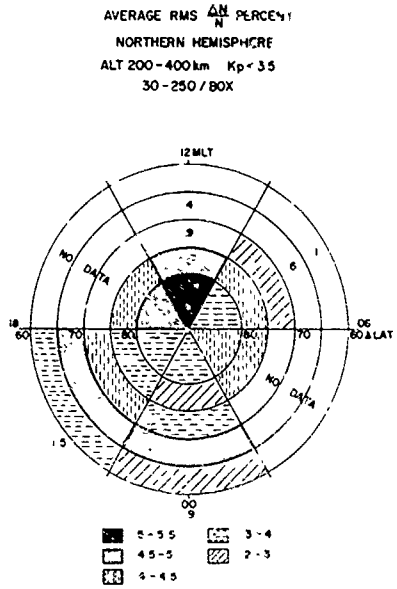


Figure 3a. Variation of the average irregularity amplitude as a function of the invariant latitude and magnetic local time under magnetically quiet conditions obtained from the AE-D satellite in the altitude interval of 200-400 km. The different tones correspond to different irregularity amplitudes expressed in percentage.

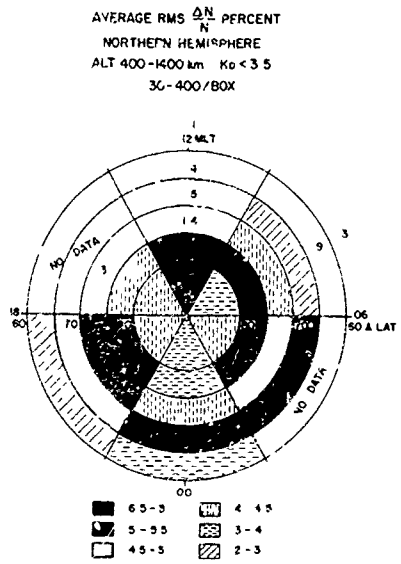


Figure 3b. Same as in Figure 3a except for the satellite altitude interval of 400-1400 km.

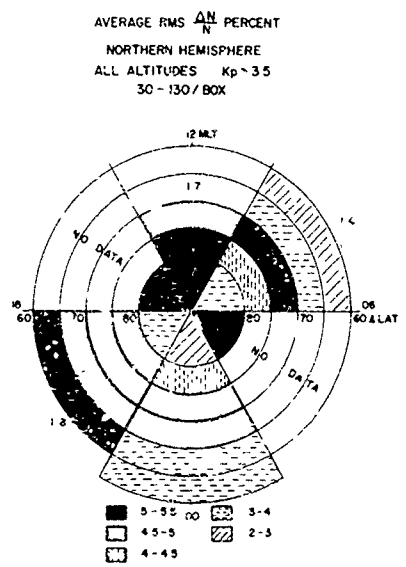


Figure 3c. Variation of the average irregularity amplitude with magnetic local time and invariant latitude under magnetically disturbed conditions (Kp > 3.5) considering all data obtained between 200-1400 km altitude.

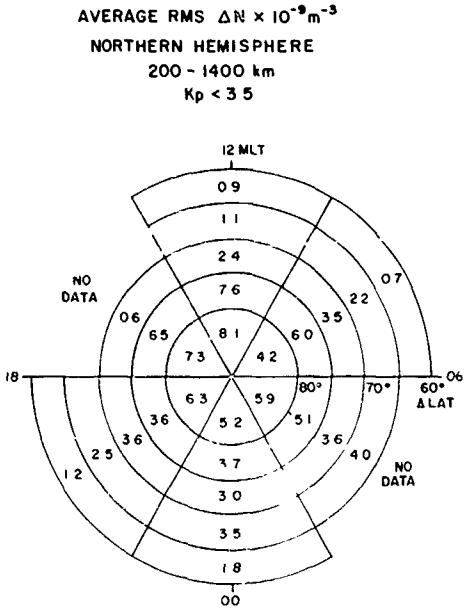


Figure 4a. Morphology of electron density deviation (ΔN) in units of 10^9 m^{-3} under magnetically quiet conditions.

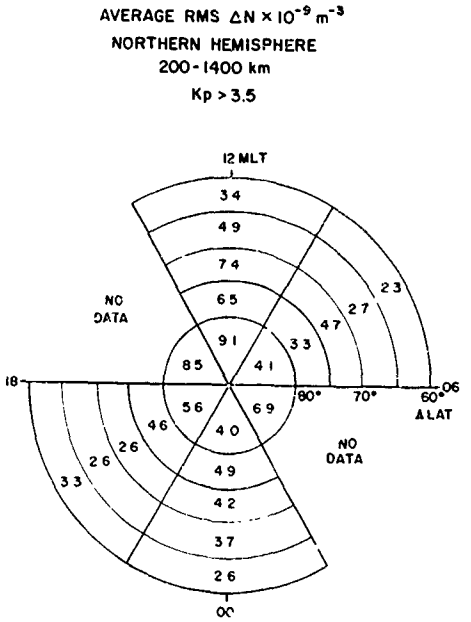


Figure 4b. Same as in Figure 4a for magnetically disturbed conditions.

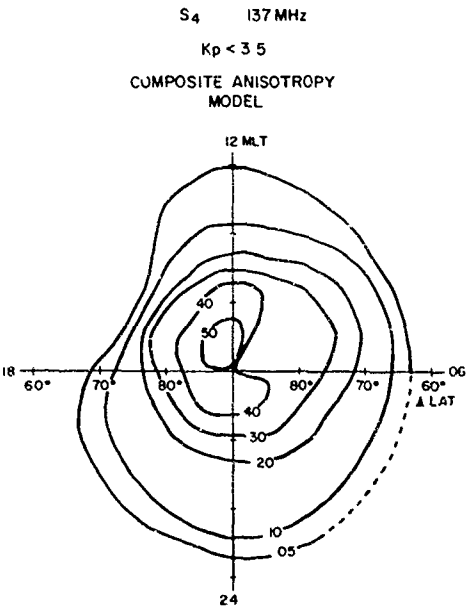


Figure 5a. Model of intensity scintillation (S_4) at 137 MHz under magnetically quiet conditions.

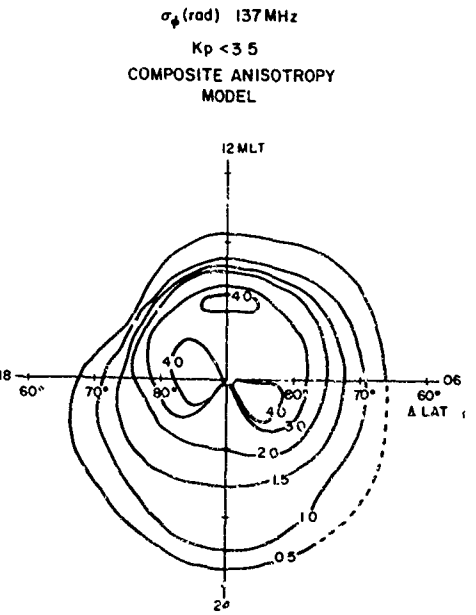


Figure 5b. Model of rms phase scintillation at 137 MHz valid for a beacon satellite orbiting at 1000 km altitude over the magnetic meridian. A composite form of anisotropy (see text) and detrend interval of 10 sec are used.

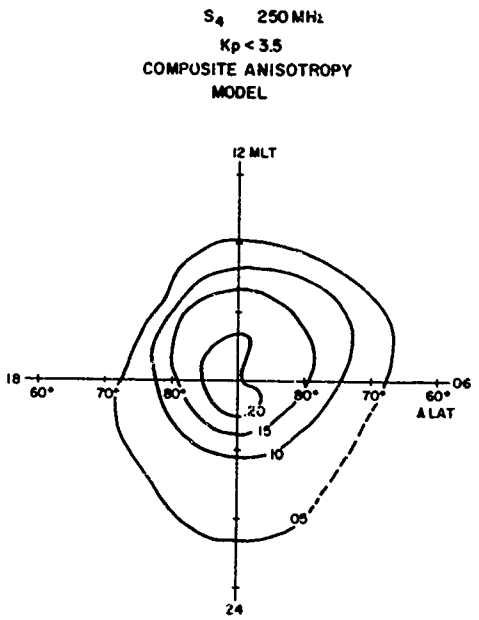


Figure 6a. Same as in Figure 5a except for 250 MHz.

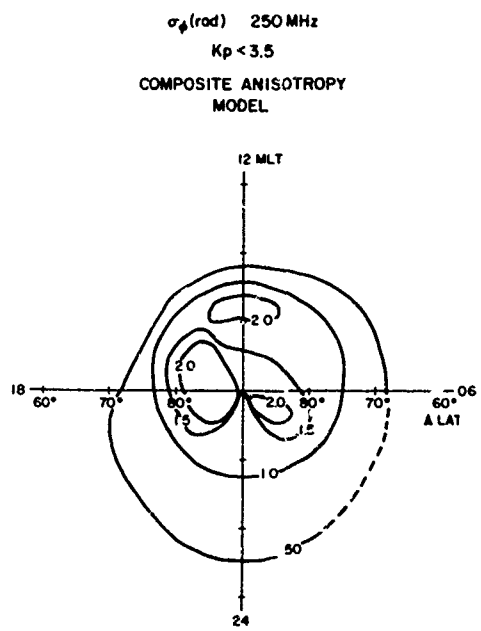


Figure 6b. Same as in Figure 5b except for 250 MHz.

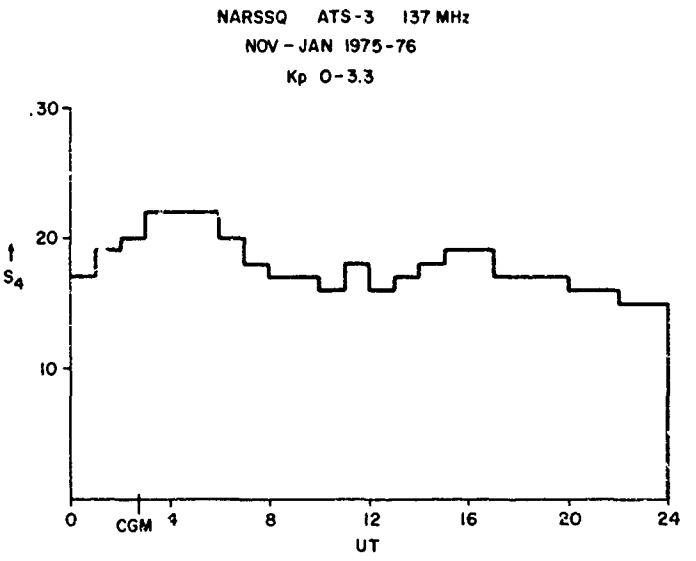


Figure 7a. Average value of intensity scintillation index (S_4) using ATS-3 at 137 MHz observed st Narssarssuaq for magnetically quiet conditions during Nov-Jan 1975-76.

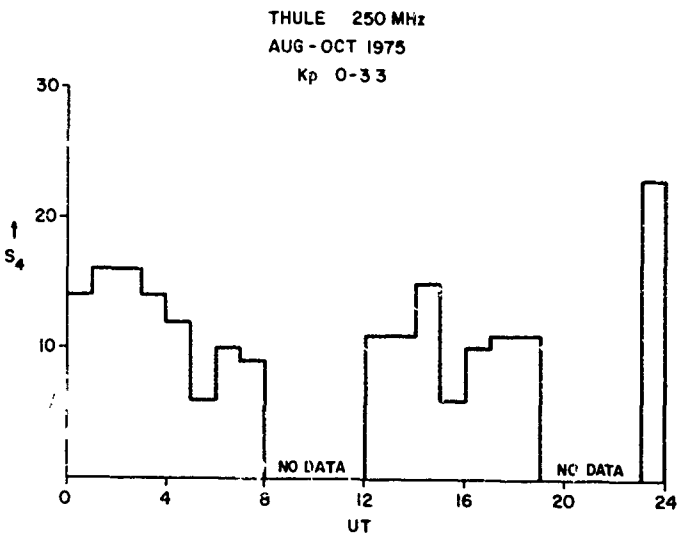


Figure 7b. Average value of intensity scintillation index (S_4) at 250 MHz observed at Thule for magnetically quiet conditions dur'ng Aug-Oct., 1975.

MF SKYWAVE PROPAGATION

Barbara G. Melander
Raymond H. West

Boeing Aerospace Company
P.O. Box 3999
Seattle, Washington 93124

INTRODUCTION

Low MF frequencies are receiving increased attention for possible survivable communications usage with both ground-to-ground and ground-to-air link geometries. Ionospheric reflections, or skywave signal components, are often ignored because of the small groundwave attenuation over short paths and the attenuation of the lower ionosphere during daytime and disturbed conditions. However, a potential exists for self interference between the groundwave and nighttime skywave at the larger distances of interest - particularly in mountainous terrain. Buried dipole antennas are utilized with short-path skywave modes efficiently excited relative to the groundwave. Also, long distance nighttime skywave interference or jamming must be considered.

The lower ionosphere is highly variable and not well modeled for predicting MF reflections. Predictions of skywave interference or jamming are correspondingly unreliable making estimates of interference conditions and effects on system performance unreliable. Estimates are most often made using procedures defined in CCIR document 264-3 which are based on best fits to limited empirical propagation data. These data are largely from Europe and North Africa and not sufficiently representative of conditions in the United States that high confidence in system performance can be attained (Wang, 1977). This paper describes an effort to relate skywave signal strength to lower ionosphere modeling rather than using propagation data only. Because of the variability of the ionosphere, its prediction must be described in statistical terms. The resulting signal is then specified as a mean estimate with confidence intervals assumed to be normally distributed.

A collection of lower ionosphere electron density profiles obtained from the World Data Center, Boulder, Colorado, has been used as the foundation of the present study. These data represent a range of locations, seasons, sunspot numbers and times of day and have previously been utilized for VLF/LF propagation studies (Berry and Davis, 1976). Regression analyses have been performed on subsets of these data to obtain models for electron density variability as a function of location and time. Reflection coefficients are then computed to obtain signal variability. Ionospheric modeling and predicted sensitivities are reported here with illustration of application in estimating field strength variability. The need for an expanded data base is demonstrated. Work on deriving signal strength prediction techniques of greatest use to system designers is continuing.

IONOSPHERIC MODEL

Data Base and Regression Analyses

The basic aim in the present development of a D-region ionospheric model is to obtain electron density and collision frequency profiles useable in calculating MF skywave propagation characteristics. Observed fluctuations in electron density due to measurement errors and real fluctuations in the ionosphere make a statistical approach to modeling very desireable. Using regression analysis techniques, a least squares fit to the electron density data obtained from the World Data Center has been done with standard deviations calculated about the mean profiles. The standard deviations of the electron density profiles allow one to calculate the expected standard deviations of MF skywave signal characteristics. The ultimate outcome of this D-region modeling will be mean MF skywave signal strength predictions as functions of local time, solar activity, month and location with the expected standard deviations of the signals also calculated.

Other statistical models include the CCIR MF signal strength predictions and an ionospheric model developed by Berry and Davis (1976). The CCIR predictions are yearly averages and do not have high enough time resolution to satisfy the present interests. The model by Berry and Davis was developed with regression analysis techniques from the same data set used on the present analysis, however, their interest was in the VLF/LF frequency range.

The data base used in this analysis was obtained from the World Data Center, Boulder, Colorado. The collection is electron density profiles derived from existing literature and digitized. The data is taken all over the world at many different values of sunspot number, latitudes, solar zenith angles, months and times of day. The methods used to derive the profiles are also varied from rocket measurements to profiles deduced from reflection measurements and thus vary in accuracy.

The main interest in this study is in U.S. nighttime and sunrise/sunset ionospheric profiles due to the enhanced reflections of MF from the ionosphere at these times. To focus on different local time regions of the ionosphere, the data were divided into three subsections, daytime, nighttime and transition regions. The divisions were made on the basis of solar zenith angle rather than local time in order to specify sunset and sunrise independent of latitude or month. The divisions are as follows:

Daytime	$0 < \cos x \leq 1.0$
Transition	$-0.5 < \cos x \leq 0$
Nighttime	$-1.0 < \cos x \leq -0.5$

where x = solar zenith angle. The divisions correspond to the transition region lasting from an hour or two before and after sunrise or sunset in the D-region at midlatitudes.

The D-region is not well understood and only fairly recently has much data been available or any modeling been done. The electron density in this region is generally believed to be a function of season, solar activity, solar zenith angle or local time and latitude. The specific dependences of the electron density on any of these parameters has not been well established.

Various height regimes in the D-region seem to respond differently to the above parameters (see Mitra and Somayajulu (1978); Bremer and Singer (1977); and Wratt (1977). To take these variations into account, separate regression analyses have been done for various height regimes. The D-region is divided into five distinct sections: (1) 50-65 km; (2) 65-72 km; (3) 72-85 km; (4) 85-95 km; and (5) > 95 km. The choice of these divisions was made on the basis of the average of observations in the literature and from inspection of the data obtained from the World Data Center. These choices are not necessarily the only available but seem to be a good average. For example, a ledge at 65 km has been observed by Deeks (1966), Bain and Harrison (1972), and Rowe (1970). This ledge has also been absent in some data as pointed out by Bremer and Singer (1977). Likewise structures have been seen a large percentage of the time at heights near 70-75 km, 80-90 km, and 90-100 km as noted by various authors; for example, Rowe (1972), Rowe et al (1974), Mitra and Somayajulu (1978), Bremer and Singer (1977) and Wratt (1977). By taking an average of these literature observations, the above height regimes were picked for the initial modeling.

The more general problem of allowing these height boundaries to be picked as part of the regression analysis makes the problem non-linear, thus harder to solve. The regression analyses performed here will be confined to linear functions. Therefore, these height divisions must be preselected.

Within each height division for each subsection (day, night, transition) of the main data set, regression analyses were performed to obtain mean and standard deviations of the electron density profiles. A linear form for the regression analyses was chosen as

$$\text{Log } N_e = A_0 + A_1 X_1 + A_2 X_2 + A_3 X_3 + A_4 X_4 + (B_0 + B_1 X_1 + B_2 X_2 + B_3 X_3 + B_4 X_4)\, h/100$$

where

N_e	=	electron density
A_{0-4}, B_{0-4}	=	regression constants
h/100	=	normalized altitude
X_{1-4}	=	regression functions dependent on month, solar zenith angle or local time, sunspot number, and latitude.

This form allows straight line fits in each altitude range with the intercept and slope of the line separately dependent on the month, local time, sunspot number and latitude. Different height regimes are allowed to respond independently to the various parameters as expected from observations.

To make N_e continuous at the boundaries of these layers, the following prescription is followed. If the density increases with increasing altitude across the layer, join points at ± 1% of layer thicknesses at each side of layer boundary. This forms a ramp not a sharp discontinuity. If the density decreases with increasing altitude across the boundary, use a weighted average of log N_e to find the new density at the boundary. This makes N_e continuous although the derivations remain discontinuous.

To choose appropriate functions X_1 - X_4 for the regression analyses, the literature was searched for expected variations with month, sunspot number, local time and latitude. No well established functional dependences exist at the present for any of these parameters.

The most consistent agreement between various data involves sunspot number. Generally electron density is enhanced throughout the D-region for enhanced solar activity (Bremer and Singer, 1977) although decreased electron density for altitudes below 65-70 km may also occur (Deeks, 1966; Metchly, et al, 1972). Following Berry and Davis (1976), X_3 is chosen as a linear function of sunspot number (X_3 = SSN).

The monthly variation of the electron density is less well established. Periodicity throughout the year seems to occur with summer having the largest electron densities and winter the lowest (Deeks, 1966; Rowe, 1972; Mechtly and Smith, 1968). This trend is not always consistent, however. Observations also show other times of the year to have maximum electron densities, e.g., spring at altitudes less than 69 km (Mechtly and Smith, 1968) and at 76.5 km (Wratt, 1977). The initial choice for seasonal variation was the same as Berry and Davis (1976), X_2 = cos (2π(m-0.5)/12) where m = month. This was varied and a sensitivity study was done on 85-95 km heights to see what differences existed in the regression analysis. Table 1 shows the various functions of month tried in analyses (1)-(5). The best fit for daytime and transition regions was the original choice mentioned above ((1) in Table 1). Nighttime had a higher correlation with analysis (2) where X_2 = cos (2π/m-2.5)/12), slightly shifted extremum from the day and transition choices.

TABLE 1

REGRESSION SENSITIVITY STUDY

$$\log N_e = A_0 + A_1 X_1 + A_2 X_2 + A_3 X_3 + A_4 X_4 + (h/100)(B_0 + B_1 X_1 + B_2 X_2 + B_3 X_3 + B_4 X_4)$$

Functions	Correlation Coefficients Night	Transition	Day		Correlation Coefficients Night	Transition	Day
(1) X_1 = cos (2π(m-.5)/12) X_2 = cos x X_3 = SSN X_4 = cos	.493	.729	.722	(6)X_2 = log (cos x)	.533	.718	.708
(2)*X_1 =cos(2π(m-2.5)/12	.559	.667	.722	(7)X_2= x	.524	.719	.715
(3)X_1=cos(2π(m-4.5)/12)	.529	.710	.726	(8)$X_2=\left\|\frac{L.T.-12}{12}\right\|\pi$	.705	.713	.679
(4)X_1=cos(2π(m-6.5)/12)	.534	.720	.718	(9)$X_2=\cos(\frac{L.T.-12}{12}\pi)$	.540	.705	.665
(5)X_1=cos(π(m-0.5)/12)	.516	.680	.727	(10)$X_4=\|\lambda\|$	.517	.719	.718

*Only parameters which are different than listed in (1) are listed in (2) -(10).
m=month, x=solar zenith angle. SSN=sunspot number, λ =latitude, L.T.=local time.

The variation expected in electron density with local time or solar zenith angle is also not well established. The ionosphere does change rapidly near sunset and sunrise, however. Electron density has been suggested to be proportional to $\cos^n x$ where x = solar zenith angle and n = 0-3 (Mitra and Somayajulu, 1978; Bremer and Singer, 1977). The value of n is found by these authors to vary for different values of x, e.g. $n = 2\text{-}3$ for $x < 40^0$, $n = 1$, $40^0 < x < 70^0$, $n \simeq 0$ for $x > 70^0$. Plots given by Bremer and Singer (1977) can be fit by $\log N_e \sim \cos x$. This is similar to results given by Rowe (1972). The initial choice of local time or solar zenith angle variation was $X_2 = \cos x$. Other choices were tested as shown in Table 1, anlayses (6)-(9). The daytime and transition regions agree with the original choice. Nighttime shows a much better correlation with local time. Berry and Davis (1976) found the entire data set to be better correlated to local time, not completely consistent with our findings since nighttime data consist of only 22 out of 551 total profiles.

Latitude variations are not well defined in the literature. Mitra and Somayajulu (1978) suggested no latitudinal effect on electron density (except that which enters through solar zenith angle dependence). Mechtly, et al (1969) found no systematic variation with latitude in a study done with constant solar zenith angle. A latitude dependence was included in our analysis to see if regression analyses could indicate some trend. The form used was $X_4 = \cos \lambda$ where λ = latitude. The form $X_4 = |\lambda|$ was also tested (shown on Table 1 as analysis 10). Only for the nighttime data section are results significantly different than with the cosine function.

The final choices for the regression analyse at all heights were selected by the above search to be

day and transition regions	X_1	$= \cos (2\pi (m-0.5)/12)$
	X_2	$= \cos x$
	X_3	$=$ SSN/100.
	X_4	$= \cos \lambda$
night region	X_1	$= \cos (2\pi (m-2.5)/12)$
	X_2	$= \left\| \frac{l.t.-12}{12} \right\| \pi$
	X_3	$=$ SSN/100.
	X_4	$= \|\lambda\|$

where

m = month

l.t. = local time

SSN = sunspot number

λ = latitude

Using the above functions, regression analyses in each height division were performed. The regression constants, correlation coefficients and standard deviations are listed in Table 2. For height regimes 50-65 km and 65-70 km not enough data for nighttime anlayses was available so no results are listed. The correlation coefficients for some regions is low (below 0.70) and standard deviations high. These facts point to the need for more data (especialy nighttime), more care in picking accurate profiles for the analyses, and more though to other possible functional dependences at various heights.

Results of Modeling

Plots of electron density profiles as predicted by the regression analyses as functions of various parameters are shown in Figures 1-5. Figure 1 shows the comparison between prediction and some of the data used in the analyses. Also shown is the Berry and Davis (1976) fit to the same profiles. These profiles were not chosen as best examples of regression fit to the data. They are only as examples of several curves spanning most of the 50-120 km height range for a given parameter set. Not all the density profiles used in the analyses span the entire altitude range under consideration. Our data fit is good overall and certainly better than Berry and Davis for altitudes over ~ 80 km. The MF frequency range reflects at altitudes near 80 km or higher at night. Berry and Davis work was designed for VLF/LF work which needs better fit at slightly lower altitudes.

TABLE 2

Height Divisions (km)	DAYTIME Based on 446 Curves		NIGHTTIME Based on 22 Curves		TRANSITION Based on 82 Curves	
	Regression Constants	Correlation Coefficients & Standard Deviations	Regression Constants	Correlation Coefficients & Standard Deviations	Regression Constants	Correlation Coefficients & Standard Deviations
50-65	A0 - 0.47817 A1 2.2919 A2 8.7048 A3 - 0.20960 A4 - 9.6620 B0 4.0904 B1 - 3.6529 B2 - 13.049 B3 0.76746 B4 13.964	.621 (±.336)			- 2.4726 - 0.23257 - 12.170 - 8.4634 - 1.1316 10.195 0.57513 28.316 11.946 - 2.1997	.846 (±.291)
65-72	A0 3.4657 A1 - 0.11594 A2 - 2.8506 A3 1.9768 A4 - 4.4486 B0 - 2.3118 B1 0.055597 B2 5.2055 B3 - 2.6089 B4 5.8978	.585 (±.309)			1.8513 - 1.7683 40.868 - 3.2253 4.0103 2.4643 2.7682 - 53.254 4.3756 - 3.9064	.664 (±.384)
72-85	A0 - 2.1084 A1 - 0.77119 A2 2.4518 A3 0.14426 A4 - 1.9462 B0 5.3415 B1 1.0284 B2 - 1.9407 B3 0.095515 B4 2.2956	.676 (±.348)	- 12.881 3.3963 0.0070438 0.77948 0.18231 15.175 - 4.6542 - 0.0064341 - 0.86122 - 0.16740	.861 (±.683)	4.7290 - 2.9622 8.8289 - 5.0990 - 6.2435 - 1.4759 4.4506 - 10.654 6.6591 4.8927	.732 (±.584)
85-95	A0 - 5.4507 A1 - 0.20919 A2 - 1.5212 A3 - 2.7086 A4 - 0.72156 B0 9.4635 B1 0.32329 B2 2.6353 B3 3.2564 B4 0.82581	.722 (±.424)	- 20.689 - 6.6021 0.010618 - 4.0742 0.30780 24.824 6.9831 - 0.0099482 4.4781 - 0.31987	.732 (±.677)	1.2189 7.3379 - 10.440 2.7786 - 20.450 2.3185 - 7.8363 12.426 - 3.0526 22.346	.729 (±.492)
95-120	A0 - 2.1917 A1 0.58171 A2 3.1335 A3 1.8609 A4 1.9848 B0 6.1401 B1 - 0.49200 B2 - 1.9504 B3 - 1.6498 B4 - 2.1185	.783 (±.291)	2.7809 3.1795 0.010217 - 7.9359 0.027576 0.087992 - 3.3887 - 0.0089133 8.2692 - 0.021516	.710 (±.493)	0.34097 - 1.0102 - 3.4120 2.9344 0.12224 3.8938 1.1213 4.9557 - 3.0410 - 0.31826	.607 (±.403)

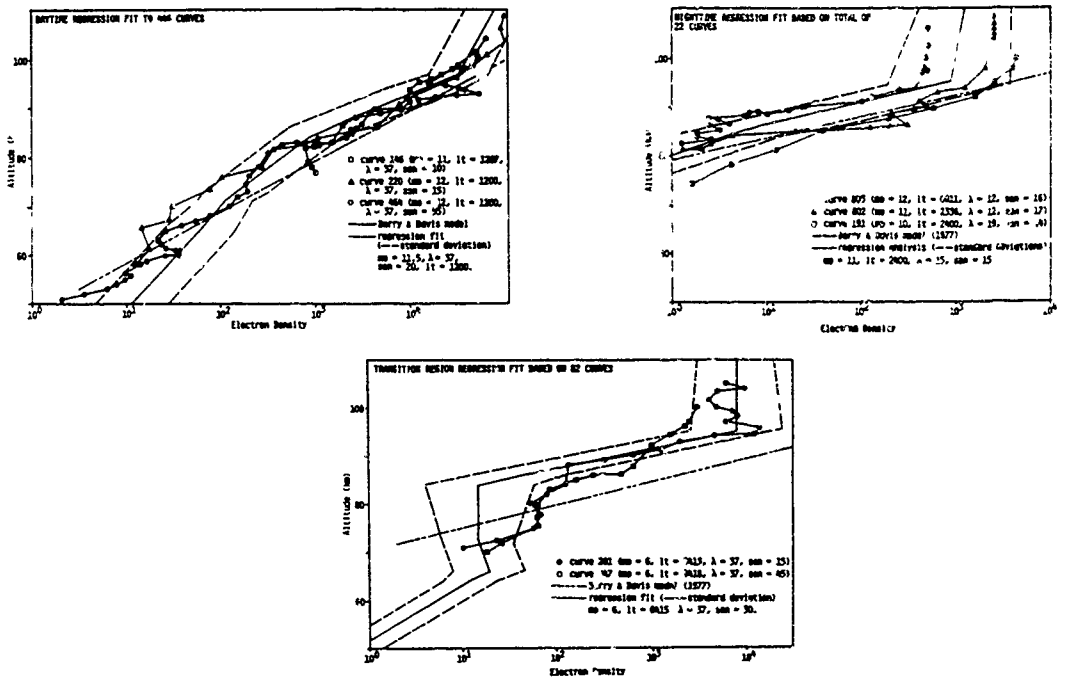


Figure 1. Comparison of regression analysis fits to data.

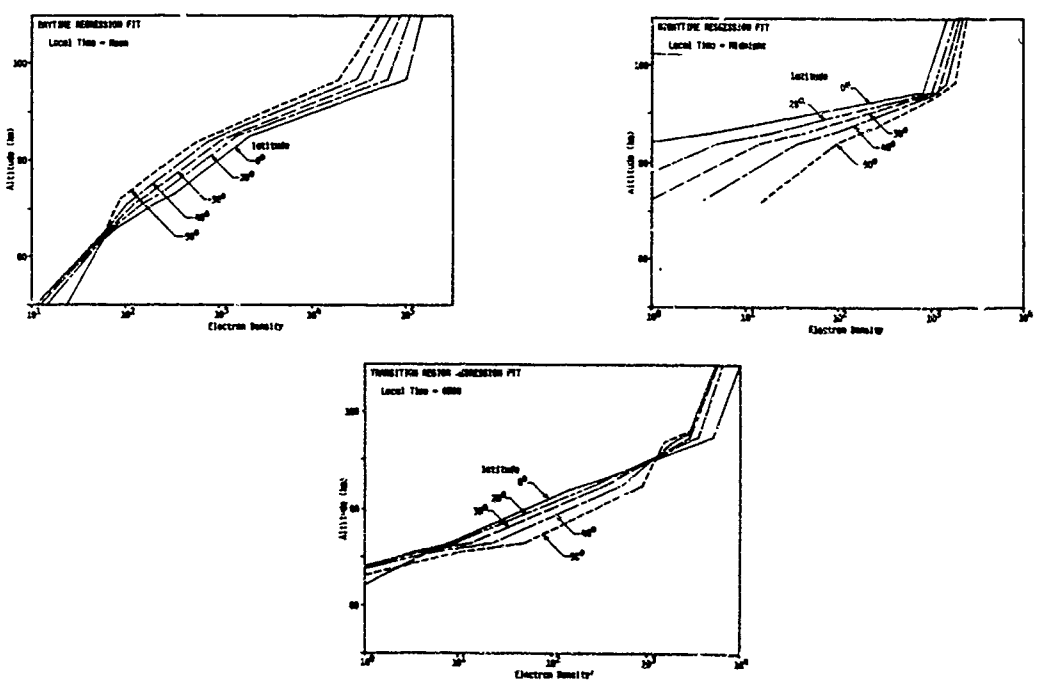


Figure 2. Model prediction of latitude dependence for June and sunspot number = 25.

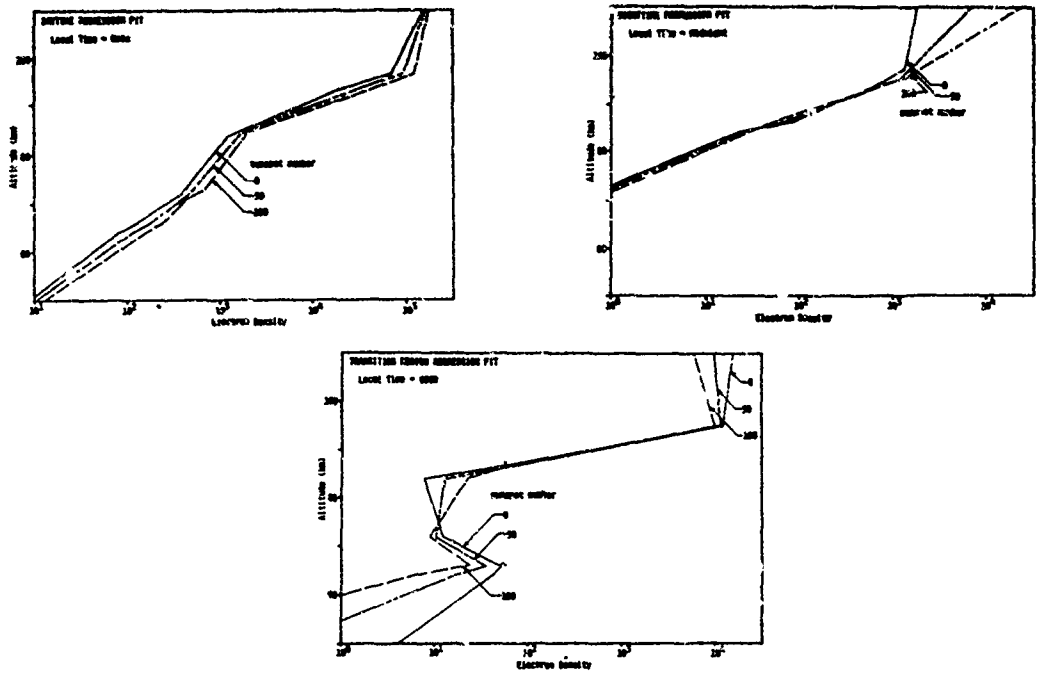


Figure 3. Model prediction of sunspot number dependence for June and 30° latitude.

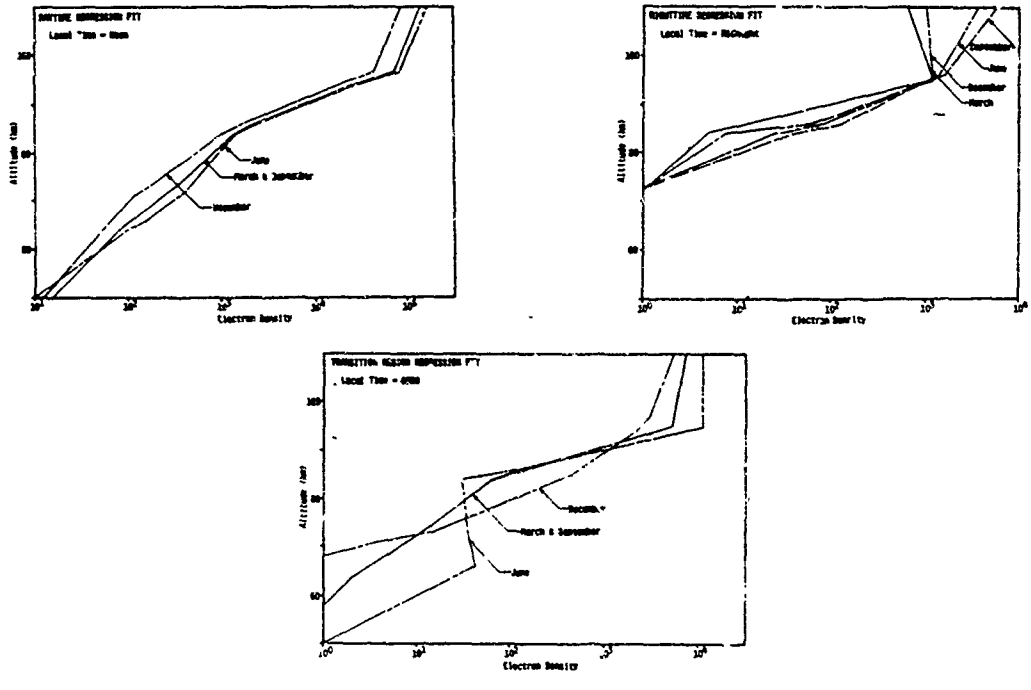


Figure 4. Model prediction of monthly dependence for 30° latitude and sunspot number = 25

Figure 2 shows the model prediction of election density profile variation with latitude for constant local time. Daytime profiles are seen to steadily increase in electron density for all altitudes as latitude increases. Nighttime shows opposite behavior, the density increases with increasing latitude. Transition profiles have mixed behavior with high altitudes exhibiting similar characteristics to daytime and low altitudes similar to nighttime predictions. No data studies have been found to predict anything about latitude dependence.

Figure 3 shows the predicted sunspot number dependence for the three data sets. Daytime profiles are enhanced by increasing sunspot number. Nighttime shows little variation except above 95 km where the density increases with increasing sunspot number. The transition profile shows opposite behavior. The density decreases with increasing sunspot number for altitudes below 75 km and above 95 km. Both Deeks (1966) and Metchly, et al (1972) see in daytime studies enhanced electron density with increasing solar activity at higher altitudes (not consistent with our model). The crossover point differs in each study. Deek's crossover pont is near 82 km while Metchly's is 62 km. Bremer and Singer (1977) see density enhancement for increased solar activity over the entire height range (although they state some uncertainty in their data below 65 km).

Monthly dependence of the density profile is shown in Figure 4. Daytime profiles show enhanced density in summer and decreased density in winter over the entire height range. The equinoxes are identical and at intermediate values to winter and summer. Nighttime profiles show spring as a minimum, fall a maximum with summer smaller than fall but larger than winter for all heights. The transition region varies with summer profiles enhanced above 90 km and below 95 km and smallest in between. Experimental results are varied. Deeks (1966) shows daytime profiles consistently lower than the equinoxes which are lower than summer for all altitudes, consistent with our daytime results. Bremer and Singer (1977) shows summer greater than winter which is greater than spring for altitudes below ~ 85 km. Above 85 km the winter profile is greatest. Rowe (1972) shows results similar to Deeks at all altitudes although he displays one winter profile which is enhanced over summer above 70 km, similar to Bremer and Singer. Mecntly and Smith (1968) show spring and summer with the most enhanced profiles below 85 km and no systematic behavior above 85 km.

Figure 5 shows the local time variation of the profiles. Daytime regression plots are used for local times of 6-12, transition regression fit is used for a local time = 4; and nighttime fits for 0 and 2. An overall decrease in density is seen as the local time decreases at all altitudes with midnight being the lowest profile. Faster local time variation is seen at lower altitudes before sunrise (local time < 6). These results are similar to those seen by Rowe (1972) and Deeks (1966).

Collision Frequency

The collision frequency profile used in the calculation of reflection coefficients is a fit to the profiles given by Bremer and Singer (1977) in their figure 2. The collision frequency is higher in summer below 84 km but lower than winter and spring above 84 km. Below 80 km these profiles are similar to those given by Rowe (1972), Deeks (1966) and Wait and Spies (1964). This last profile was used by Berry and Davis in their ULV/LF work.

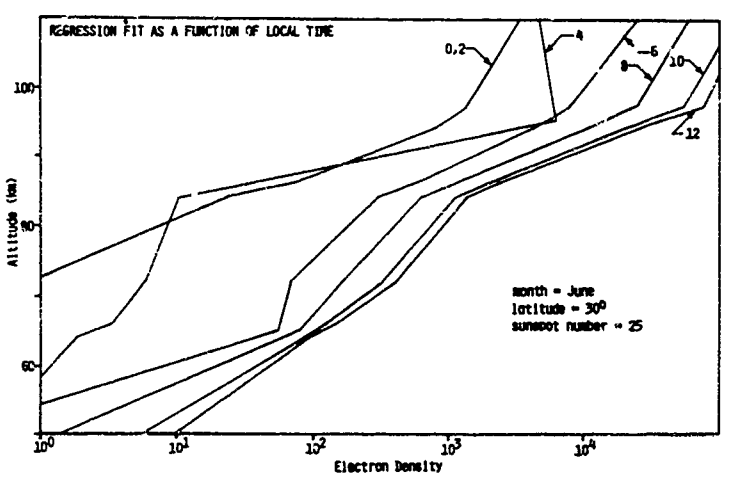


Figure 5. Model prediction of local time dependence.

PROPAGATION PREDICTION

Estimation of the ionospheric contribution to signal strength losses requires selection of a method to calculate MF propagation through the model ionosphere. MF frequencies are particularly difficult to deal with due to the wavelengths involved (close to scale lengths in the ionosphere) and to the fact that the frequencies involved are close to collision and cyclotron frequencies. These effects can lead to large absorption, scattering or coupling of ionospheric propagation modes. Because of the large absorption, ray-trace methods are not suitable except when reflection is expected to be significant, e.g., at night. Methods using multiple reflecting layers are useable but require thin layers due to the small wavelengths. Ideally a full wave analysis would be used but has not yet been implemented.

Initially a simple reflection code was developed to handle the case of east-west propagation for vertical polarization. This allows us to test the sensitivity of the reflection coefficients to changes in the election density profiles. A sample calculation is shown in Figure 6 with the reflection coefficient in dB plotted against the cosine of the incident angle relative to the ionosphere. The curves are parametric in variation of the electron density profile as defined by log N_e (Δ in figure 6). Typical standard deviations of the profiles are found to fall within the range $\Delta = \pm 0.4 - \pm .6$ when expressed in this manner. Although it is possible, and intended, to express reflection coefficients in terms of mean and standard deviations, this has not been implemented at this point. The curves illustrate the type of variability to be expected, however. The shapes are similar to reflections from a dielectric earth. These curves show considerable sensitivity to shifts in the electron density profile, especially near the curve minima. Using a mean and standard deviation of an electron density profile, a range of reflection coefficient values can be specified including confidence limits as functions of location, time of day, solar activity and frequency.

Eventual incorporation of the results of this ionospheric modeling into a more comprehensive system design and performance predictive capability requires that other system components be described in a compatible form. Principal among these, are antenna performance and atmospheric noise. Buried MF dipole antenna performance is dependent on local ground conductivity which can be shown to be reasonably represented by a normal distribution expressing siting and weather variability about the mean of a particular locale. A similar statement can be made for uncertainties in groundwave signal strength resulting from the inability to accurately specify path variability (of interest when treating ground-wave-skywave interference). Atmospheric noise is expressed as a noise factor with uncertainties of prediction described (generally) by independent normal distributions. Our system model will, then, include estimates of each system parameter expressed as mean values which are dependent on specification of location, time, and frequency and also estimations of the errors of prediction as functions of these same parameters. Independence of the statistical variables is assumed. Confidence in predictions are available as a result of this approach on which to assign performance margins.

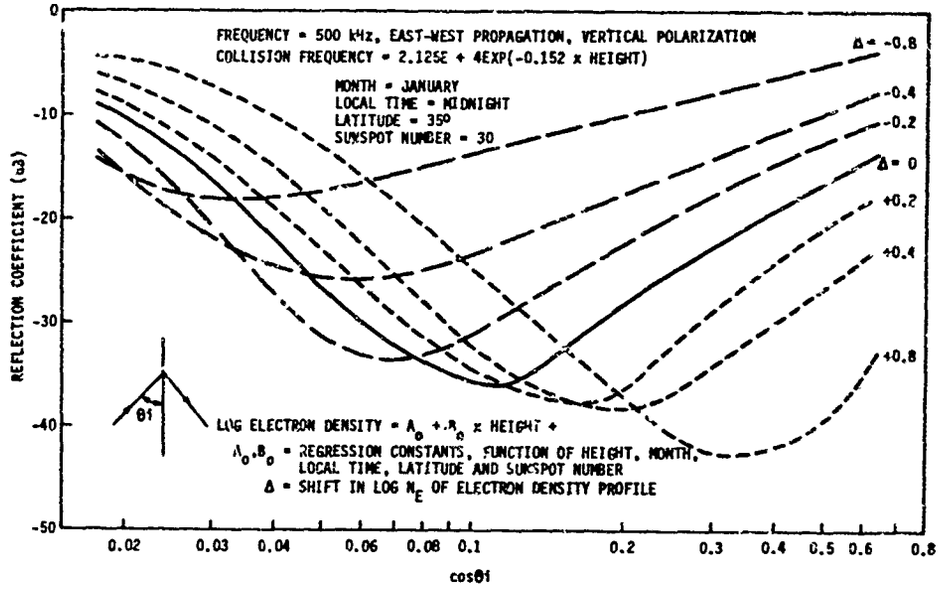


Figure 6. Ionospheric reflection coefficient for mean electron density profile & variation about mean.

CONCLUSIONS

Regression analysis of lower ionosphere electron density profiles has been shown to be a reasonable approach to developing MF signal estimates in a form compatible with system prediction requirements. The resulting model and hence skywave signal strength estimates are critically dependent on the detail of available ionospheric data. This has been demonstrated to be somewhat lacking - particularly for the times of greatest interest (night, sunrise and sunset) so that expansion of the data base is necessary for improved confidence. Within the limitations of the data utilized, a multilayered model of the D and lower E regions has been developed which allows estimates of variability as means and standard deviations. Illustrations of the various sensitivities of the ionosphere and comparisons with other data in the literature have been presented. Application of the ionospheric model has been illustrated by computing reflection coefficients which can be interpreted to signal strength variability. Additional model development in this area is required.

The conclusion is reached that the approach followed will allow estimation of skywave signal strengths, their variability and the errors of prediction in a form useful to the system designer. This approach will allow insight into physical processes more completely than the commonly used CCIR approach. Limitations on available data leave the prediction open to question but some light will hopefully be shed with continuing effort.

REFERENCES

Bain, W. C.; Harrison, M.D.; Proc. Inst. Elect. Engr.; 119; 790, 1972

Barry, L.A.; Davis, R.M.; "A Statistical Model of the Lower Ionosphere;" Defense Communications Agency No. C600-TP-76-2; 1976.

Bremer; J.; Singer, W.; "Diurnal, Seasonal and Solar-Cycle Variations of Electron Densities in the Ionosphere D- and E-Regions;" J. of Atm. Terr. Phys.; 39; 25; 1977.

CCIR Report 264-3; "Skywave Propagation Curves Between 300 km and 3500 km at Frequencies Between 150 kHz and 1600 kHz in the European Broadcasting Area."

Deeks, K. G.; "D-Region Electron Distributions in Middle Latitudes Deduced from the Reflexion of Long Radio Waves;" Proc. Royal Soc.; A291; 413; 1965.

Metchley, E. A.; Smith, L.G.; "Seasonal Variations of the Lower Ionosphere at Wallops Island during the IQSY;" J. of Atm. Terr. Phys., 30; 1555; 1968.

Metchley, E. A.; Rao, M. M.; Shaperdas, D. O.; Smith, L. G.; "Latitude Variation of the Lower Ionosphere;" Radio Sci.; 4; 517; 1969.

Metchley, E. A.; Bowhill, S. A.; Smith, L. G.; "Changes of Lower Ionosphere Electron Concentrations with Solar Activity;" J. of Atm. Terr. Phys.; 34; 1899; 1972.

Mitra, A. P.; Somayajulu, V. V.; "A Global Model of D-Region Ionization;" Cospar Space Research Plenary Meeting 21; 269; 1978.

Rowe, J. N.; "Model Studies of the Lower Ionosphere;" NTIS; AD-755-447; 1972.

Rowe, J. N.; Mitra, A. P.; Ferraro, A. J.; Lee, H. S.; "An Experimental and Theoretical Study of the D-Region - II. A Semi-empirical Model for the Mid-Latitude D-Region.;" J. of Atm. Terr. Phys.; 36; 755; 1974.

Wait, J. R.; Spies, K. P.; "Characteristics of the Earth-Ionosphere Waveguide for VLF Radio Waves;" NBS Technical Note 300.

Wang, J.C.H.; "Prediction of Medium-Frequency Skywave Field Strength in North America;" IEEE Trans., Vol. BC-23, No. 2; 1977.

Wrott, D. S.; "Variations on Electron Density in the Middle Lattitude D-Region Above Urbana, Illinois;" J. of Atm. Terr. Phys.; 39, 607; 1977.

LATERAL BENDING EFFECTS AT THE IONOSPHERIC HEIGHT TRANSITION

David Mannheimer
Norden Systems, Inc.
Melville, New York 11747

I. BACKGROUND

Extensive usage of the Omega long range radio location system on commercial air routes has revealed a type of anomalous position error which is systematic in nature, tending to recur on successive dates at the same time and position, when both aircraft receiver and Omega transmitter are in twilight. This paper surveys the relevant VLF propagation aspects, and offers an explanation for its cause. It also points out key co-terminator geometric relations useful for extraction of ionospheric height data from phase recordings over known paths.

Significant phase errors at VLF frequencies can be expected due to lateral bending over long paths nearly parallel to the terminator because the corresponding ionospheric height transition separates regions of measurably different propagation velocity.

At 10.2 kHz, for example, if day and night phase velocities differ by 0.4%, a critical refraction angle of 5.1° is predicted for waves passing through the twilight velocity gradient Consequently, over a long path with one end in daylight and the other in darkness, but with terminator/path angle less than 5°, critical bend angles will occur, irrespective of the detailed profile of the transition gradient both for the dominant and possible higher order modes. Such "straddle" geometric situations have been identified in a number of flight tests as coincident with large position errors, removable only by deselection of the causal terminator station. This type of geometry has frequent incidence somewhere along commercial aircraft route/schedules varying with time of year, but may seldom if ever be apparent from a fixed location laboratory. Yet the error effect is a minute-by-minute phenomenon, difficult to isolate, except by use of an absolutely stable (rubidium) phase reference at a fixed location, and when individual measurement can be made on a single transmitter. Also in worst case operational situations involving only a few used stations with poor bearing geometry, position error can greatly exceed a single causal radio error, due to reference clock shift in an airborne receiver.

The general effect of bending of the wave normal is to cause a modest increase in phase delay compared to that which would be expected over a direct great circle path. More importantly, it causes a significant lateral displacement of the actual propagation path away from the direct great circle. This causes a change in timing (GMT) of the rapid diurnal phase shift expected over long grazing paths leading to transient carrier phase errors, and results in large dispersion errors in multi-frequency tracking system like Omega.

II. BENDING ERROR PREDICTION FROM A SIMPLE IONOSPHERIC MODEL

An estimate of phase error due to bending at small terminator angles can be determined by considering the terminator as an extensive step interface between day (morning) and night ionospheric wave guides. Figure 1 shows such an elementary model, and defines the coordinate system and terminology used to consider lateral bending using the flat earth and plane wave approximations. This model is only appropriate when transmitter and receiver locations straddle the ionospheric height gradient over long paths, but is useful to predict upper error bounds when critical bending occurs, of prime interest herein.

Bend angle (δ) for any propagation mode can be expressed as a function of angle of incidence (Ψ) by the use of Fermat's principle which reduces to Snell's law.

$$\tan \delta = \tan (\Psi' - \Psi) = \frac{\tan \Psi \left[\sqrt{\frac{\mu^2 - 1}{\sin^2 \Psi} + 1} - 1 \right]}{1 + \tan^2 \Psi \sqrt{\frac{\mu^2 - 1}{\sin^2 \Psi} + 1}} \qquad (1)$$

This formulation is identical to that used by Wait[1] if due allowance is made for differences in format and terminology.

At VLF frequencies $\mu^2-1 \approx 2\left(\frac{v_d-v_n}{v_n}\right) = \frac{2\Delta v_{d/n}}{v_n}$

Thus at large angle of incidence, which will occur at large terminator angles, bending is nil.

$$\lim_{\Psi >> \Psi_c} \tan \delta \approx \frac{\mu-1}{\tan \Psi} \quad \begin{matrix} \to 0 \\ \Psi \to 90^\circ \end{matrix}$$

This is equivalent to Wait's result[1].

However, for cases of interest here, where terminator angles, and consequently angles of incidence are small

$$\lim_{\Psi << \Psi_c} \tan \delta \approx \sqrt{\mu^2-1} \approx \Psi_c$$

Thus, at grazing terminator angles, the bend angle approaches the critical angle for the frequency and propagation mode involved. This critical behavior predicts a surface wave along the terminator on the day side. This markedly changes the day/night path length ratio as compared to the ratio computed conventionally along the direct great circle path – the greatest path shift occurring for example at transition onset at sunrise, tapering to nil at the end of the twilight transition. This is preceded by growing reflective/multipath errors maximizing at onset, and possibly total reflection. Figure 2 shows typical error trends.

A quadratic solution in $\tan \Psi$ can be written from figure 1, as a function of terminator angle (α), and relative refractive index (μ) in terms of the parameter L_n/L_d whose range is zero to infinity during terminator transit along the radio path.

$$[1-(\mu L_n/L_d)^2] \tan^2 \Psi - 2 \tan \alpha (1+L_n/L_d) \tan \Psi - (\mu^2-1) (L_n/L_d)^2 + (1+L_n/L_d)^2 \tan^2 \alpha = 0 \qquad (2)$$

At nightside end of path refraction is nil.

$$\lim_{L_n/L_d \to 0} \Psi = \alpha$$

A dayside end of path critical refraction occurs for small terminator angles, involving a finite L_n/L_d ratio

$$\lim_{\substack{\Psi \to 0 \\ \Psi' \to \Psi_c}} L_n/L_d \approx \frac{\tan \alpha}{\tan \Psi_c - \tan \alpha} \qquad (\Psi < \Psi_c)$$

This enables prediction of the maximum possible lateral path shift and the critical refraction error (Δt_c)

$$\Delta t_c = \frac{L \sin \alpha}{900} \left(1 - \frac{\tan \alpha}{\tan \Psi_c}\right) \text{ hrs} \quad \left(\Delta T_{12} = \frac{L \sin \alpha}{900}\right) \qquad (3)$$

where Δt_c is timing error, and ΔT_{12} is time differential between sunrise (or sunset) at transmitter and receiver. The peak refractive phase delay error in fractional wavelengths then becomes (for a constant α)

$$\Delta \phi_c = 100\, L/\lambda \, [\cos (\Psi_c - \alpha) - 1] \qquad (4)$$

where L/λ is number of path wavelengths.

[This error prediction, for example, is of the order of a half wavelength at 10.2 kHz over a 6000-mile path at a terminator angle of 2°, while Δt_c is about 15 minutes.]

The above expressions are derived using the daylight velocity for the critical leg for the dominant mode. The slower surface wave or, on the other hand, greater bending at higher order modes will modify this simple estimate. Detailed reduction shows a steady reduction in error during transition. In the following, only this peak error is discussed, because it provides a good measure of significance, and greatly simplifies calculations. This occurs at sunrise onset (1) assuming constant terminator angle. When the actual variability of terminator angle is considered, it will occur at sunrise onset if one end of path is on the equator, but will occur in mid-transition for trans-equatorial paths. [Then it is identified as $\Delta \phi_p$.]

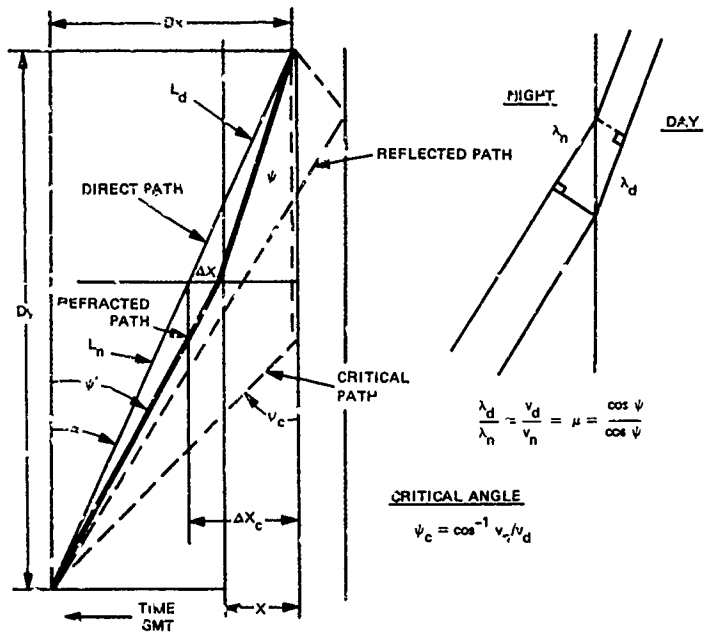


Figure 1. Refraction Geometry

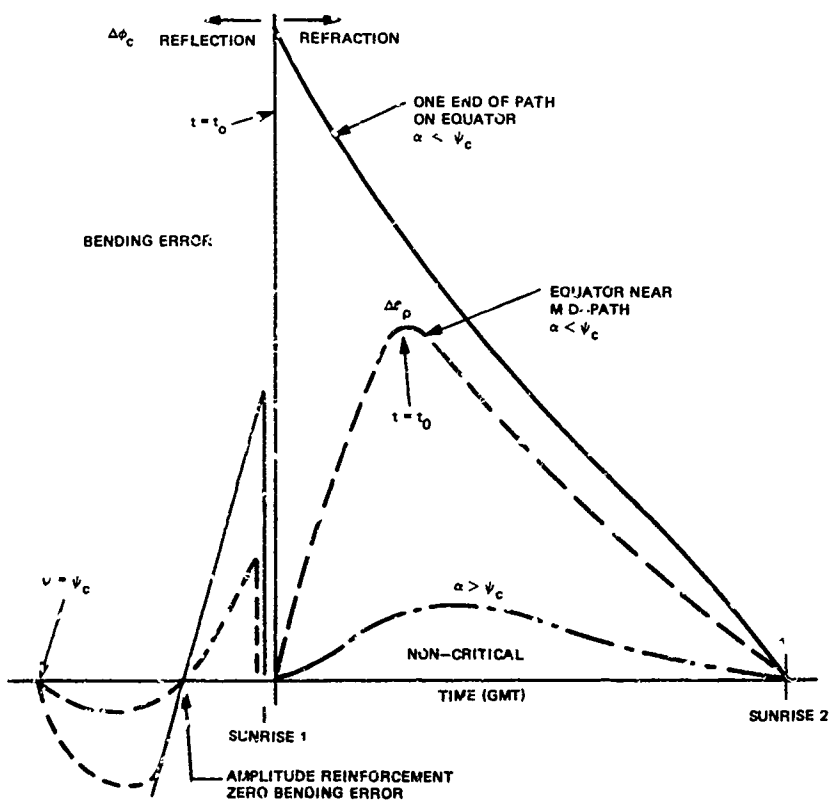


Figure 2. Typical Bending Error Trends

The actual terminator angle during transition is variable, but of course can be accurately predicted as a function of time between known locations by taking accurate account of sun location. The intersection velocity of the terminator along the radio path and consequently the diurnal phase rate becomes very high over long paths at small terminator angles. It is consequently necessary to use precise geometric and time normalization if attempt is made to identify bending effects on actual phase recordings. In fact, co-terminator geometry and timing alone determines several important characteristics of bending error. Accordingly the relevant navigational relationships are summarized in the following section.

III. CO-TERMINATOR GEOMETRY

The direct radio path great circle can be defined in terms of its equatorial longitude ($Lo_p + 90°$) and its angle from the meridian at the equator (L_p), here termed path declination. Also Lo_p, L_p are coordinates of the axis or pole of the path great circle as illustrated in figure 3. These are calculable for any path.

$$L_p = \sin^{-1} [\sin \phi, \cos L_1] \tag{5}$$

$$Lo_p = Lo_1 + 90° (270°) + \cos \phi_1 / \cos L_p \tag{6}$$

where L_1, Lo_1 and ϕ_1 are location and path bearing at any point along the path great circle, usually taken at the receiver.

The terminator great circle can be similarly defined in terms of sun-noon position [sun latitude or declination (D) and sun longitude (Lo_s)], corresponding with the axis of the terminator great circle.

Sun declination and consequently terminator bearing at the equator varies ±23.4° periodic in one year. Terminator equatorial longitude ($Lo_s + 90°$) increases at 15°/hr*, phased from 0600 GMT-sunrise at the Greenwich meridian at the equator.

IV. TERMINATOR ANGLE TIME FUNCTION

Terminator angle then can be expressed as a time function depending only on path and sun declination (date) for a given path. It is also the great circle distance between sun and path pole position evident from the following expression.

$$\cos \alpha(t) = \sin D \sin L_p + \cos D \cos L_p \cos \Delta Lo \tag{7}$$

where. $\Delta Lo = 15°/hr\ (t - t_0)$

$t_0 = 0600\ (1800) + Lo_p/15$

Time differentiation will reveal that minimum terminator angle occurs when the terminator intersects the direct path great circle at the equator, that is, at a fixed time of day for any fixed laboratory – station path, since Lo_p is the only variable in the above result for t_0.

Further, minimum terminator angle occurring at this time depends only on the difference between sun and path declinations.

$$\alpha_{min} = D - L_p \tag{8}$$

V. PATH DISTANCE FUNCTION

The angular distance between the terminator path intersection and the equator along the station path can also be expressed as a time function.

$$\sin B(t) = \frac{\cos D \sin \Delta Lo}{\sin \alpha} \tag{9}$$

Differentiation will yield terminator velocity along the station path – which determines diurnal phase rate.

The above co-terminator time functions have been illustrated in figure 4 for a typical case involving 3° minimum terminator angle for a path between RAE Farnborough and Omega station 6 in Argentina.

---

*The small variations in sun longitude rate expressed by the Equation of Time are not considered herein.

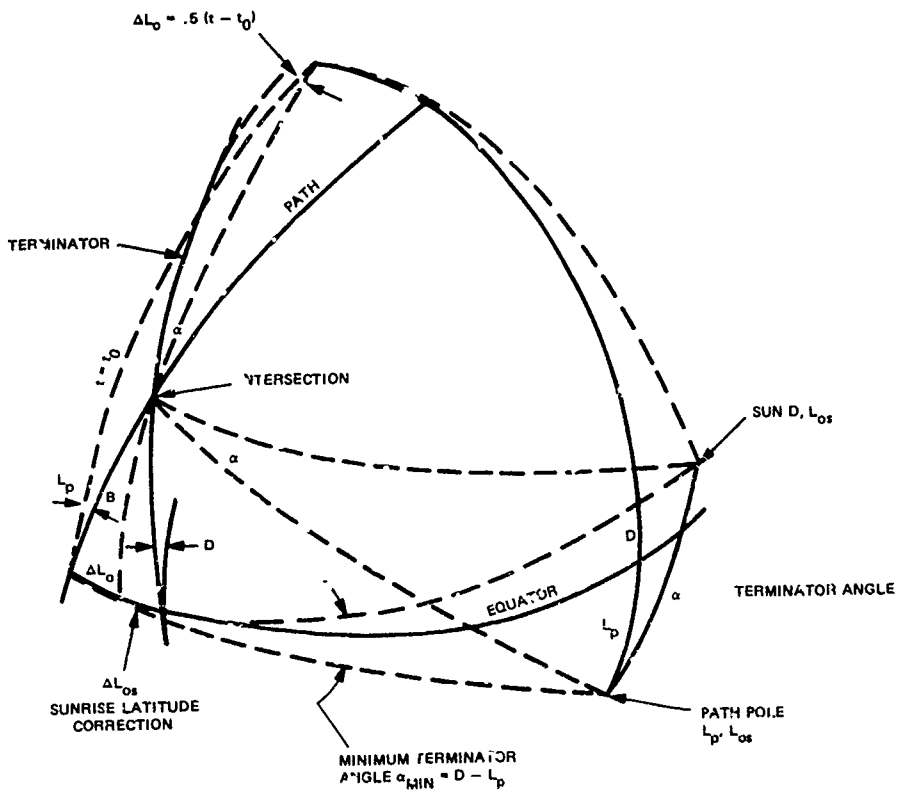


Figure 3. Co-Terminator Geometry

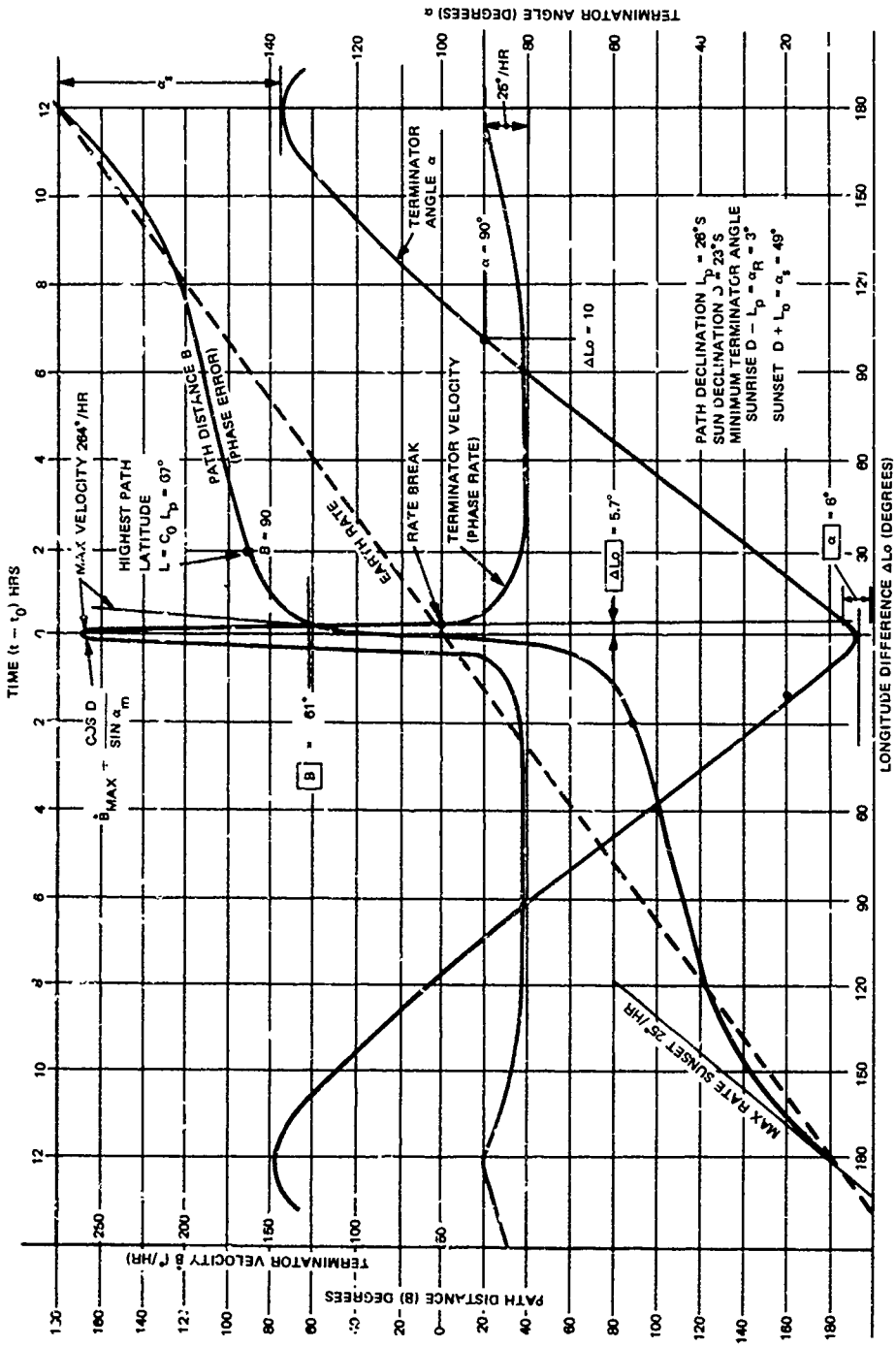


Figure 4. Co-Terminator Time Functions

For the cases of interest during fast transitions, along paths including or nearly including the equator where terminator angle is small, terminator angle and path distance functions can be accurately approximated in simpler fashion [angles are expressed in degrees].

$$\alpha \approx \alpha_m \sqrt{1 + \left(\frac{\Delta Lo}{\alpha_m}\right)^2} \tag{10}$$

$$\sin B \approx \frac{\cos D}{\sqrt{1 + \left(\frac{\alpha_m}{\Delta Lo}\right)^2}} \tag{11}$$

This shows the simple generic nature of the terminator angle time function in the situations of interest, and more readily enables simple determination of the maximum and average intersection rates.

$$dB/d\Delta Lo \text{ max} = \frac{\cos D}{\sin (D - L_p)} \qquad \Delta Lo \to 0 \tag{12}$$

VI. PHASE RATE FUNCTION

This enables an expression of maximum diurnal phase and phase rate functions in terms of day/night phase velocity fraction $\left(\frac{\Delta v_{d/n}}{v_n}\right)$ and wavelength (λ) yielding the maximum phase rate result.

$$\dot{\Phi}_{max} = \frac{90000 \cos D}{\sin (D - L_p)} \frac{\Delta v_{d/n}}{v_n \lambda} \text{ (cec/hr)} \tag{13}$$

From this result, assuming propagation travels along a fixed direct path crossing the equator, maximum phase rate occurs at the same time throughout the year, while its magnitude depends only on day and night phase velocities and date. The diurnal phase shift for small time differences from t_0, can be approximated similarly.

$$\Phi \approx \frac{90000 \cos D \, \Delta v_{d/n} (t - t_0)}{\sin \alpha \, v_n \lambda} \text{ (cec)} \tag{14}$$

An approximation for peak error at the equator is also expressed

$$\Delta\phi_p \approx \dot{\Phi}_{max} \Delta t_c \tag{15}$$

These simple but accurate geometric/time predictions can be directly correlated with a Nautical Almanac and leave only the day/morning phase velocity fraction as uncertain. This then becomes measurable from phase recordings over known geometry paths. It corresponds with the terminal phase velocity ratio at night and morning ends of the path, or morning/night height ratio

Figure 5* shows phase recordings for successive dates for a fixed path, overlaid on the same time-of-day scale. It illustrates well the above predicted maximum phase rates both as regards magnitude and timing. The scatter in timing of maximum phase rate apparent in the recordings then is presumed to be due to bending effects due to lateral path shifts, as compared with stationary and random phase anomalies.

It is apparent from the foregoing that error maxima may occur at the time of minimum terminator angle or at the time of dayside onset – not far separated in time during fast transitions.

Figure 6 compares an actual phase recording for the Argentina – Farnborough path on January 3, with direct and bent path predictions, using the sun elevation zero definition. It illustrates that the simple step ionospheric model is quite adequate to predict the basic error magnitude and timing effects at 10.2 kHz. If the same calculations are made at other frequencies, the dispersive nature of refraction is evident – the phase errors in miles becoming much larger at the difference frequencies, as used in Omega to resolve carrier phase ambiguities.

Calculations from the observed maximum phase rate (249 cec/hr) shown in figure 6, using equation (13) yield a realistic value of night/morning fraction at 10.2 kHz

$$\frac{\Delta v_{n/m}}{v} = 0.0025 \text{ (compared with 0.004 for day/night ratio)}$$

*These are selections from the extensive file of large scale, high sampling rate recordings made available through the courtesy of J. Chamberlin and D. White of the Radio and Navigation Department, RAE Farnborough.

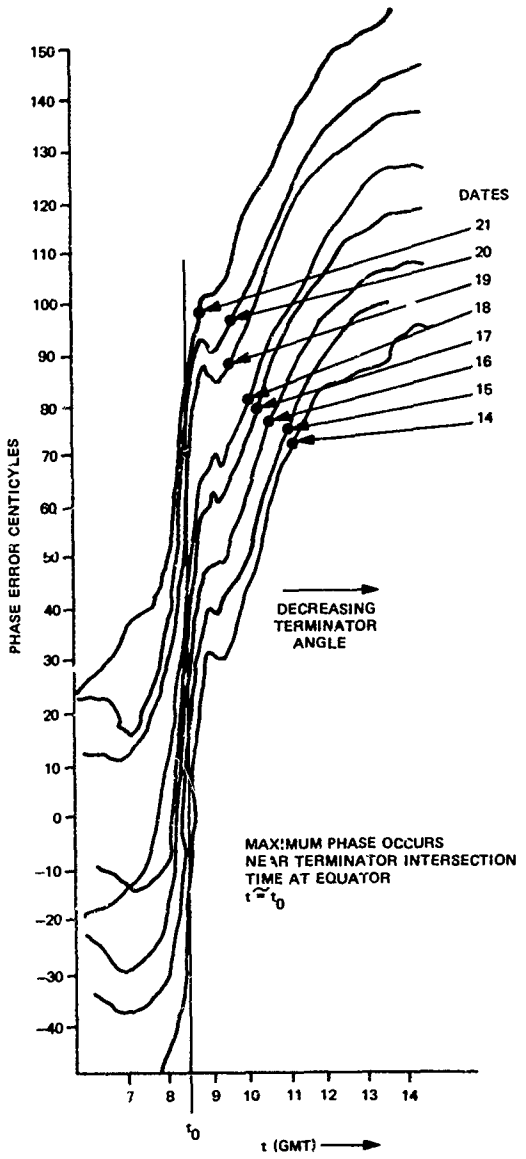


Figure 5. Argentina at Farnsborough on Successive Dates – Sunrise January 1977

From equation (3), Δt_c = 9 minutes, and $\Delta\phi_p$ = 37 cec, both quite reasonable approximations. The direct path prediction (without bending) is a scaled version of the B(t) function near t_c, thus corresponding to a theoretical case with bending effects removed, using a step model.

$$\underset{\text{cec}}{\phi\,(t)} \approx 6000\ B/\lambda\ \frac{\Delta v_{n/m}}{v_n} = 0.95\ B(t) \tag{16}$$

The comparison between the actual recording and this direct path prediction, illustrates approximately the quadrature phase error advance for nearly total reflection preceding onset, and a point of zero error when the reflection path difference is a half wave longer, as predicted by the step model.

The step model assumes an instantaneous change in propagation velocity moving along the intersection path. If instead the direct path prediction is convolved with a smoothing kernel, shaped as the actual terminator gradient, and of finite width, the prediction will start earlier and end later, yet cause little timing change at mid transition. The dotted line marked "without bending" corresponds with such a smoothed version of the direct path prediction. The difference between the actual recording and the dashed line is thus a better approximation of the bending error time function because it accounts for actual terminator width.

This illustrated case involves a terminator angle ranging from 3° to 5° during transition over a 6600-mile path. It should be apparent that at smaller terminator angles and in particular when one end of the path is on the equator, onset errors will be much larger. A review of RAE Farnborough phase recordings through the year – shows a characteristic step on the diurnal phase transition indicative of bending whenever terminator angle is small at sunrise. A similar effect at the end of transition occurs at sunset. In fact, the sunrise and sunset cases for the same geometry (6-month-time difference) are quite similar.

Rasmussen and Lewis have detected bending using interferometric techniques with spaced receivers[3,4]. Their analysis assumed a path entirely within the twilight gradient, using a linear gradient model and parabolic ray path. Application of the analytic technique described herein to their case for a short (2000 nm) path at 18 kHz gives comparable results to a linear gradient analysis.

VII. CONCLUSIONS

The above results using a step model confirm the applicability of geometric optics to prediction of phase errors due to (very long wave) propagation through the terminator and account for serious errors due to lateral bending in VLF radio location systems. Further refinement can be expected by use of a continuous gradient model, e.g., with the "scattering" matrix approach developed by Wait[1,2] extended by use of co-terminator geometry described herein.

VIII. REFERENCES

[1] Wait, James R., "Mode Conversion and Refraction Effects in the Earth-Ionosphere Wave Guide for VLF Radio Waves", Journal of Geophysical Research, June 1968.

[2] Wait, J.R., "On the Theory of VLF Propagation for a Step Model of the Nonuniform Earth-Ionosphere Waveguide" Canadian J. of Physics, 46, p. 1979, 1968.

[3] Lewis, Edward A. and Rasmussen, J. E., "Relative Phase and Amplitude Shifts of VLF Signals Received on Two Paths Almost Parallel with the Sunrise Line".

[4] Lewis, Edward A. and Rasmussen, J. E., "Phase Comparison of VLF Signals Propagated Over Adjacent Paths", I.E.E. 1967, London, England.

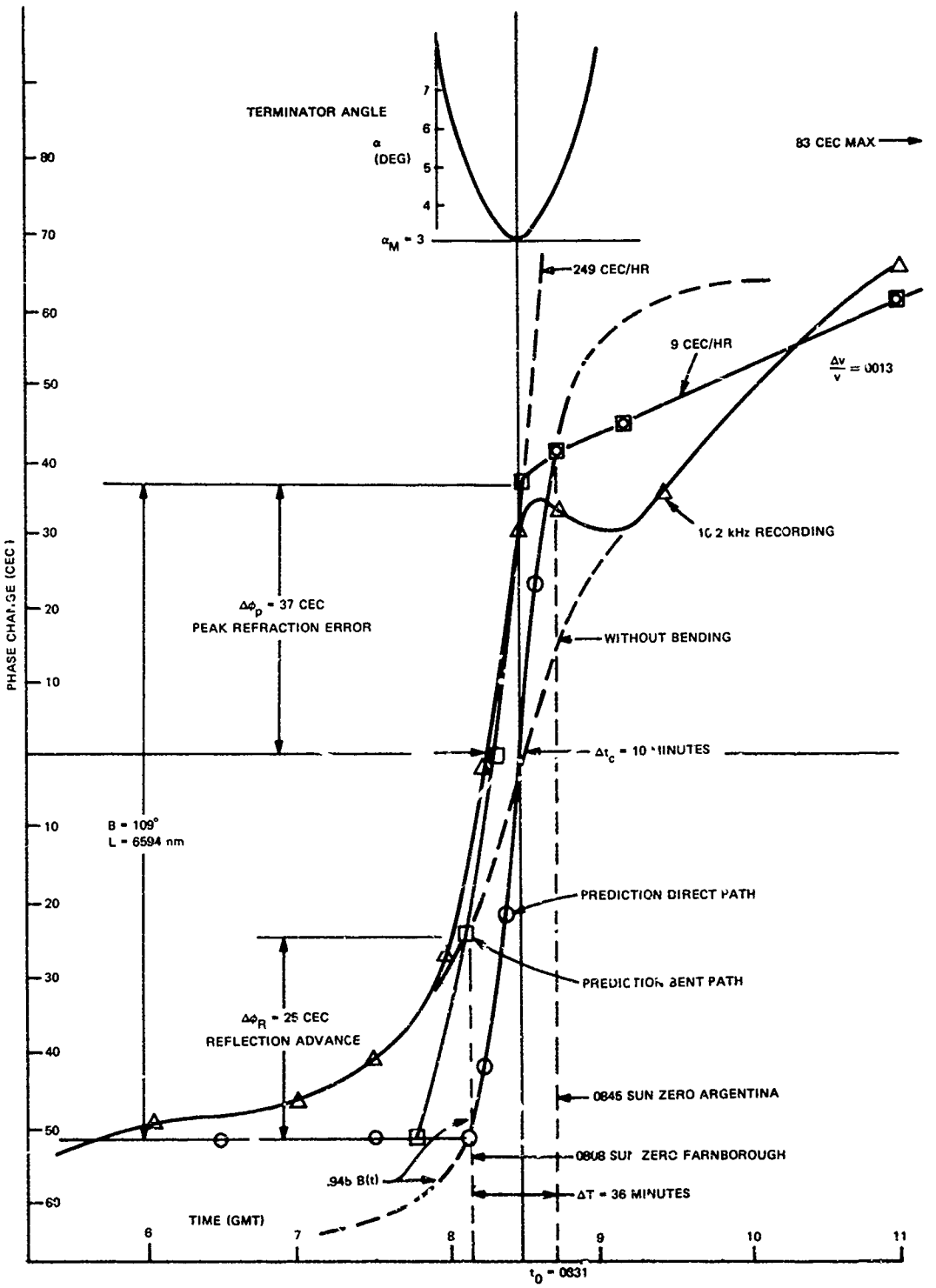


Figure 6. Calculated Bending Error

RESULTS OF A DIFFERENTIAL OMEGA EXPERIMENT

T.M. Watt, L.E. Abrams, and F.G. Karkalik
Systems Control, Inc.
Palo Alto, California 94304

I. INTRODUCTION

Omega, a VLF navigation system based on a worldwide network of eight transmitters operated at 10.2, 11.05, 11.333 and 13.6 kHz, provides global coverage for users. These frequencies are synthesized from a common source and are maintained in the exact ratio 1, 13/12, 10/9, 4/3. Phase coherence and emission timing are tightly controlled in the transmitter network. Table I lists Omega transmitter letter designations and locations.

Table I. Omega Transmitting Stations

Station Letter Designation	Location	Latitude/Longitude
A	Aldra, Norway	66°25'N/13°08'E
B	Monrovia, Liberia	06°18'N/10°40'W
C	Haiku, Hawaii	21°24'N/157°50'W
D	LaMoure, North Dakota	46°21'N/98°20'W
E	LaReunion	20°58'S/55°17'E
F	Golfo Nuevo, Argentina	43°03'S/65°11'W
G	Gippsland, Australia	38°29'S/146°50'E
H	Tsushima, Japan	34°37'N/129°27'E

A user wishing to navigate measures the phase difference between signals at one frequency from a pair of transmitters and thus establishes a line of position (LOP). Repeating the process with two more pairs of transmitters leads to a unique navigation fix. Navigational ambiguities can exist with Omega since any one phase difference corresponding to a pair of transmitters defines a family of hyperbolic LOPs. Along a baseline, LOPs occur every half wavelength. The region between adjacent LOPs is known as a lane and Omega accuracies are frequently described in units of centilanes (0.01 lane).

Omega is a VLF system and it is therefore subject to all the propagation anomalies normally associated with VLF. Some of the more important error sources associated with Omega are; (a) diurnal and seasonal ionospheric variations, (b) transient ionospheric phenomena such as Polar Cap Absorption (PCA) events and Sudden Ionospheric Disturbances (SIDs) that give rise to propagation anomalies, (c) modal interference, and (d) noise.

Diurnal and seasonal ionospheric variations cause phase shifts on the order of 50-100 centilanes at most user locations. These variations can be predicted and modeled to within a reasonable accuracy and are usually provided to users either as tables of correction values or as algorithms within a receiver's navigation processor.

Sudden phase anomalies are associated with SIDs caused by solar flare x-rays. These are daytime events and typically last about fifty minutes. Solar protons, associated with large flares, may be guided into the polar regions and produce PCA events. These events may affect polar region propagation for several days.

Modal interference describes the effect that occurs when more than one waveguide propagation mode is excited by a radiated signal. When this happens, the modes received by a user combine constructively and destructively and cause anomalous signal variations. This phenomenon occurs most commonly near a transmitter and when the propagation path crosses a twilight region.

Noise at VLF is mostly of atmospheric origin, although man-made noise can dominate in certain local regions. Noise effects can be diminished by integrating received signals over long periods, but the period of integration must be consistent with dynamic requirements of the user and expected transients in the signals.

Quoted accuracy for Omega under nominal conditions is 1-2 NM [1,3,4]. Nominal conditions include favorable geometry of available signals, the use of propagation prediction corrections (PPCs) to compensate for regular ionospheric variations, the absence of SIDs and PCA events, the absence of modal interference, and the absence of excessive noise. Under less favorable conditions, Omega accuracy degrades. Marine users on the high seas may find 1-2 NM accuracy acceptable and may even be able to tolerate limited periods of degraded accuracy. On the other hand, marine navigation in restricted waterways and aircraft navigation near terminals requires a higher level of accuracy and reliability.

The Differential Omega concept arises from the observation that many Omega navigation errors associated with propagation effects are highly correlated in time and space. For example, consider two Omega users navigating independently a short distance apart. The absolute error of each user's fix may be 1-2 NM, but the relative positional error will be perhaps an order of magnitude smaller. If a realtime data link could be established between the two users so that both sets of phase measurements could be correlated, then the two users could maintain a positional relationship accurate to within a fraction of a mile. This concept is known as Relative Omega. If we now consider that one user remains fixed at a known, surveyed location and provides real-time phase measurement data to the second user, then the second user can obtain absolute navigational accuracy to within a fraction of a mile. This concept is known as Differential Omega, the fixed user is called the monitor and the moving user is called the navigator.

The ability of Differential Omega to eliminate correlated errors points to a significant practical benefit, namely that the navigator need not provide or compute PPCs since such corrections are intrinsic to the differential corrections received over the data link. This means that the computational task of the navigator can be greatly reduced, thus yielding a corresponding reduction in size, power, weight and cost of the navigator's system.

Differential Omega, as a concept, has been recognized for at least 14 years [1-3]. Experimental verification of the concept has been somewhat limited [4-5]. Swanson and Davey [5] have described the results of a marine Differential Omega experiment conducted in the coastal waters off Galveston, Texas. Figure 1 illustrates results of navigational accuracy as a function of range from the monitor obtained by Swanson and Davey. These results indicate an accuracy of 0.2 NM at close ranges and a gradual degradation in accuracy with increasing range. At very long ranges, the error obtained

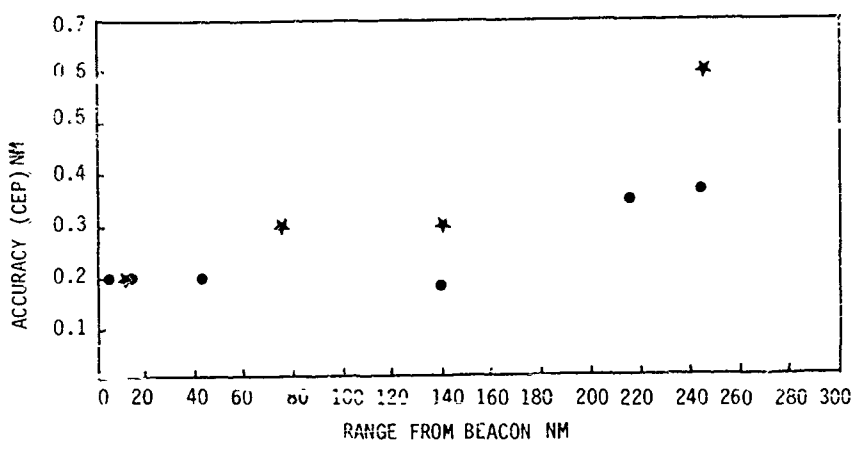


Figure 1. Differential Omega Accuracy vs. Range from Monitor (Reference 5).

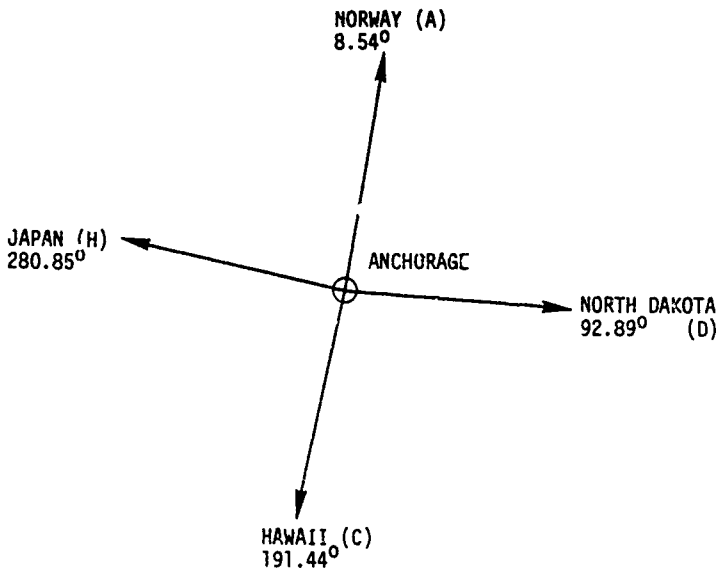


Figure 2. Azimuth Direction to Omega Transmitters from Anchorage, Alaska.

with Differential Omega may exceed the error obtained with ordinary Omega. The radius of the applicable region is limited both by the propagation range of the data link and the tolerable decorrelation error.

Navigational aids for aviation users include a broad spectrum of systems, that range from a simple non-directional beacon (NDB) to VOR/DME (and its military counterpart, TACAN) and radar. The most comprehensive and complex systems such as radar and VOR/DME are expensive, require frequent maintenance and consume a high level of electrical power. Less comprehensive navigational systems such as NDBs require very little maintenance and power and are relatively inexpensive. Throughout much of Alaska and Northern Canada, many remote communities depend entirely on aviation for supplies and transportation to and from the outside world. The facilities available at these remote sites frequently consist of little more than a landing strip large enough to accommodate a small aircraft. Energy is always a problem at such sites since fuel must be flown in, consequently, most remote airstrips do not even have the luxury of a colocated NDB. The costs of providing VOR/DME at every remote airstrip in North America would be prohibitive. Even NDBs with their lower costs and more modest energy requirements do not offer a completely attractive solution, since an NDB provides directional information only, and in order to be effective, must be located at the site being sought by the navigating aircraft.

The FAA in the United States and the Ministry of Transport in Canada (Transport Canada) have been seeking a solution to the requirement for a low-cost, accurate navigation system that will meet the needs of small aircraft flying in and out of remote locations in the northern part of the continent.

Differential Omega is considered to be a potential answer to this requirement. Studies [6] have shown that Differential Omega in the Alaska/Yukon region should be able to provide two-dimensional navigation accurate to within a fraction of a mile over a region within a hundred-mile radius from a monitor. It has been suggested that it is practical to colocate a Differential Omega monitor and an NDB and to use the NDB as a carrier for the differential correction information. This means that Differential Omega monitors could be deployed at existing NDB sites at relatively low cost and at very little increase in electrical power requirements. It also means that NDBs used in conjunction with Differential Omega would not have to be located at every airstrip, but could support navigation over a region that might include several airstrips.

II. DESCRIPTION OF EXPERIMENT

The experiment described herein has been primarily operational rather than research oriented. Under sponsorship of the FAA, an experiment has been conducted in Alaska to test the feasibility of Differential Omega for general aviation. The program represents a joint effort between the FAA and Transport Canada, where Transport Canada has provided design and development of the monitor, and Systems Control, Inc. (SCI) has provided program management, system engineering and flight test on behalf of the FAA. Tracor, Inc. has provided the airborne Omega equipment under contract to SCI. The location of the experiment, Alaska and Northwest

Canada, has the interesting characteristic that the azimuth directions of signals from stations A, C, D, and H intersect at nearly right angles, as illustrated by Figure 2.

The experimental concept calls for the use of operational NDB signals as carriers for differential correction data. Since NDBs have a dual mission of automatic direction finding (ADF) and weather broadcast, a ground rule for the Differential Omega experiment was that any additional modulation on the beacon signal (in this case, biphase modulation of Omega phase correction data) was not to degrade the quality of the signal received by other NDB users. This requirement mandated a maximum modulation level of about 20% for phase correction data which limited the effective operating range between the monitor and the navigator aircraft.

Figure 3 illustrates the experimental configuration. The avionics were mounted on a special pallet in an FAA Convair 580 aircraft based at Anchorage International Airport. Monitor stations were located at Merrill Field about 5 miles east of Anchorage International, and at Deadhorse, on Prudhoe Bay. Omega receivers and NDB transmitters were approximately colocated at both monitor stations. Reference location information for the navigator aircraft was provided by surveyed pads for ground tests, and by DME instrumentation for flight tests. Figures 4 and 5 illustrate the geography of the experiment and available instrumentation. These figures display every VOR/DME within 250 miles of Anchorage and Deadhorse. Although these navigation aids provide reasonably redundant coverage for enroute navigation of aircraft flying at high altitudes, their line-of-sight range is proportionately less at low altitudes so that they degenerate to short-range homing aids for general-aviation users who are limited to altitudes less than about 10,000 ft. It can be seen that for the low-altitude user, there are vast areas in Alaska where there is no effective coverage by VOR/DME.

It is instructive to provide a brief description of the operation of the monitors and the determination of differential corrections. We begin by expressing the known location of a monitor in terms of standard phases. A standard phase is defined by a monitor location, an Omega transmitter location, an Omega frequency, an assumed propagation velocity, and a geodetic model. First, the propagation range between a monitor and an Omega transmitter is calculated using an appropriate geodetic model. Next, the propagation range is expressed in wavelengths for the particular frequency. Finally, the integer number of wavelengths is discarded and the fractional wavelength is retained. This fractional wavelength is known as a standard phase and it is a highly sensitive indicator of monitor location. Standard phases from three Omega transmitters define the location of a monitor uniquely except for lane ambiguities.

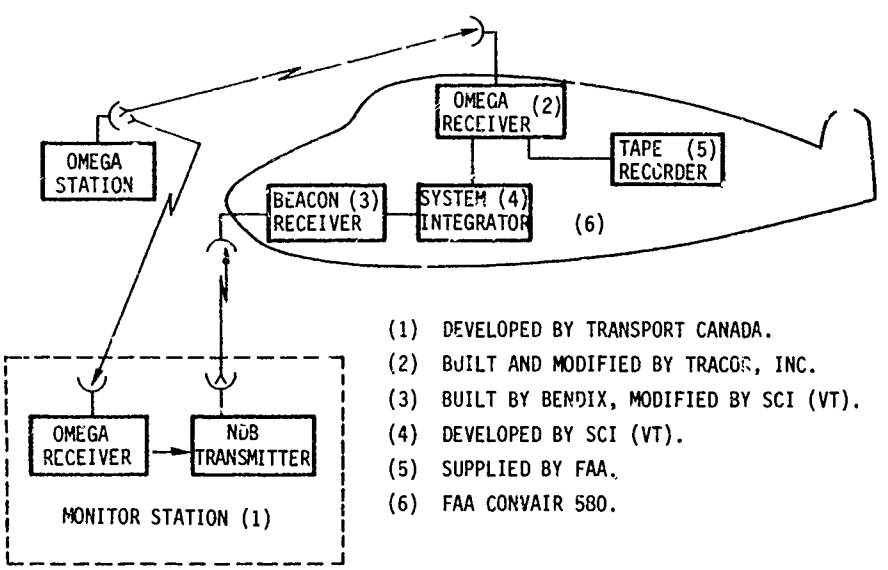


Figure 3. Differential Omega System Configuration

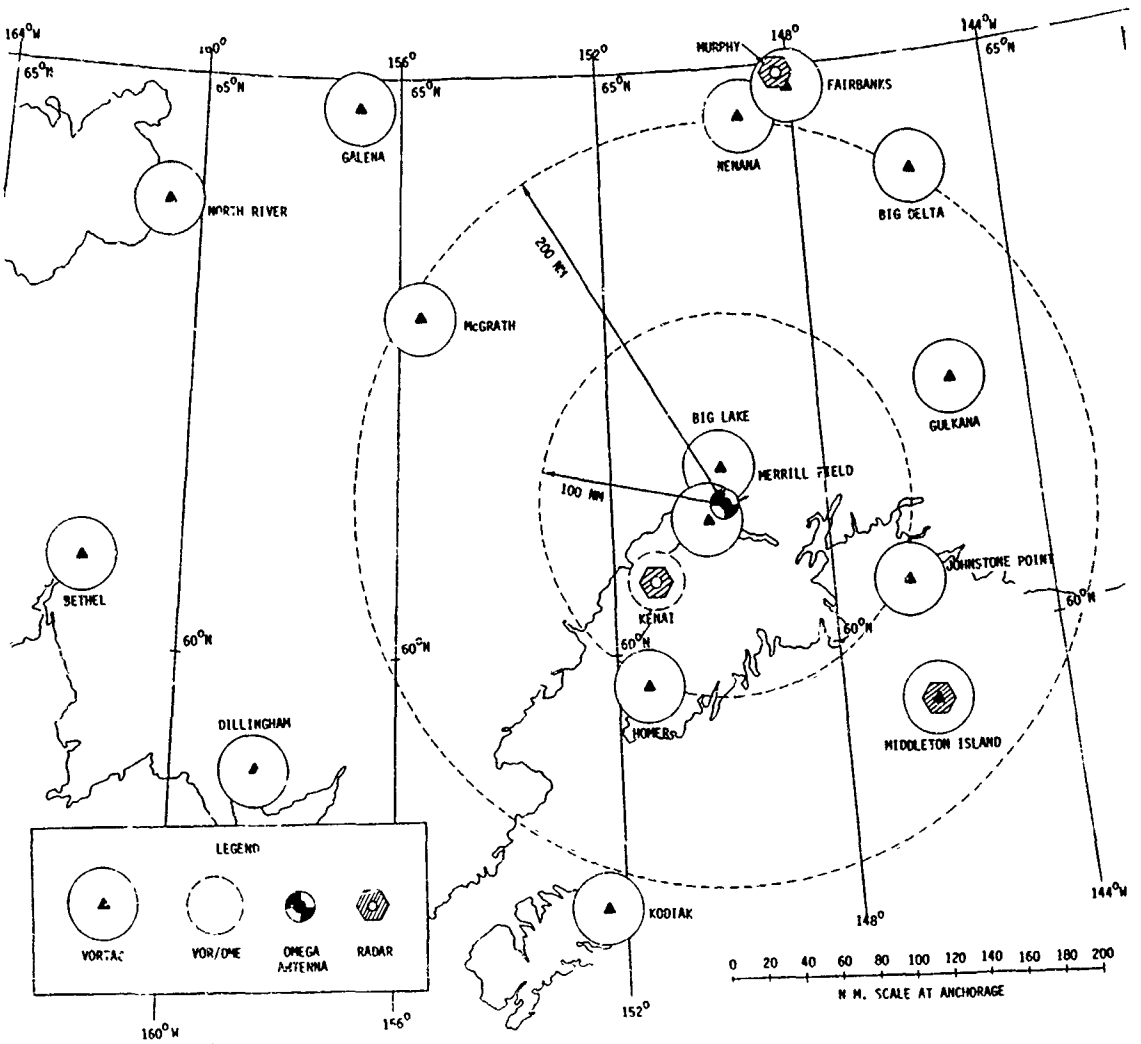


Figure 4. Instrumentation Available for Flights Near Anchorage.

We next consider the arithmetic involved in providing a differential correction. Each monitor receiving system consists of three Omega receivers, a Rubidium frequency standard driving a frequency synthesizer and a microcomputer. The frequency synthesizer provides stable outputs at 10.2, 11.33 and 13.6 kHz. Each Omega receiver is tuned to one of the three frequencies and usually receives signals from Stations A, C, D, and H. At each frequency, the phase of a signal from one transmitter is compared with the phase of the locally synthesized signal, and the phase difference is transmitted as a correction. We have, for each signal,

$$\phi_s - (\phi_m - \phi_r) = \Delta \qquad (1)$$

where ϕ_s = standard phase
ϕ_m = measured signal
ϕ_r = local reference phase
Δ = differential correction.

Ideally, the local reference phase ϕ_r would be identical to the phase at the Omega transmitter in which case Eq. (1) would express the relationship: True Range - Measured Range = Range Error. In fact, the local reference phase differs from the transmitter phase by an arbitrary unknown value. Because of the precision of

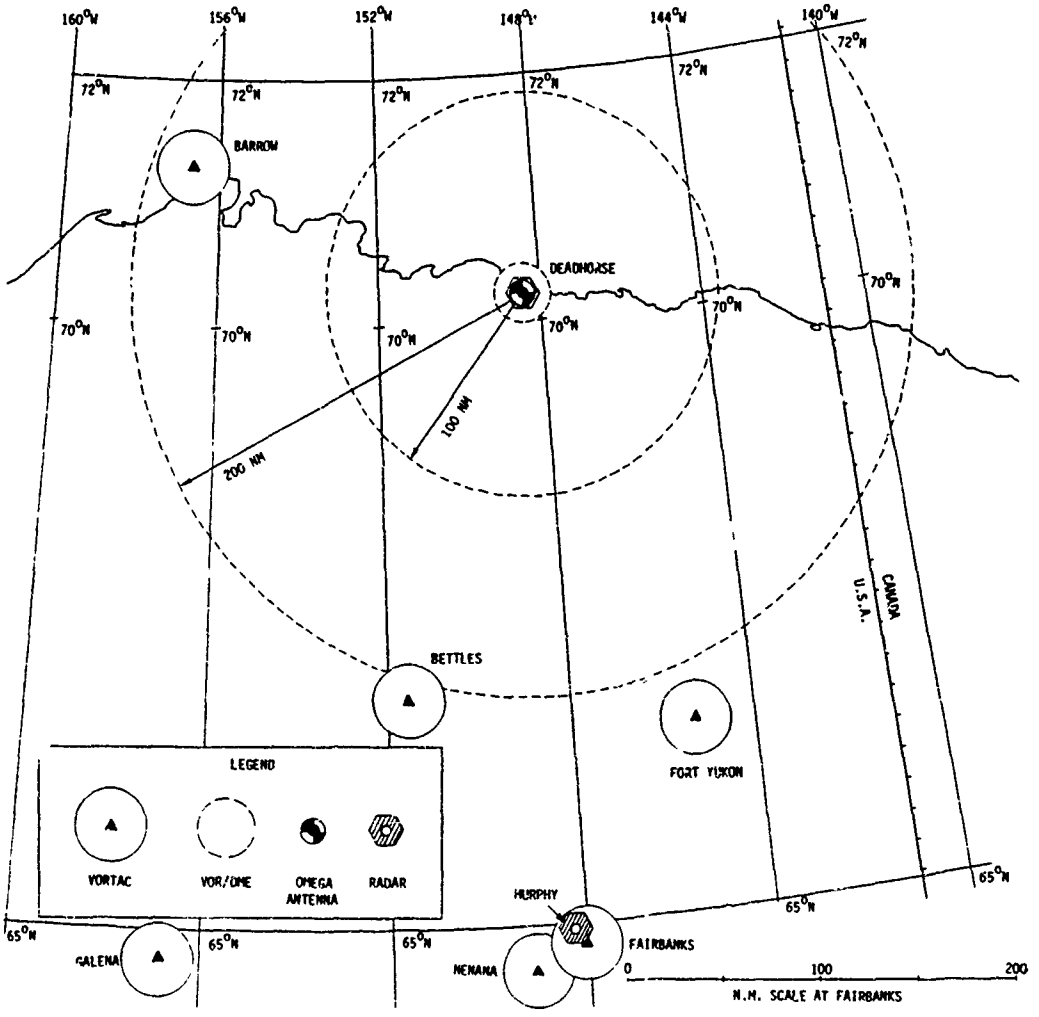


Figure 5. Instrumentation Available for Flights Near Deadhorse.

the local frequency reference, however, the unknown phase difference between the monitor local reference and the transmitter varies quite slowly. The local reference phase disappears in the process of forming an LOP which, as we have stated, involves forming phase differences between signals measured from two transmitters. If we apply this process to Eq. (1) for any two Omega transmitters labeled No. 1 and No. 2, we have

$$\phi_{s1} - (\phi_{m1} - \phi_r) = \Delta_1 \tag{2}$$

$$\phi_{s2} - (\phi_{m2} - \phi_r) = \Delta_2 \tag{3}$$

Subtracting Eq. (3) from Eq. (2) yields a quantity $\Delta_1 - \Delta_2$ which is a correction to be applied to an LOP and which is independent of the local reference phase at the monitor.

In this experiment each phase correction as expressed by Eq. (1) was transmitted via the NDB. The appropriate combinations, as expressed by the difference $\Delta_1 - \Delta_2$, were performed by the navigation computer within the avionics according to the LOPs being computed. A complete correction message was transmitted every ten seconds and consisted of twelve correction values; i.e., four transmitters at three frequencies each.

The experiment described here took place during the period October 16 through October 24, 1980, and February 23 through 27, 1981, and was conducted in four parts. The first part involved Differential Omega navigation while the aircraft was parked at a known location. Ten-second samples of navigational solutions from the avionics were recorded for fifteen minutes each on several occasions, yielding statistical performance data at a fixed location free of the complications associated with flight testing. The second part of the experiment consisted of recording phase data in one-minute samples for several days as received by the monitors at both Merrill Field and Deadhorse. The measurement data taken from each monitor provided information on diurnal variations in phase associated with regular ionospheric behavior. Comparison of the phase data between the two monitors yielded information on range decorrelation error for Differential Omega. The third part of the experiment involved flight tests performed during routine flights of the aircraft out of Anchorage International Airport. These tests yielded information on in-flight performance of Differential Omega in terms of accuracy and maximum range of the data link. The fourth part of the experiment involved dedicated flights of the aircraft along a localizer beam at Elmendorf Air Force Base in order to obtain estimates of Differential Omega navigation accuracy under conditions approximating non-precision approach operations.

Differential Omega Ground Tests

The Differential Omega ground tests were performed as follows. The aircraft was parked on a pad at Anchorage International Airport. The monitor computer at Merrill Field was loaded with the appropriate standard phases. The system was operated in the Differential Omega mode using correction data from the monitor at Merrill Field. Navigation solutions at the aircraft based on ten-second sampling periods were recorded for fifteen minutes. The standard phase values in the monitor computer were then modified to simulate a displacement of the monitor two miles north of its actual position. Ten-second samples of navigation solutions at the aircraft were again recorded for fifteen minutes. The standard phase values in the monitor computer were then modified to simulate a displacement of the monitor two miles west of its actual position. Ten second samples of navigation solutions at the aircraft were again recorded for fifteen minutes. The three sets of measurements were performed twice, once between 11 AM and 12 noon, and once between 6 PM and 7 PM local time.

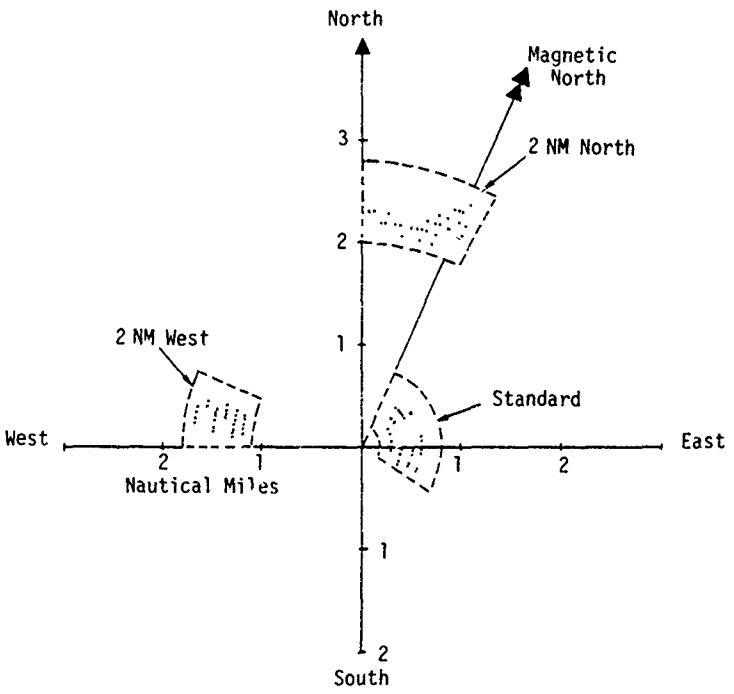


Figure 6. Representative Results of Differential Omega Ground Tests, October 19, 1980.

Figure 6 illustrates an example of the results obtained from the Differential Omega ground tests. The origin of the plot is defined to be the pad location, 61°10'27"N, 149°58'06"W, and the plotted points represent the navigation solutions obtained during the tests. The three groups of solutions correspond to the three sets of standard phases loaded into the monitor computer. The results shown in Figure 6 are representative of all results obtained from the Differential Omega ground test.

Analysis of the results of these tests yields the following observations.

(a) The 2-D RMS value (95%) of the ten-second navigation solutions was about 0.25 nautical miles. Since the sampling rate was not adjustable, there was no opportunity to investigate the dependence of random error statistics on sampling period.

(b) Mean error of the test data was about 0.5 nautical miles eastward, 0.25 nautical miles northward. This error is not necessarily attributable to Differential Omega per se, since it could easily be caused by survey errors at the monitor or the aircraft pad.

(c) Mean error at any single location can be zeroed out by adjusting the standard phases at the monitor. Adjustment of the mean error had no observable effect on the random error component of the navigation solutions.

Monitor Phase Correlation Tests

Received Omega phases at the Merrill Field and Deadhorse monitors were sampled and recorded at one-minute intervals nearly continuously for several days. Phase data at each monitor yielded information on (a) long-term drift of the monitor frequency standard relative to the standard of the Omega transmitter network, (b) diurnal phase variations associated with regular ionospheric effects, and (c) the presence of phase anomalies with periods greater than one minute. In addition, comparison of phases between the two monitors yielded information on the range decorrelation error of Omega navigation solutions between the two monitor locations.

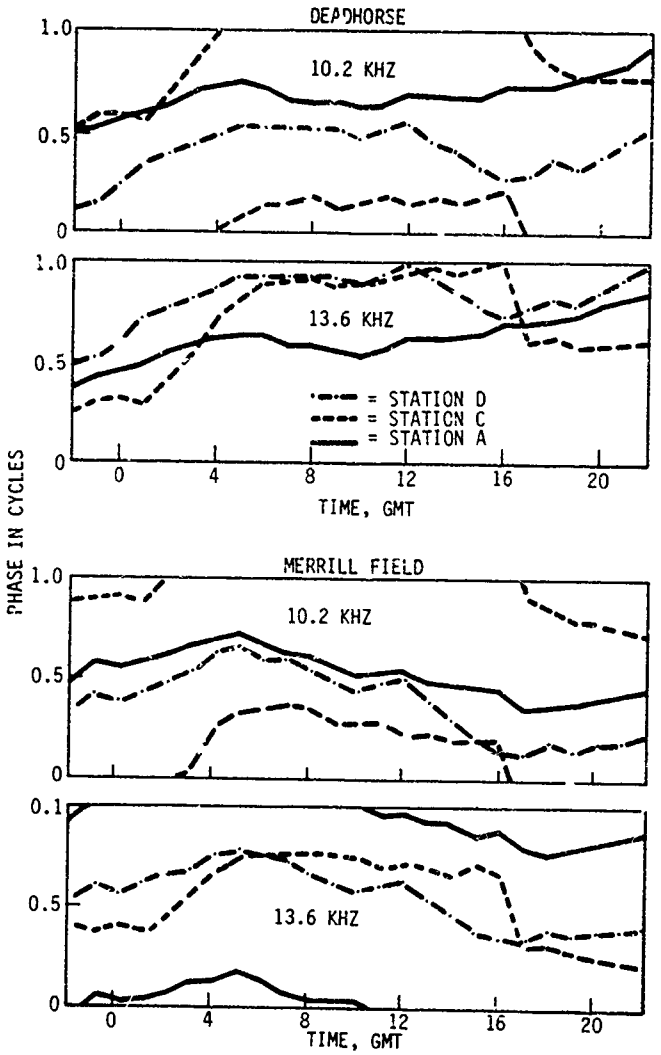


FIGURE 7. OMEGA MONITOR STATION PHASE MEASUREMENTS, OCTOBER 23, 1980

Figure 7 illustrates 24 hours of phase data recorded at the Merrill Field and Deadhorse monitors. Curves are presented for signals at 10.2 kHz and 13.6 kHz from stations A, C, and D. Station H was temporarily off the air during these measurements. Phase is measured modulo one cycle and cycle rollovers are reflections of continuous phase variations.

Examination of Figure 7 reveals the following features:

(a) There is great similarity in the gross behavior of corresponding signals at 10.2 kHz and 13.6 kHz although fine structure appears to be uncorrelated.

(b) Diurnal phase variations are most pronounced and most regular from station D, and least pronounced and least regular from station A. These observations are consistent with the facts that propagation from D is

essentially through midlatitudes along a meridian where ionospheric behavior is well-behaved and daily solar zenith angle variations are large; whereas propagation from A is through the polar cap which is less regular and where daily solar zenith angle variations are small.

(c) There is evidence of a slow (one-half cycle per day) drift in the phase of the Deadhorse frequency standard with respect to the standard of the Omega transmitter network. This effect is evident at both 10.2 kHz and 13.6 kHz. Any drift at Merrill Field is much smaller and is not readily discernable over a 24 hour period.

Comparison of the phases received at the two monitors can provide information on the correlation of Omega signals between the two monitors. The degree of correlation is a direct measure of the utility of Differential Omega over the distance between the two monitors.

The description of range correlation can be approached in the following way. Suppose, using the monitor at Merrill Field, one formed a differential correction from signals at 10.2 kHz received from stations A and C. Using a form of Equations (1)-(3), one obtains

$$\delta_{ACM} = \Delta_{AM} - \Delta_{CM} \qquad (4)$$

where δ_{ACM} is a differential correction to the A-C LOP formed at Merrill Field and Δ_{AM} and Δ_{CM} are obtained from Equation (1) as applied to signals from stations A and C, respectively. The number represented by δ_{ACM} will be applied by a nearby navigator as a differential correction to the A-C LOP formed by the navigator. The navigator then expects that the corrected LOP will be nearly errorfree. The A-C LOP can be combined with two other LOPs in the navigator's computer to form the desired fix. Let us now repeat the process, using the Deadhorse monitor. We have, analogous to Equation (4)

$$\delta_{ACP} = \Delta_{AP} - \Delta_{CP} \qquad (5)$$

where subscript P refers to the Deadhorse monitor. Let us now imagine a navigator hovering directly over the Deadhorse monitor who purports to navigate with Differential Omega using either the Merrill Field corrections described by Equation (4) or the Deadhorse corrections described by Equation (5). If there were perfect correlation between Merrill Field and Deadhorse, the result should be independent of which monitor is used, thus perfect correlation implies that, at each instant of time, $\delta_{ACP} = \delta_{ACM}$. The extent of disagreement between the two sets of differential corrections is, therefore, a measure of the lack of correlation between the two locations and may be described as range decorrelation error for the pair of locations.

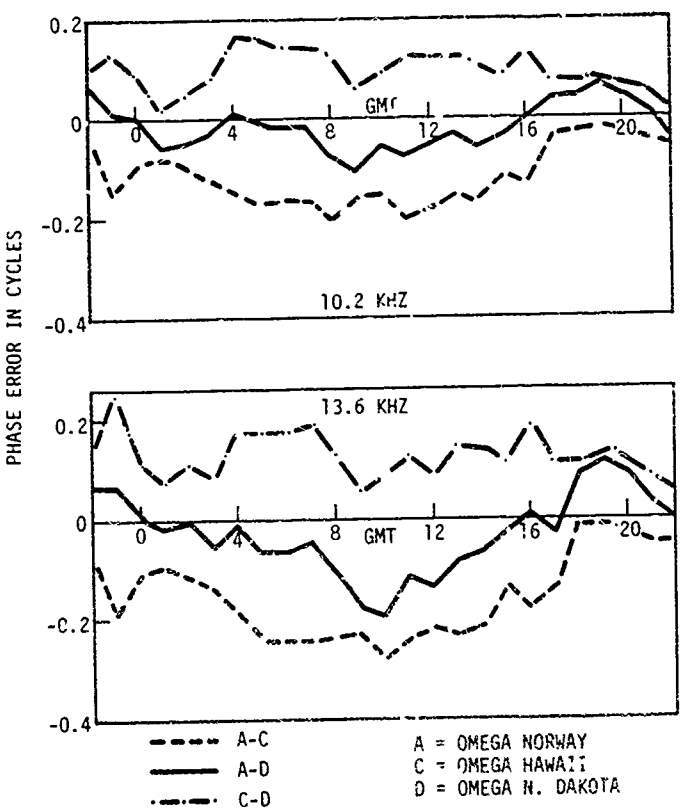


Figure 8. Decorrelation Errors Between Deadhorse and Merrill Field Monitor Stations, October 23, 1980.

Figure 8 illustrates values of $\delta_{ACP} - \delta_{ACM}$, $\delta_{ADP} - \delta_{ADM}$ and $\delta_{CDP} - \delta_{CDM}$; that is, range decorrelation errors between corresponding LOP corrections obtained at Deadhorse and Merrill Field for the three possible

pairs AC, AD and CD. Range decorrelation errors between Deadhorse and Merrill Field are seen to exhibit the following characteristics:

(a) Gross behavior is similar between 10.2 kHz and 13.6 kHz.
(b) A diurnal pattern is evident for each LOP, but the patterr is complex and is not the same for all LOPs.
(c) The total range of error observed during the 24-hour period is about ±0.2 NM. The maximum excursion of any LOP error is about 0.2 NM.

Although Deadhorse and Merrill Field are separated by about 550 NM, a much greater range than is considered for Differential Omega validity, the results illustrated by Figure 8 suggest that even at this range, the accuracy achievable from Differential Omega would be comparable to that achieved with ordinary Omega. As discussed below, however, Differential Omega operation was limited to lesser ranges by signal quality on the data link.

Data-Link Range Tests

Flight tests were performed to verify Differential Omega operation and performance under actual flight conditions. The flight tests provided information on two primary indicators of performance; maximum range of the data link and accuracy of the navigation.

Four round-trip flights out of Anchorage International Airport were flown during the experiment, as follows: October 16, Anchorage to Galena to Anchorage; October 17, Anchorage to McGrath to Anchorage; October 22, Anchorage to Bettles to Anchorage; October 23, Anchorage to Deadhorse; October 24, Deadhorse to Anchorage. Table II lists the flights and the maximum ranges of the data link that were observed. Maximum range was defined in terms of received signal quality according to an algorithm that measured error rate in the differential correction data. When error rates exceeded a preset threshold, the differential correction message was rejected and the maximum range of the data link was deemed to have been exceeded.

Table II. Observed Differential Omega Maximum Range in NM

Date	Monitor	Outbound Flight	Inbound Flight
October 16	Merrill Field	55	107
October 17	Merrill Field	128	151
October 22	Merrill Field	68	103
October 23	Merrill Field	44	---
October 23	Deadhorse	---	No data
October 24	Deadhorse	198	---
October 24	Merrill Field	---	147

The results shown in Table II suggest the following:

(a) Maximum range of the data link varied by a factor of two from day to day. Since the routes and times associated with the flights on October 16, 17, and 22 were similar, the most likely explanation for the daily variation appears to be the level of radio noise present on the data link. In fact, concurrent measurements of precipitation static in a collateral experiment suggest support for this explanation.

(b) The smallest maximum range observed during this experiment was 44 NM. There is no reason to believe this figure could not be improved by using higher levels of modulation at the monitor NDB or by means of other changes designed to optimize the performance of Differential Omega.

(c) There is a strong indication of a non-uniform gain pattern in the beacon antenna of the aircraft. The ratio of average maximum range on inbound flights to average maximum range on outbound flights for the flights on October 16, 17 and 22 is about 1.4 and suggests a front-to-back gain ratio of 3 dB in the aircraft antenna.

Measurements were made to determine the accuracy performance of Differential Omega during the flights of October 16-24. The technique used was to record simultaneously samples of DME measurements and navigation outputs of the Omega equip-

ment. DME measurements were of slant range from the aircraft to the DME transponder being interrogated by the aircraft. DME accuracy is considered to be about one percent of the range being measured.

The accuracy measurements were hampered by several problems in the aircraft. First, data from one of the two DMEs failed consistently to record properly, thus determinations of Omega accuracy could only be made along the direction defined by the other DME measurement. Second, true airspeed data from the aircraft to the Omega were not available for this experiment. True airspeed is an essential input to the Omega system, and missing or incorrect data causes significant error in the navigation solution. Because of these difficulties, none of the flight tests during October yielded results that are considered representative of the navigation potential of Differential Omega.

In order to obtain performance data on navigation accuracy, additional flight tests were scheduled for the week of February 23-27, 1981. Equipment malfunctions negated testing until February 27 when a single, dedicated flight was scheduled for the exclusive benefit of Differential Omega testing.

During the portion of the flight of interest here, the aircraft was flown three times at low altitudes along a localizer beam at Elmendorf Air Force Base. During each pass, the aircraft DME measured slant range to the localizer/TACAN antenna and special equipment aboard the aircraft measured azimuth location of the aircraft with respect to the center of the localizer beam. Following the first and second pass along the beam the aircraft proceeded counter-clockwise to intersect the beam again. Data were recorded every ten seconds. During these passes the data-link range was about 10 NM.

Figure 9 illustrates results representative of these tests. In Figure 9 the tips of the arrows describe the location of the aircraft at various times as determined by the localizer and TACAN. The square dots at the bases of the arrows describe the Differential Omega solutions obtained at the same times.

The results illustrated in Figure 9 suggest the following:

(a) The magnitude of navigation error in the Differential Omega solutions varied from about 1.0 NM to about 0.1 NM during these tests.

(b) The error increased suddenly after each aircraft meneuver and then converged to near zero with an effective time constant of 1-2 minutes. The prevalent error was in the form of an overshoot in aircraft position.

III. CONCLUSION

The experiment described here is part of an ongoing effort to examine the utility of Differential Omega navigation for general aviation in Alaska and Northwest Canada. The focus of the experiment has been on operational issues rather than on scientific research.

The primary objective of the experiment was to evaluate the performance of a prototype Differential Omega system using a monitor nearly colocated with an NDB and an Omega navigator mounted on an FAA aircraft. Performance parameters were (a) maximum range of the data link, (b) mean and variance of navigation error, and (c) range decorrelation error.

The maximum range of the data link was measured on four round-trip flights out of Anchorage, Alaska. The observed maximum ranges varied from 44 NM to 198 NM and appeared to depend on the level of radio noise present on any given day.

Mean and random components of navigation error were tested at a surveyed site about five miles from a monitor. Mean error was observed to be about 0.5 NM and was adjustable at the monitor. The random error was observed to be about 0.25 NM 2-D RMS.

Range decorrelation error was observed by comparing Omega phases at two monitors separated by 550 NM. Maximum range decorrelation error observed during any 24 hour period was 0.2 NM.

The ground tests generally yielded results that were consistent with the results of earlier experiments. The random errors illustrated in Figure 6 agree

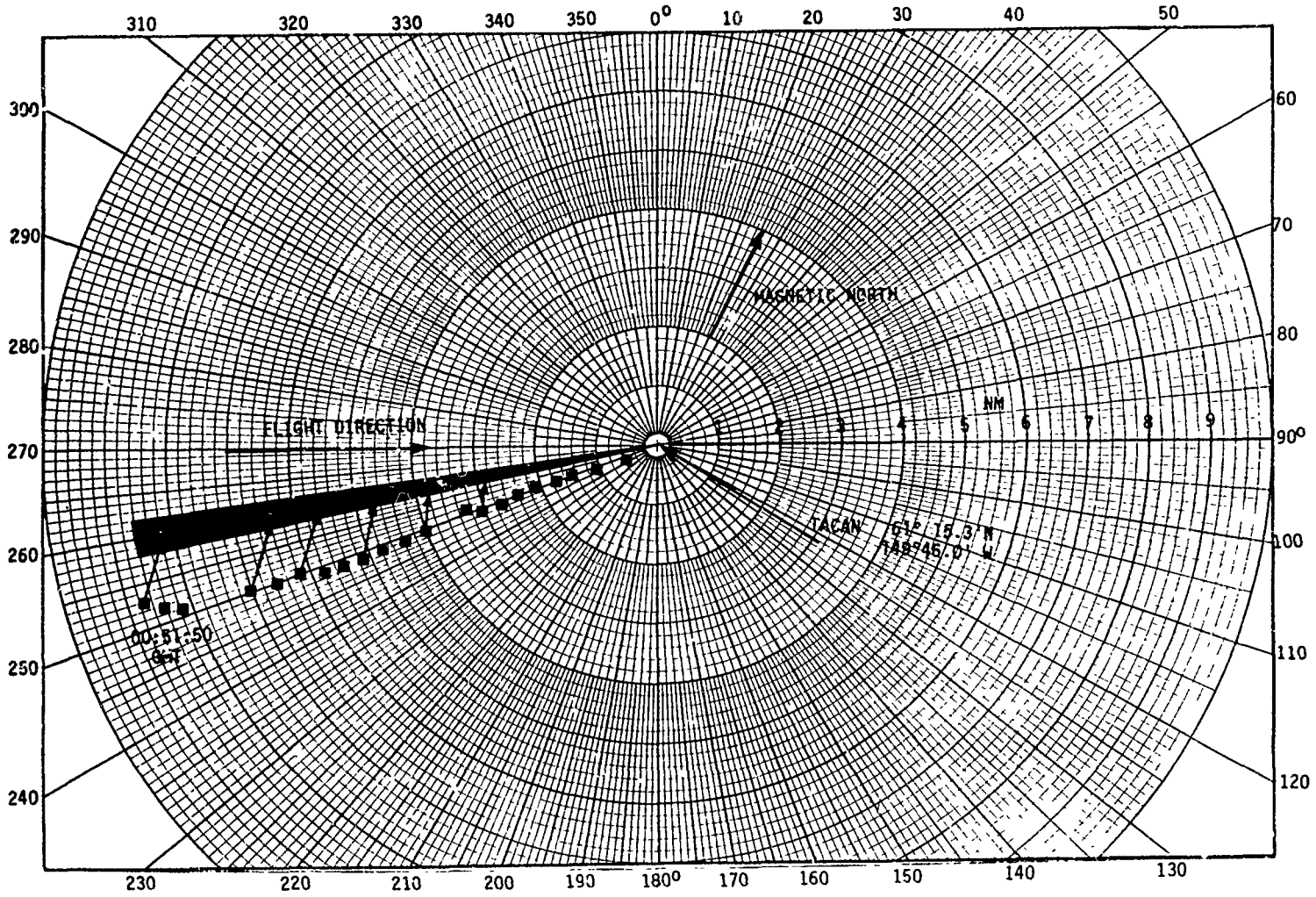


Figure 9. Comparison of Differential Omega and DME Positioning During Flights Along the Localizer Beam at Elmendorf Air Force Base on 27 February 1981, Third Approach.

quite well with the minimum-range results illustrated in Figure 1 [5]. The range decorrelation errors illustrated in Figure 8 are also consistent with an extrapolation of the results of Figure 1 to a range of 550 NM.

The accuracy results obtained during flight tests were poorer than expected. Some of this navigation error may be attributable to instrumentation errors which are known to have been present. However, a significant component of Differential Omega error is due to inadequate compensation of aircraft dynamics.

Many of the results obtained during this experiment corroborated and extended earlier results. In addition, several areas were identified that require more study. In particular, additional flight testing is required to gain more information on the behavior of maximum range from the monitor as a function of precipitation noise, aircraft antenna design, and modulation level at the NDB. Additional flights with improved instrumentation will provide more data on navigation accuracy of Differential Omega under various flight conditions. Finally, subsequent experiments should include a transportable monitor that can be colocated at various NDB sites to permit extended measurements of range decorrelation error at relatively low cost.

The program is still active and more experiments are planned. It is expected that answers to many of the questions raised here will be answered in the months to follow.

IV. REFERENCES

1. Brogden, J.W., and Luken, K.O.L., "Differential Omega," NRL Memorandum Report 1716, 8 August 1966.

2. Tracor, Inc. Report 67-135-U, "Final Technical Report-Differential Omega Test and Evaluation Program," Contract CG-14, 091-A, 18 January 1967.

3. Luken, K. Brogden, J.W., and Meyers, W.D., "Accuracy Studies of the Differential Omega Technique," NRL Report 7102, Naval Research Laboratory, Washington, D.C., 29 June 1970.

4. Swanson, E.R., Adrian, D.J., and Levine, P.H., "Differential Omega Navigation for the U.S. Coastal Confluence Region," NELC/TR 1905, Parts I and II, Naval Electronics Laboratory Center, San Diego, California, 2 January 1974.

5. Swanson, E.R., and Davey, J., "Differential Omega Evaluation," NOSC/TN 187, Naval Ocean Systems Center, San Diego, California, 27 June 1977.

6. Foltz, J.M., et al, "Differential Omega in the Alaska/Yukon Region: An Analytic Assessment," Report FAA-RD-79-77, The Analytic Sciences Corporation, Reading, Massachusetts, August 1979.

MODELLING THE C & D-REGIONS

W. Swider
Air Force Geophysics Laboratory
Hanscom AFB, MA 01731

ABSTRACT

Our steady state model has been shown previously to give good results with respect to the disturbed, solar proton event (SPE), D-region. Twenty-one positive and eight negative ions are included in this model which requires less than one minute of computer time. The negative ion chemistry represents a simplified scheme which appears to provide reasonable electron profiles for both day and night. The model is applied here to the normal (quiet) D-region. The deduced quiet electron profile contains a lower altitude portion which may correspond to the so-called C-layer, where ionization by cosmic rays is dominant. This layer is not always evident. Past observations are discussed briefly in terms of the difficulties of making accurate measurements of electron concentrations at these altitudes, near 65 km, by means of rocket-borne instruments. We note that recent studies by T.M. York et al., which suggest theoretical modifications to certain in-situ probe data, lead to electron profiles in fair agreement with our predicted (model) results. A typical computed profile shows a C-layer upper boundary near 65 km, $d[e]/dz \approx 0$, and $[e] \approx 75\ cm^{-3}$. This region may vanish as a "layer" if NO is enhanced in the D-region and/or if ionization from precipitating particles is present. The "layer" also tends to disappear with decreasing solar zenith angle. Formation of a C-layer appears to be compatible with our determination of an altitude of about 65 km for unity λ, the ratio of negative ions to electrons. This height rises to about 80 km at night for $\lambda = 1$, the parameter being mainly a function of that solar photon radiation which gives rise to the detachment of negative ions both directly and through the formation of atomic oxygen by ozone photodissociation.

INTRODUCTION

The D-region is the smallest portion of the ionosphere but yet it is the most complex region to model and to measure. Measurements of electron concentrations in this altitude regime, about 50-90 km, have been accomplished by a number of techniques. Ground-based investigations have employed (Thrane, 1974) the ionosonde, incoherent scatter, partial reflections, wave interactions, wave absorption and VLF/LF measurements while the application of rockets has provided for data acquisition through in-situ probes, Faraday rotation and differential absorption methods, VLF/LF scalar field measurements and VLF doppler plus a few other techniques. The variety of observational approaches available undoubtedly are related to the difficulties of determining electron concentrations in this region. These problems may be attributed to the low quantity of electrons being measured (except for those times when the concentration is greatly enhanced by precipitating energetic particles) and the importance of electron collisions. For example, at higher altitudes (E and F regions), the absence of electron-neutral collisions implies that the measured in-situ probe current is directly proportional to the electron population and if a probe calibration is not performed the data can be normalized with respect to a local ionosonde determination of the peak E or F region electron concentration. The fact that electron-neutral collisions are important leads directly to chemical complications since appreciable negative ion formation, a characteristic of the lower D-region, requires the initial three-body process $e + O_2 + O_2(N_2) \rightarrow O_2^- + O_2(N_2)$. The presence of triple collisions also makes the positive ion chemistry of this region far more complex than that portion of the ionosphere immediately above, the E-region.

In this paper, we shall argue that a model developed for the highly disturbed D-region, in particular, the November 2-5, 1969, SPE, is suitable also for the quiet D-region. The seasonal, latitudinal and solar cycle variations of nitric oxide, alluded to only very briefly, are a major complication in formulating a D-region model. However, computations will be shown to compare favorably with some experimental results. It will be argued that some of the discrepancies between the model and certain data may be due in large part to the uncertainties inherent in many of the measurements (Thrane, 1974). A particular concern of this paper will be the lower D-region, sometimes called C-region which from a theoretical aeronomic standpoint is most likely to be present when high solar zenith angles are present and hence low ionization levels in the upper D-region for NO^+ formation Cosmic rays, almost non-varying, ionize the C-region.

MODEL CHARACTERISTICS

The model has been run on a CDC 6600 computer and has an execution time of only a few seconds on the central processor. A recent version of the program is described fully elsewhere (Swider and Foley, 1978). The current code now includes five reactions dealing with the collisional breakup of certain NO^+ cluster ions and thus has a positive ion reaction list very similar to that employed by Reid (1977). The program encompasses 21 positive ions and 8 negative ions: O^-, O_2^-, O_3^-, O_4^-, CO_3^-, CO_4^-, NO_2^- and NO_3^-. The total number of reactions in the program is 83, not including the fact that an ion-ion recombination rate coefficient of $3 \times 10^{-8} cm^3 sec^{-1}$ (Smith et al., 1976) has been adopted for all ion-ion neutralization processes.

The model requires an extensive input of neutral constituent data. Temperatures and total neutral particle concentrations, [M], were taken from the 1976 U.S. Standard Atmosphere, the mesopause temperature being 187°K at 86-90 km. Values for O_2 and N_2 are 0.21 [M] and 0.78 [M], respectively. Carbon dioxide and water vapor concentrations have been set at 300 ppmV and 4 ppmV, respectively, except for a few early computations where 1 ppmV H_2O was utilized. (This difference in $[H_2O]$ has an insignificant impact on the computed electron concentration, [e].) Other neutral species required in the computations are O, O_3, $O_2(^1\Delta)$, NO and NO_2. The concentrations for all these species are identical to those used by Swider et al., (1978) in their SPE-disturbed D-region calculations. However, we also shall consider here two other lower concentration nitric oxide profiles. In addition, an increase in $[O_2(^1\Delta)]$ has been instituted for the quiet D-region calculations with peak concentration of $3.5 \times 10^9 cm^{-3}$ at 86 km.

The quiet ionization sources utilized in the computations include

$$q(NO^+) = 4 \times 10^{-7} [NO] \exp\{-[O_2]H \sec \chi \times 10^{-20}\}$$

where H is the local scale height, $q(O_2^+)$, based upon the ionization of $O_2(^1\Delta)$, as given by a formula published by Paulsen et al., (1972) and an ionization rate of 10^{-17}[M] generated by cosmic rays. The latter rate is segregated into 90% $q(O_2^+)$ and 10% $q(NO^+)$ (Swider, 1969). Ionization created by the absorption of solar X rays has been omitted. Hence, the electron concentrations calculated above about 84 km may be somewhat low.

Computations were performed in 2 km intervals from 40 to 90 km. The calculations proceed from initial guesses for the [e] and [SP], total positive ion, profiles. Individual negative ion concentrations, their sum, individual positive ion concentrations, their sum, and finally a new [e] are determined. Typically five or less iterations are performed to reach the criteria that new [e] agree with the previous [e] to within one percent. Because of the algebraic nature of the solutions to the program, results can be obtained with only a hand calculator.

DISTURBED D-REGION RESULTS

Before considering the quiet D-region, we briefly review the past success of the model. Figure 1 depicts the excellent agreement that was obtained (Swider et al., 1978) between the computed and the measured electron profile for the earliest result of the 2-5 November, 1969, SPE expedition. Good agreement also was obtained between the computations and the electron profiles measured near noon on the two succeeding days. The results were very encouraging in view of the fact that the negative ion chemistry scheme is quite simple, not involving hydration processes. On the other hand, NO_2^- and NO_3^- are essentially terminal ions and the addition of water molecules to the core ions would not drastically alter the results unless the respective ion-ion recombination coefficients were changed significantly. Laboratory evidence (Smith et al., 1976) indicates that such coefficients depend only very weakly upon the type/mass of the ions interacting.

Good agreement is shown in Figure 1 between the model and the up-leg data. In general, down-leg or descent data tends to be of poorer quality than the ascent data. Photo-detachment processes are among the 83 reactions considered, and although not evident from this figure, the discussion in the original paper (Swider et al., 1978) reports that photodetachment rates twice the normal one-way rates have been adopted because they agree better with the data. Note that CIRA 1972, 60°N, temperature and [M] were used in the computations for the SPE.

An important aspect of the success of the match illustrated in Figure 1 may be attributable to the fact that the excessive electron content of the disturbed D-region provides for superior observational results.

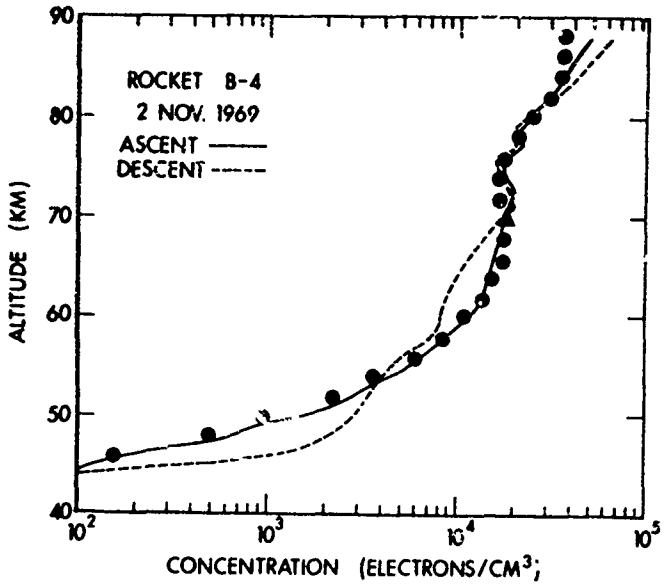


FIGURE 1. Computed steady-state model (dots) compared with SPE data.

QUIET D-REGION RESULTS

Figure 2 illustrates the principal ions determined with the exact model, including [M] and T, used for SPE conditions but without the ionization generated by precipitating protons. The computed electron profile has a distinct sudden change in the electron gradient near 65 km which suggests a separation of the entire profile into a D-region (65-90 km) dominated by the ionization of NO and a C-region (65 km and below) dominated by a cosmic ray ionization source.

The ratio of negative ions to electrons, λ, is unity at 62 km, the same as for the SPE situation since the same minor neutral, T and [M] profiles have been adopted. The enhancement of $[O_2(^1\Delta)]$ and the lower [NO] profiles mentioned two sections ago, which will be utilized later, were instituted because of the fact that in Figure 2 too little $[O_2^+]$ is present near 90 km and the electron profile at 75-85 km, particularly is somewhat higher in concentration than typical data for this altitude region (e.g. Mechtly and Smith, 1968).

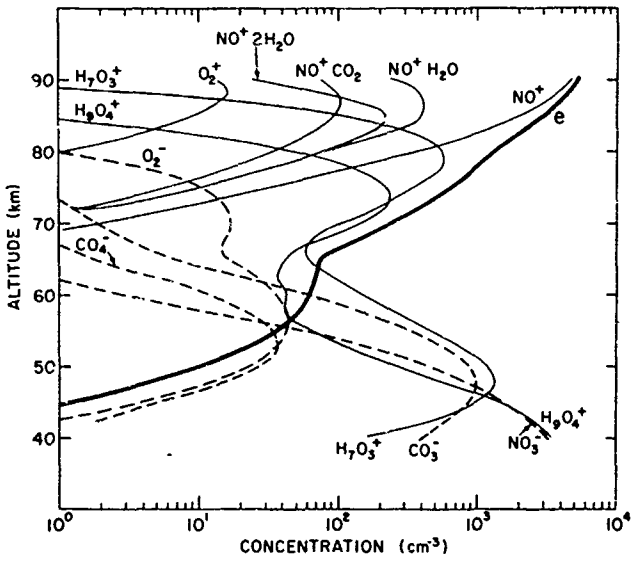


FIGURE 2. Model results for quiet conditions at a 60° solar zenith angle.

The C-layer in Figure 2 is more prominent than those evinced by the aforementioned typical data. On the other hand, ground-based observations of the quiet D-region often include a "C-region." For example, Bain and Harrison (1972) derived an electron profile for summer noon and sunspot maximum conditions with a C-layer below 68 km having a peak electron concentration of about 90 cm^{-3} at 59 km. They found this layer provided best for a fit to measured 16 kHz ground interference patterns. Full wave analysis of 16 kHz data led Deeks (1966) to conclude that there is a (local) peak electron concentration of about $10^2 cm^{-3}$ near 64 km. More recently, using VLF/LF pulse ionosonde data Rasmussen et al. (1980) found support for a C-layer about 6 km thick with estimated electron concentrations of 60-160 cm^{-3} at 63 to 69 km in the daytime (December) ionosphere. There are various complications inherent in determining electron concentrations from ground-based observations and Sternina (1975) claims that [e] cannot be determined explicitly from distant VLF fields.

The situation is much less clear when it comes to in-situ data obtained with rockets, which although probably the best data above 70 km, may suffer from interpretation difficulties and insensitivity problems below 70 km (Thrane, 1974). There are data (Bowhill and Smith, 1966) documenting the development of a distinct C-layer near 65 km as the sun rises. However, this layer generally appears to vanish with respect to daytime data obtained with rockets (Mechtly and Smith, 1968). There are exceptions even within the large U. Illinois data base. An early result from a rocket (Nike Apache 14.143) launched from Wallops Island, Virginia, on April 16, 1964, at a solar zenith angle of 60° evinces a clear C-layer with a peak concentration of about 60 electrons cm^{-3} at 64 km and with 10 cm^{-3} at 55 km (Mechtly et al., 1967). A so-called "anomalous day" profile and a summer day profile for the same location and solar zenith angle contain a C-layer similar to the model in Figure 2 except that the data is about a factor of two lower in concentration for the anomalous case and somewhat worse for the other case (Sechrist et al., 1969).

York (1979) has recently re-examined some D-region data obtained with rocket borne blunt probes (U Illinois results are derived from Faraday rotation and differential absorption). Figure 3 is taken from York's (1970) third figure. The circles represent electron concentrations reported by Mechtly (1974) which are determined from differential absorption below 83 km. The curves marked with squares and triangles plus solid dots correspond to blunt probe current data interpreted, respectively, in terms of a flux limit and certain mean free path limits. The point we wish to stress is that both interpretations lead to a C-layer structure similar to that provided by the model, the solid curve, which was shown also in Figure 2.

One feature of the model which may require further revision is the low total ion concentration near 65 km, about 70 cm^{-3} (Figure 2). Eddy transport, if incorporated into the model, might tend to fill in this valley. However, the eddy diffusion time constant, $\tau = H^2/2D$, where H is the local scale height ($\sim$ 6 km) and D the eddy diffusion coefficient, $\sim 3 \times 10^5$ cm^2/s, has a value of about 6×10^5 s. This time appears to be much longer than the positive ion lifetime, about one hour for this altitude, but the negative ion lifetimes are as great as 1.7×10^5 s.

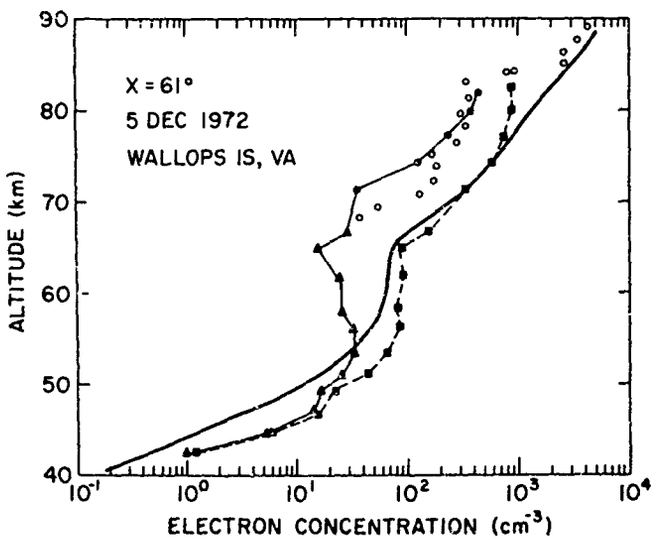


FIGURE 3. Electron distribution model (solid curve) compared to data (see text).

Three different D-region profiles and their related nitric oxide profiles are exhibited in Figure 4. The largest [NO] curve was based upon the analysis of the 2-5 November 1969 SPE observations and provides for the largest [e] profile shown, which is precisely the profile illustrated in the previous two figures. The middle [NO] and [e] distributions represent an attempt to match better the rocket data (e.g. Mechtly and Smith, 1968) while the lowest [NO] profile yields an electron profile in even better agreement with the rocket-associated data which shows an average [e] of only 200-300 cm^{-3} for a solar zenith angle $\chi = 60°$ at Wallops Island, Virginia.

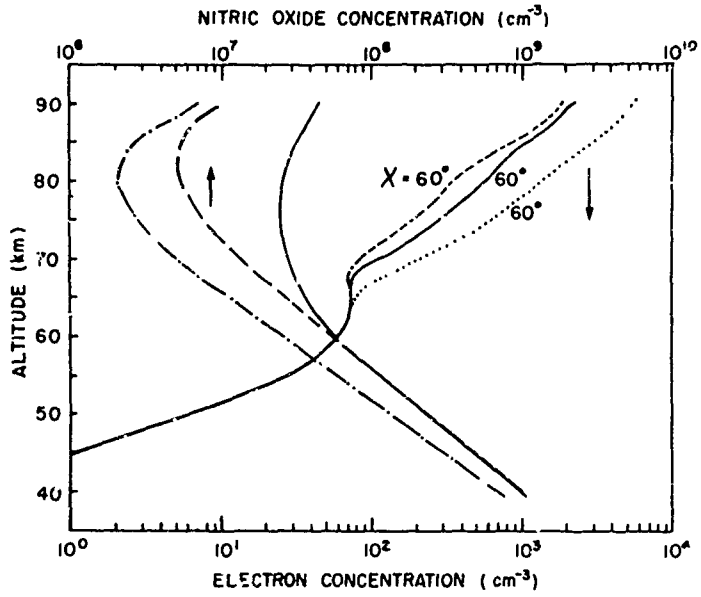


FIGURE 4. Electron profiles computed for a 60° solar zenith angle using three different [NO] profiles.

The lowest [NO] distribution is taken from a study by Swider (1980) and corresponds to a rough best fit of the nitric oxide profiles obtained through the use of rockets in the mesosphere. This distribution was weighted heavily near 80 km in terms of the lowest concentrations observed. Interestingly enough, this curve is quite similar in shape to an [NO] distribution determined theoretically by Ogawa and Shimazaki (1975) except that their profile lies even lower, by about a factor of two.

The natural variation of nitric oxide with latitude, season and solar activity cannot be overemphasized. Nitric oxide concentrations derived for 105 km from the UV data of Atmosphere Explorer satellite C have been found (Cravens and Stewart, 1978) to be 2-3 times larger and more variable at high latitudes (up to 68° geographic) as compared to equatorial latitudes. During strong magnetic storms, [NO] can be enhanced even at the equator. The first nitric oxide band absorption measurements obtained with a satellite yield a factor of 4 sunrise to sunset variation in [NO] at 85 km for $\pm$30° latitude compared to a predicted value (Massie, 1980) of 1.9.

In Figure 5, the median profile given in Figure 4 is reproduced. In addition, electronic distributions are given for solar zenith angles of 0° and 75.5° (sec χ = 4) using the identical nitric oxide concentrations utilized in computing the 60° case. This figure demonstrates how the C-layer tends to vanish with high solar zenith angles as $q(NO^+)$ increases prominently in the region above 60 km as the sun rises higher in the sky. Vertical, one-way non-deviative linear absorption at 3 MHz is 5.6, 4.4 and 3.7 dB for χ = 0°, 60° and 75.5°, respectively.

Another variation that could be, but has not been, introduced is that associated with variations in the cosmic ray ionization source. The ionization rate selected, $10^{-17}[M]$, is appropriate for mid-latitude conditions, but the actual local rate, which is mainly a function of latitude can be one-fifth lower at the equator for high solar activity and twice as great at the highest latitudes for low solar activity (Swider, 1969). These ionization rates would alter the electron concentration at 65 km by about a factor of 1/ 5 and 2, respectively, i.e., to about 30 cm^{-3} and 10^2 cm^{-3}, since electrons are lost primarily by dissociative recombination at this height, where λ = 1. Thus, the formation of the C-layer is more likely at mid-to high-latitudes where not only

the solar zenith angle is generally higher but also where the ionization rate due to cosmic rays is greatest. However, the presence of a high latitude C-layer may be eliminated by a sufficient number of precipitating particles creating ionization in the 65-80 km region.

In the section, "Quiet D-region Model", an altitude of 62 km was cited for the case where λ, the ratio of negative ions to electrons, is unity. However, this result was obtained from calculations which were associated with the SPE, 60°N CIRA 1972 densities and temperatures. The application of the U.S. Standard Atmosphere densities and temperatures for subsequent computations using the middle and low [NO] distributions resulted in an altitude of 65 km for unity λ. (This is mainly related to the different [M].) Recent determinations of this altitude from Thomson scatter spectra yeild an altitude of about 69$\pm$1 km for summer noon at Arecibo, Puerto Rico (Ganguly et al., 1979).

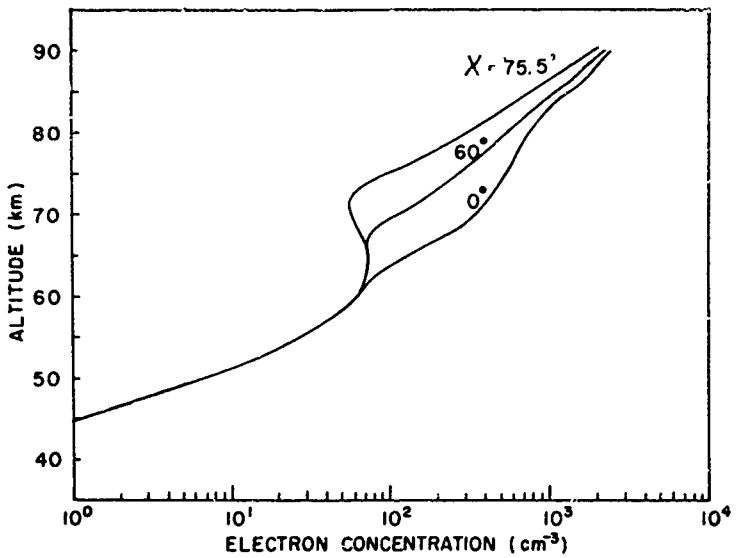


FIGURE 5. Electron profiles computed for several solar zenith angles.

NIGHTTIME CONDITIONS

In general, there is basically no D-region at night except under sufficiently disturbed conditions. For the quiet D-region at night only ionization generated by cosmic rates is present. (Some scattered HLy α and HLy β radiations are present in the E-region.) The absence of photoprocesses results in a rise of the altitude for unity λ to about 80 km at night (Ganguly et al., 1979) with a similar altitude under SPE conditions, 78$\pm$3 km (Swider et al., 1971). The altitude of unity λ is strongly related to the region where atomic oxygen virtually vanishes at night. Atomic oxygen inhibits negative ion formation by destroying the initial negative ion made in the D-region. The process is $O_2^- + O \rightarrow O_3 + e$ and $O^- + O_2$ and although each path is approximately equal, $O^- + O \rightarrow O_2 + e$. Atomic oxygen typically has a large concentration, $\gtrsim 10^{11}$ cm^{-3} near 90 km by night and day, but at night the concentration drops to $\gtrsim 10^9$ cm^{-3} at 80 km and less than 10^5 cm^{-3} at 75 km. Where atomic oxygen is negligible at night, negative ion initiation and subsequent ionic processes proceed much more rapidly than in the daytime towards the formation of highly stable, clustered, ions. For the quiet nighttime D-region, only a few electrons exist at 80 km (Knight, 1972) with virtually none below this altitude.

SUMMARY

We have presented a steady-state model of the D-region which has been successful in describing data from the 2-5 November, 1969, SPE (Swider et al., 1978). We have shown here that this model may be suitable for the quiet D-region also and that a C-region may be present at times. The latter region is most likely when ionization at altitudes of 65-80 km is low as, for example, at high solar zenith angles (high latitudes thereby favored) with low or moderate NO levels and, of course, no significant precipitation by energetic (ionizing) particles. Ionization generated by cosmic rays, the source of the C-region, is greater at high latitudes.

Comparison of our model with experimental data is fair, with observations obtained through use of rockets being less likely to evince a C-l than ground-based VLF/LF measurements. However, there may be sufficient uncertainty in the D-region data base, which we estimate must exceed at least 100 published profiles, to date, to bring into question the precision of the data, particularly near 55-65 km, with respect to the results obtained with the model.

We find that the model predicts few electrons at night for the D-region which appears to be in accord with the available information. The altitude of the ratio of negative ions to electrons, λ, rises, for a value of unity, from about 65 km by day to about 80 km at night.

REFERENCES

Bain, W.C., and M.D. Harrison (1972) Model ionosphere for D region at summer noon during sunspot maximum, Proc. Inst. Elec. Eng., 119, 790-796.

Bowhill, S.A., and L.G. Smith (1966) Rocket observations of the lowest ionosphere at sunrise and sunset, Space Res., 6, 511-521.

Cravens, T.E., and A.I. Stewart (1978) Global morphology of nitric oxide in the lower E-region, J. Geophys. Res., 83, 2446-2452.

Deeks, D.G. (1966) D-region electron distributions in middle latitudes deduced from the reflexion of long radio waves, Proc. Roy. Soc. A, 291, 413-437.

Ganguly, S., J.D. Mathews, and C.A. Tepley (1979) Thomson scatter radar detection of D-region negative ions at Arecibo, Geophys. Res. Lett., 6, 89-92.

Knight, P. (1972) A classification of night-time electron-density profiles, J. Atmos. Terr. Phys., 34, 401-410.

Massie, S.T. (1980) Nitric oxide delta band absorption measurements in the lower thermosphere, J. Geophys. Res., 85, 2155-2164.

Mechtly, E.A. (1974) Accuracy of rocket measurements of lower ionosphere electron concentrations, Radio Sci., 9, 373-378.

Mechtly, E.A., S.A. Bowhill, L.G. Smith, and H.W. Knoebel (1967) Low ionosphere electron concentration and collision frequency from rocket measurements of Faraday rotation, differential absorption, and probe current, J. Geophys. Res., 72, 5239-5245.

Mechtly, E.A. and L.G. Smith (1968) Seasonal variation of the lower ionosphere at Wallops Island during the IQSY, J. Atmos. Terr. Phys., 30, 1555-1561.

Ogawa, T., and T. Shimazaki (1975) Diurnal variations of odd nitrogen and ionic densities in the mesosphere and lower thermosphere: simultaneous solution of photochemical-diffusive equations, J. Geophys. Res., 80, 3945-3960.

Paulsen, D.E., R.E. Huffman, and J.C. Larrabee (1972) Improved photoionization rates of O ($^1\Delta_g$) in the D-region, Radio Sci., 7, 51-55.

Rasmussen, J.E., P.A. Kossey, and E.A. Lewis (1980) Evidence of an ionospheric reflecting layer below the classical D region, J. Geophys. Res., 85, 3037-3044.

Reid, G.C. (1977) The production of water-cluster positive ions in the quiet daytime D region, Planet. Space Sci., 25, 275-290.

Sechrist, C.G., Jr., E.A. Mechtly and J.S. Shirke (1969) Coordinated rocket measurements on the D-region winter anomaly - I. Experimental results, J. Atmos. Terr. Phys., 31, 145-153.

Smith, D., N.G. Adams, and M.J. Church (1976) Mutual neutralization rates of ionospherically important ions, Planet. Space Sci., 24, 697-703.

Sternina, I.M. (1975) Impedance model of the daytime ionosphere for an electron density profile of arbitrary shape, Geomag. Aeron., 15, 507-509.

Swider, W. (1969) Ionization rates due to the attenuation of 1-100 A nonflare solar X rays in the terrestrial atmosphere, Rev. Geophys., 7, 573-594.

Swider, W. (1980) Nitric oxide in the upper stratosphere and mesosphere, pp. 263-272 in Proc. NATO Adv. Study Inst Atmospheric Ozone, FAA-EE-80-20.

Swider, W., R.S. Narcisi, T.J. Keneshea, and J.C. Ulwick (1971) Electron loss during a nighttime PCA event, J. Geophys. Res., 76, 4691-4694.

Swider, W., and C.I. Foley (1978) Steady-state multi-ion disturbed D-region model, AFGL-TR-78-0155.

Swider, W., T.J. Keneshea and C.I. Foley (1978) An SPE-disturbed D-region model, Planet. Space Sci., 26, 883-892.

Thrane, E.V. (1974) Ionospheric profiles up to 160 km: a review of techniques and profiles, Space Res., 14, 1-21.

York, T.M. (1979) Measurement of electron densities in the middle atmosphere using rocket borne blunt probes, ASL-CP-79-0100-5.

THE APPLICATION OF AN IONOSPHERIC MODEL FOR PREDICTING RADAR-PROPAGATION EFFECTS

GEORGE H. MILLMAN
GENERAL ELECTRIC COMPANY
SYRACUSE, NEW YORK 13221

ABSTRACT

The Penn State Mark I ionospheric model is employed for the prediction of propagation effects such as angular bending, range-time delay and Faraday polarization rotation on radar performance. A suggestion for improvements in ionospheric modeling is presented.

INTRODUCTION

The performance of HF radar systems that are designed to detect targets at over-the-horizon distances are being evaluated utilizing the ITS-78 (Barghausen et al., 1969) and the IONCAP (Lloyd et al., 1978) ionospheric models. These models, which are in the form of computer programs originally developed by the Institute for Telecommunication Sciences for HF communication systems, are not readily applicable for the evaluation of the performance of VHF and UHF radar systems.

In the design of VHF and UHF radar systems for detecting and tracking earth's satellites and ballistic missiles, it is necessary to take into account the detrimental effects of both the troposphere and the ionosphere since both media can impose limitations on system capability and performance. In this paper, only the effects of the ionospheric medium on radar propagation are considered.

The analysis of propagation effects on transionspheric radar systems can be accomplished by employing ionospheric models such as those developed by Nisbet (1970), Damon and Hartranft (1970), Bent et al. (1972) and Ching and Chiu (1973).

In this paper, Nisbet's model, i.e., Penn State Mark I ionospheric model, is used in the estimation of the range and elevation angle bias errors and the angular polarization rotation that could confront a radar located in the midlatitude. An evaluation is made of the ionospheric propagation effects as a function of time, season, sunspot number and various azimuth-elevation angle orientations.

MATHEMATICAL FORMULATION

Range Error

Time delays or range errors are always inherent in the measurement of positional data of space vehicles by radio waves. This is due to the fact that the velocity of electromagnetic propagation in the ionosphere is slightly less than the free space velocity. In other words, the presence of the ionosphere introduces an increase in the effective group path length relative to free space.

The range error, ΔR, is the difference between the apparent range, R_a, and the geometric range, R, or

$$\Delta R = R_a - R \tag{1}$$

Utilizing Snell's law for a spherically symmetrical surface, it can be readily shown (Millman, 1965) from Figure 1 that

$$R_a = \frac{1}{n_o r_o \cos E_o} \int_{r_o}^{r} r \left[\left(\frac{n r}{n_o r_o \cos E_o} \right)^2 - 1 \right]^{1/2} dr \tag{2}$$

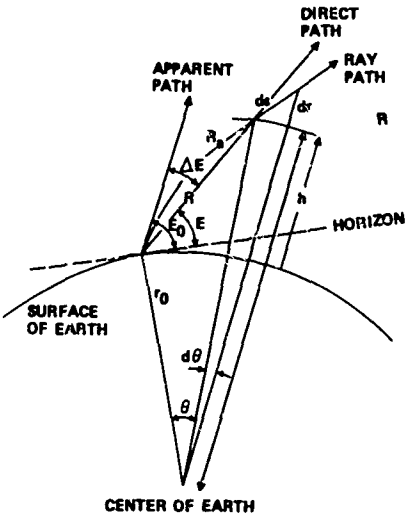


Figure 1. Ray Path Geometry for a Spherical Earth

and

$$R = \left[r_o^2 + (r_o + h)^2 - 2r_o (r_o + h) \cos \theta \right]^{1/2} \tag{3}$$

where r_o is the radius of the earth, n_o and n are the indices of refraction at the earth's surface and at the altitude, h, respectively and E_o is the apparent elevation angle of the ray at the earth's surface.

The central angle, θ, is given by

$$\theta = \int_{r_o}^{r} \frac{1}{r} \left[\left(\frac{n\, r}{n_o\, r_o \cos E_o} \right)^2 - 1 \right]^{-1/2} dr \tag{4}$$

and the radial distance, r, is merely

$$r = r_o + h \tag{5}$$

The index of refraction in the ionosphere is defined by

$$n = \left[1 - \frac{N_e\, e^2}{4\pi^2\, \epsilon_o\, m_e\, f^2} \right]^{1/2} \tag{6}$$

where N_e is the electron density (electrons/m^3), e is the electron charge (1.6×10^{-19} C). ϵ_o is the electric permittivity of free space ($10^{-9}/36\pi$ F/m), m_e is the electron mass (9.1×10^{-31} kg) and f is the frequency of the incident wave (Hz).

It is noted that the ionospheric refractive index is also a function of both the electron collision frequency and the earth's magnetic field. For frequencies on the order of 10 MHz and above and, at altitudes greater than 80 km, the effect of the collision frequency term on the index of refraction is negligible (Davies, 1965). The refractive index, as defined by Equation (6), is applicable for estimating the magnitude of ionospheric propagation anomalies such as range error and angular-refractive bending.

It can be shown that, for frequencies in the VHF range and above, the ionospheric range error can also be expressed by the relationship

$$\Delta R = \frac{e^2}{8\pi^2 \epsilon_o m_e f^2} \int_0^s N_e \, ds \tag{7}$$

where ds is the differential distance in the direction of propagation.

It is evident that the range error is inversely proportional to the square of the frequency and directly proportional to the integrated electron density, i.e. electron content, along the transmission path.

Refractive Bending

When radio waves traverse the ionosphere, they undergo a change in direction or refractive bending. This phenomenon, which results from the nonhomogeneous characteristics of the medium, introduces an error in the measurement of the angular position of a space vehicle.

The elevation angle error, ΔE, can be evaluated from the expression

$$\Delta E = E_o - E \tag{8}$$

where E, the true elevation angle, is given by

$$E = \cos^{-1}\left[\frac{r_o + h}{R} \sin\theta\right] \tag{9}$$

Faraday Polarization Rotation

A linearly polarized wave when propagating through the ionosphere separates into two components, both in the general case elliptically polarized with opposite senses of rotation. However, for frequencies in the VHF range and above, the waves are circularly polarized. Since the two waves travel with different velocities of propagation, their phase relationships are continuously being altered. On emerging from the ionosphere, the waves recombine to form a linearly polarized wave whose plane of polarization has been rotated with respect to that of the incident wave. This phenomenon is the result of the interaction of the electromagnetic wave with the electrons in the presence of the earth's magnetic field.

The amount of angular rotation, Ω (rad), experienced by a linearly polarized wave traversing a one-way path in the ionosphere can be represented by the function (Millman, 1974)

$$\Omega = \frac{e^3}{8\pi^2 c \epsilon_o m_e^2 f^2} \int_0^s N_e B \cos\phi \, ds \tag{10}$$

where c is the free space velocity (3×10^8 m/s), B is the magnetic induction (Wb/m^2) and ϕ is the propagation angle, i.e., the angle between the magnetic field vector and the direction of propagation. This expression is valid for the quasi-longitudinal mode of propagation.

In the computation of the polarization rotation, the earth's magnetic field was approximated by a series of spherical harmonics (Chapman and Bartels, 1940). The set of 80 spherical harmonic coefficients for epoch 1965 derived by the International Association of Geomagnetism and Aeronomy (IAGA, 1969) was used to specify the magnetic potential function.

According to Equation (10), the angular rotation is inversely proportional to the frequency squared and directly proportional to the integrated product of the electron density and the component of the earth's magnetic field in the direction of propagation.

DISCUSSION

The Penn State Mark I Ionospheric Model which was used to estimate the range and elevation angle bias errors and the magnitude of the Faraday polarization rotation was developed by Nisbet (1970). The model is in the form of a computer program that describes the electron density profile between 120 and 1250 km altitude as a function of geographic coordinates, day number, time of day, and solar activity. Although the ionosphere is dynamic in nature, a static-equilibrium model is employed. The program is reasonably accurate for the mid-latitudes although it can be used with discretion to model the ionosphere at other locations.

The electron density profiles are basically controlled by the local production and loss of ionization. This assumption is reasonable up to an altitude of approximately 180 km. Above this altitude, the validity of the assumption becomes questionable until about two scale heights above the maximum ionization level where the electron density is, for the most part, controlled by diffusion and gravity.

The model was used to generate, for a sunspot number (SSN) of 10 and 60 and for the months of June and December, the hourly values of electron density profiles at an assumed geographic coordinate 38°N, 40°E, which corresponds to a geomagnetic latitude of approximately 35°N, and for three locations north of the site and three locations south of the site, each separated by a 10° geographic latitude.

The data for the six additional locations yielded latitudinal-electron density coverage and information on the existence of north-south electron density gradients. Since the hourly values of the electron density profiles can be coverted directly to 15° longitudinal values, the east-west gradients were also available.

For specified antenna beam orientations, i.e., azimuth angles of 0°, 90°, 180°, and 270° and elevation angles of 2°, 5°, 10°, 20°, 30°, 40° and 60°, the electron densities along the ray paths were determined by linear interpolation within the electron density grid.

The vertical distributions of electron density at approximately local noon along the 40°E meridian for the months of June and December, and for a SSN of 60 are plotted in Figures 2 and 3, respectively.

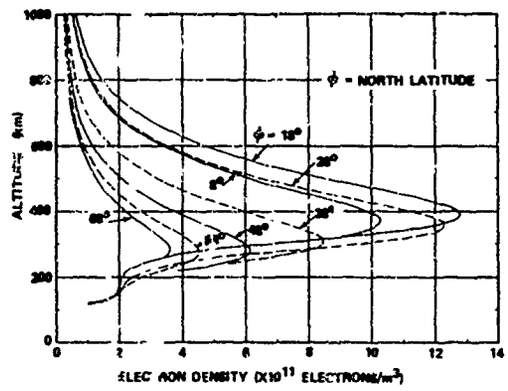


Figure 2. Vertical Distribution of Electron Density at Approximately Local Noon Along 40°E Meridian for June and Sunspot Number of 60

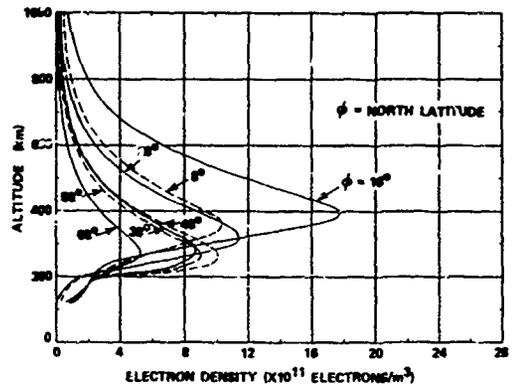


Figure 3. Vertical Distribution of Electron Density at Approximately Local Noon Along 40°E Meridian for December and Sunspot Number of 60

It is apparent that the magnitude and height of the maximum ionization, i.e., F-layer peak, is latitude dependent, the minimum electron density and layer height existing in the northern latitudes.

The maximum electron density which appears at 18°N latitude can be readily explained. Studies of the behavior of the ionosphere have revealed that the geographic variations of the F-layer electron density are geomagnetically controlled. That is, when the sun is over the geomagnetic equator, the electron density of the F-layer peak is found to maximize at approximately ±10° geomagnetic latitude. This is in the vicinity of the geographic coordinates 18°N, 40°E, with a trough located near the geomagnetic equator.

Figures 4 and 5 illustrate the diurnal variations of the vertical distribution of electron density for June and December at a SSN of 60, respectively. It is seen that the electron density reaches a maximum at local noon in June while, in December, it occurs at 1000-h local time. This asymmetry in the electron density-diurnal curve is a typical characteristic of the ionosphere. It is known that, depending on the geographic location, season and epoch of the sunspot cycle, the diurnal maximum of the F-layer electron density can occur practically at any time of the day.

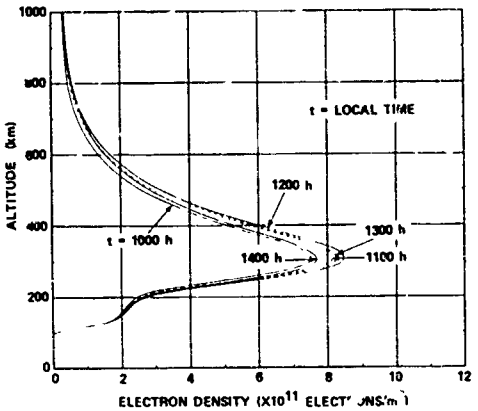


Figure 4. Diurnal Variation of Vertical Distribution of Electron Density at Geographic Coordinates of 38°N, 40°E for June and a Sunspot Number of 60

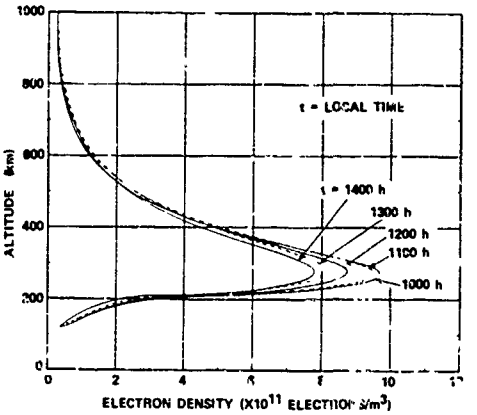


Figure 5. Diurnal Variation of Vertical Distribution of Electron Density at Geographic Coordinates of 38°N, 40°E for December and a Sunspot Number of 60

The seasonal variation of the electron density distribution at the radar site for a SSN of 10 and 60 is presented in Figures 6 and 7, respectively. The June and December noontime values are comparatively the same while the nighttime values show a marked difference. It is noted that the electron density is a function of solar activity, increasing with increasing sunspot number. Evidence of the winter anomaly phenomenon, i.e., the daytime value of electron density is higher in the winter than in the summer, is shown in Figure 7 and also in Figures 4 and 5. The winter anomaly is a characteristic of only the daytime; the nighttime electron density values are always higher in the summer than in the winter.

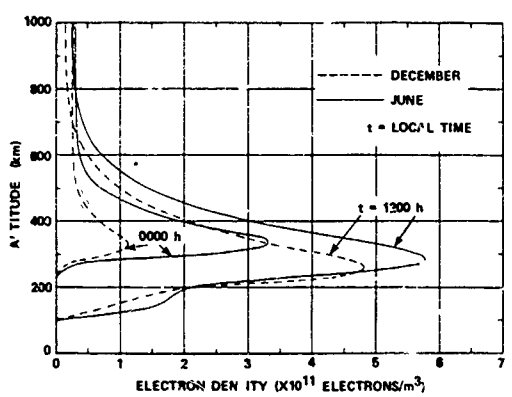


Figure 6. Seasonal Variation of Vertical Distribution of Electron Density at Geographic Coordinates of 38°N, 40°E for a Sunspot Number of 10

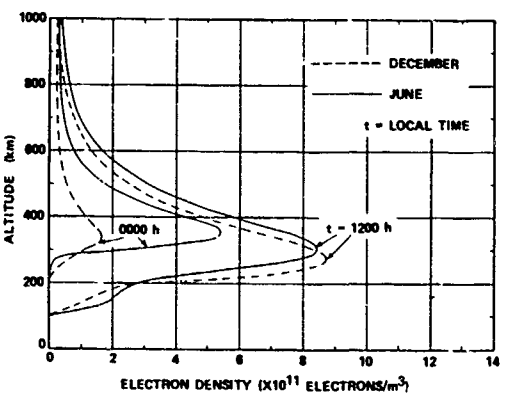


Figure 7. Seasonal Variation of Vertical Distribution of Electron Density at Geographic Coordinates of 38°N, 40°E for a Sunspot Number of 60

Theoretical estimates of the range errors at a frequency of 425 MHz that could be encountered in the sourtherly direction at a SSN of 60, at local noon and in the months of June and December, are presented in Figures 8 and 9, respectively. It is evident that, at 10° elevation angle, the maximum bias range error is on the order of 691 ft in June while it increases to about 788 ft in December.

It is found that, at 0° azimuth, the maximum range errors evaluate to 353 ft and 485 ft for June and December, respectively. In addition, the range errors are reduced by a factor of approximately one-half to one-fourth during the nighttime. For a SSN of 10, the range errors are about one-half the magnitude calculated for a SSN of 60. It is possible that the bias range errors at a SSN of about 110 could be a factor of 2 greater than the sunspot 60 results.

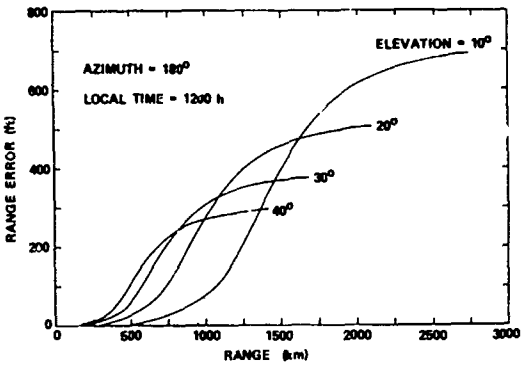


Figure 8. Range Error as a Function of Range for a Sunspot Number of 60 in June

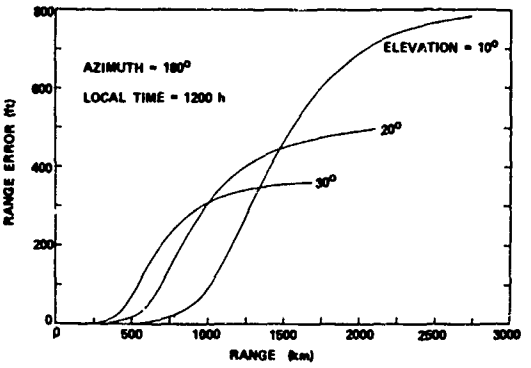


Figure 9. Range Error as a Function of Range for a Sunspot Number of 60 in December

Examples of the ionospheric elevation angle errors at 425 MHz are plotted in Figures 10 and 11 as a function of range. While the range error increases monotonically with distance, the elevation angle error attains a maximum at a range which is a function of the angle of elevation and the ionization distribution along the propagation path. The error maximizes at altitudes above the peak of the F-layer. For the June data, Figure 10, the path altitude at which the error reaches a maximum is approximately 700 km at 10° elevation angle error, 570 km at 20°, 540 km at 30° and 500 km at 40°.

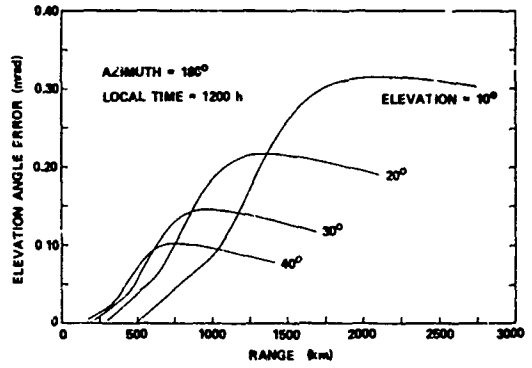


Figure 10. Elevation Angle Error as a Function of Range for a Sunspot Number of 60 in June

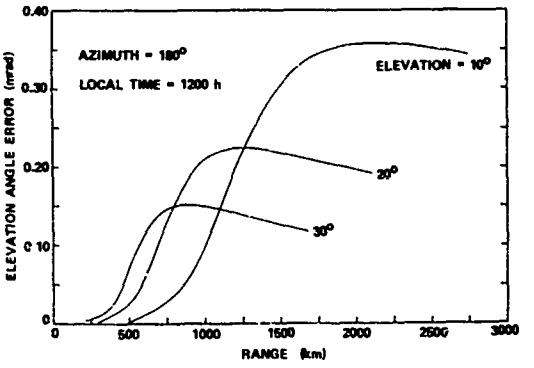


Figure 11. Elevation Angle Error as a Function of Range for a Sunspot Number of 60 in December

According to Figures 10 and 11, the maximum error at 10° elevation angle and 180° azimuth is on the order of 0.35 mrad while calculations show that the corresponding maximum error in the northerly direction is only 0.27 mrad.

For similar azimuth-elevation angle conditions and seasons, it is found that the maximum elevation angle errors at a SSN of 10 are about one-half of that at a SSN of 60. The maximum errors at 0000-h local time and a SSN of 60 in June are less than approximately 0.10 mrad.

Theoretical estimates of the Faraday polarization rotation that could be encountered in the summer at a SSN of 60 are depicted in Figure 12. The 270° azimuth angle calculations are not included since they are on the same order of magnitude as the 90° azimuth data. The interesting feacure to note in Figure 12 is the reversal in the slope of the polarization rotation vs altitude curve at 0° azimuth. The reversal point corresponds to the altitude at which the ray path experiences perpendicularity with the lines of force of the earth's magnetic field. It is seen that the angular rotation is a maximum in the southerly direction and a minimum in the northerly direction. The nighttime values are approximately a factor of 2 to 3 less than the daytime results while the sunspot 10 data are about a factor of 1.5 to 2 less than the sunspot 60 predictions.

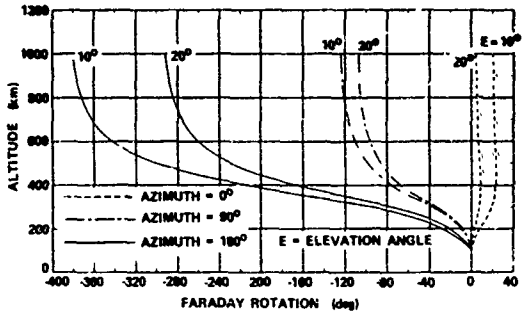


Figure 12. Theoretical Estimate of Ionospheric Faraday Polarization Rotation at Local Noon for June and Sunspot Number of 60

CONCLUSION

Ionospheric models which are dynamic in nature, i.e., electron density specified as a function of time, space and solar-geophysical conditions, can be applied to the prediction of propagation effects on transionospheric radar systems.

The range and elevation angle errors and angular polarization rotation that can be deduced reveal for the most part only estimates of average ionospheric conditions. What is required by the system analyst is an indication of the accuracy of the predictions. The statistical parameters that should also be provided by an ionospheric model are the standard deviation and the upper and lower decile values of the error estimates.

REFERENCES

Barghausen, A.F., J.W. Finney, L.L. Proctor and L.D. Schultz, "Predicting Long-Term Operational Parameters of High-Frequency Sky-Wave Telecommunication Systems", Environmental Sciences Services Administration Technical Report ERL 110-ITS 78, May 1969.

Bent, R.B., S.K. Llewellyn and M.K. Walloch, "Description and Evaluation of the Bent Ionospheric Models", DBA Systems, SAMSO TR-72-239, SAMSO Contract No. F04701-72-C-0380, October 1972.

Chapman, S. and J. Bartels, Geomagnetism, Vol. 1 and 2, Oxford University Press, New York, 1940.

Ching, B.K. and Y.T. Chiu, "A Phenomenological Model of Global Ionospheric Electron Density in the E-, F1- and F2-Regions", Journal of Atmospheric and Terrestrial Physics, Vol. 35, pp 1615-1630, 1973.

Damon, T.D. and F.R. Hartranft, "Ionospheric Electron Density Profile Model", Aerospace Environmental Support Center Technical Memorandum 70-3, July 1970.

Davies, K., "Ionospheric Radio Propagation", National Bureau of Standards Monograph 80, U.S. Government Printing Office, Washington, DC, 1965.

Lloyd, J.L., G.W. Haydon, D.L. Lucas and L.R. Teters, "Estimating the Performance of Telecommunications Systems Using the Ionospheric Transmission Channel", Institute for Telecommunication Sciences, Unpublished Report, September 1978.

IAGA Commission 2 Working Group 4, Analysis of the Geomagnetic Field, "International Geomagnetic Reference Field 1965.0", Journal of Geophysical Research, Vol. 74, pp 4407-4408, August 15, 1969.

Millman, G.H., "Atmospheric Effects on Radio Wave Propagation", in Modern Radar, Analysis, Evaluation and System Design, edited by R.S. Berkowitz, pp 317-377, John Wiley & Sons, Inc. New York, 1965.

Millman, G.H. and G.M. Reinsmith, "An Analysis of the Incoherent Scatter - Faraday Rotation Technique for Ionospheric Propagation Error Correction", General Electric Technical Information Series Report No. R74EMH2, February 1974.

Nisbet, J.S., "On the Construction and Use of the Penn State Mk I Ionospheric Model", Pennsylvania State University, Ionosphere Research Laboratory Scientific Report No. 355, May 1970.

ANOMALOUS IONOSPHERIC REFRACTION ASSOCIATED WITH THE AURORAL ZONE

J. V. Evans* and R. H. Wand+
*M.I.T. Lincoln Laboratory and +NEROC Haystack Observatory
Lexington, Massachusetts 02173 and Westford, Massachusetts 01886

INTRODUCTION

Refraction in the earth's troposphere and ionosphere can limit the precision with which a search radar can locate the position of a satellite or ballistic missile. As illustrated in Figure 1, the bending of the ray path in both regions serves to increase the apparent elevation over the true elevation. The apparent range to the target also is increased.

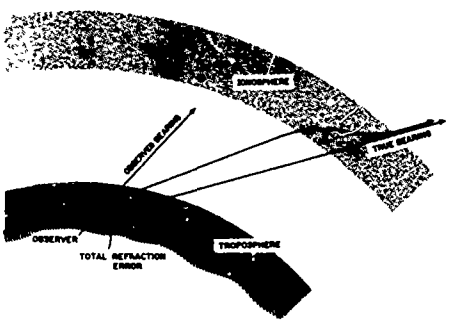


Figure 1. Diagram illustrating the effects of refraction in the troposphere and the ionosphere.

Table 1 gives the magnitude of the two errors at 2° elevation and indicates approximately how they decrease with elevation.

TABLE 1
Atmospheric Propagation Effects (2° apparent elevation)

Peak Bias	Tropospheric	Ionospheric (UHF)
ΔR (m)	$\sim$ 48 $\pm$3	25-500
Δ El (m. deg.)	$\sim$ 390 $\pm$20	2-60
Elevation Dependence	El^{-1}	$\sim El^{-\frac{1}{2}}$

The tropospheric error is independent of frequency (at radio wavelengths) and falls off fairly rapidly with increasing elevation. The ionospheric errors vary as $1/f^2$ (for frequencies greater than about f = 50 MHz) and those quoted in Table 1 are for f = 400 MHz. Ionospheric errors are much more variable and do not decrease rapidly with increasing elevation.

Between February 1969 and July 1973, the Bell Telephone Laboratories and the Lincoln Laboratory conducted a joint study to investigate the limitations imposed by naturally occurring propagation effects on defensive radar systems operating at UHF. The study encompassed a program of field measurements, largely conducted at the Lincoln Laboratory Millstone Hill Field Station (42.6°N, 71.5°W) and an associated program of data interpretation, theoretical analysis and model building conducted at BTL (under the direction of J. H. W. Unger).

The results of the program are applicable to studies of the performance of any radar system operating at VHF or higher frequency in the northern continental United States and, as such, were reported in some detail in various technical reports[1] and papers[2,3]. The experimental program conducted at Millstone Hill is described in three technical reports dealing respectively with the instrumentation employed[4], calibration of the instrumentation[5], and the results[6].

Broadly stated, the initial objectives of the Radar Propagation Study program were to determine if there are anomalous propagation effects associated with the auroral regions of the ionosphere that would degrade the performance of defensive radar systems. It was feared, for example, that severe angular refraction would be encountered for ray paths that traverse auroral arcs. These and other short-lived tracking perturbations were expected to manifest themselves in altering the apparent position of a target at UHF, but would not be seen at L-band (where ionospheric effects are considerably weaker). Accordingly, the experiments were planned to provide a simple comparison of apparent target position at these two frequencies. Since tropospheric refraction is frequency independent in the radiowave band, it was expected that this approach would serve to discriminate against spurious effects that might be introduced in the troposphere or by poor tracking.

As the program progressed, it became evident that the systematic range and angle bias errors introduced by the ionosphere were also large enough to degrade tracking performance significantly, and the scope of the work was broadened to include their study. The measurement of systematic bias errors required an order-of-magnitude improvement in the knowledge of the absolute pointing of the antenna beams at UHF and L-band, and this was achieved only after considerable effort[6].

The targets employed throughout the study were five satellites of the Navy Navigation Series. These were chosen because they carry UHF (400-MHz) beacons permitting the satellite position to be determined using a simple beam-comparison tracking system. In addition, these satellites are in polar orbits and hence may be seen rising or setting to the north through the auroral regions. The tracking was performed by means of an 84-foot diameter parabolic antenna, instrumented to make simultaneous angle-of-arrival measurements of the UHF (beacon) signals and L-band radar returns. The antenna also carried a yagi array to receive a VHF (150-MHz) signal that is a subharmonic (3/8) of the 400-MHz signal. Comparison of these two signals permitted the time rate of change of the ionosphere electron content along the line-of-sight to the satellite to be studied by means of a differential-Doppler experiment.

In this paper, we report observations of an auroral latitude feature that proved to introduce the largest bias refraction effects encountered in the study, namely a trough in the F-region electron density that appears in the daytime on some very magnetically disturbed days. The physical cause of this feature also is discussed.

DIFFERENTIAL DOPPLER MEASUREMENTS

The differential-Doppler measurements made during each pass of the Navy Navigation Satellites provided a measure of the rate of change of the total number of electrons in the ionosphere (up to about 1000 km) as the line-of-sight swept from south to north, or vice versa. These records were useful in showing latitudinal structure in the earth's ionosphere produced, for example, by gravity

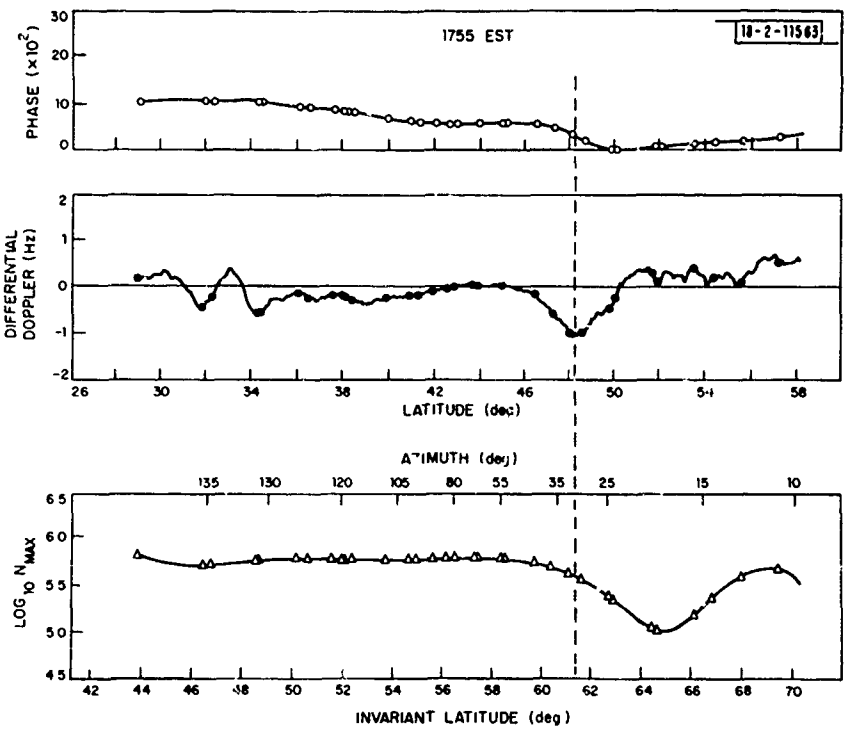


Figure 2. Example of the differential phase Φ, differential Doppler $d\Phi/dt$, and $Log_{10}N_{max}F2$ vs. latitude for a pass observed in November 1972. It is evident that the perturbation in the differential-Doppler record is associated with the equatorward edge of the trough.

waves (Traveling Ionospheric Disturbances) or related to the auroral zone. Figure 2 provides an example of one of the records obtained (center panel). In this case, a TID is clearly evident to the south of the station (near 34° latitude) and a trough to the north (near 48° latitude). By integrating the differential Doppler, the variation of the phase path length caused by the ionosphere can be recovered (top panel of Figure 2). Utilizing local measurements of F-region critical frequency as well as a model of the seasonal and diurnal variation of the ionospheric slab thickness[7], it was possible to recover an estimate of the variation of the peak electron density along the latitude of the subionospheric point[8] (taken to be the latitude where the ray path traverses the ionosphere at 300 km). This is shown in the bottom panel of Figure 2. It is clear that the differential Doppler signature of the trough corresponds to the point were the ray path traverses its equatorial edge.

Troughs in the auroral ionosphere first were recognized in ionograms gathered by the Alouette 1 satellite[9] and their morphology since has been a subject of intense study[10-18]. Using the differential Doppler records gathered at Millstone over a 2-year period, Wand and Evans[3] were able to construct maps of the average peak electron density in the vicinity of the trough for summer and winter for various levels of Kp. Figure 3ab shows two examples of these maps.

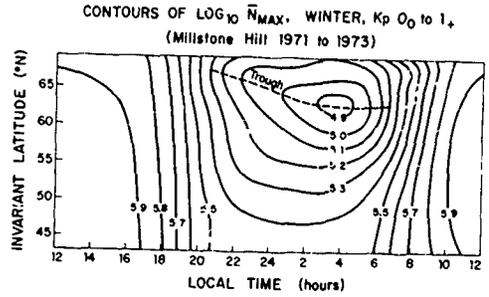


Figure 3a. Contours $\log_{10}N_{max}$, as derived from differential Doppler measurements made at Millstone Hill 1971-1973. Results are for a K_p interval 0_o to 1_+ in winter and the deduced mean position of the electron density trough is shown.

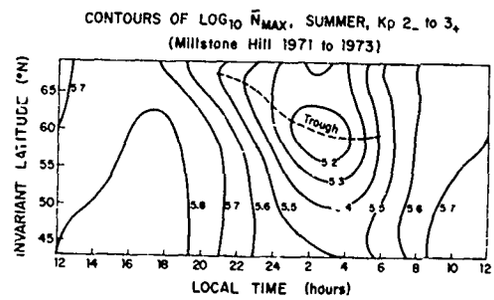


Figure 3b. Contours $\log_{10}N_{max}$, as derived from differential Doppler measurements made at Millstone Hill 1971-1973. Results are for a K_p interval 2_- to 3_+ in summer and the deduced mean position of the electron density trough is shown.

The nighttime troughs were not observed to produce significant ionospheric refraction effects[6] and this result was consistent with ray-tracing studies[19]. An unexpected feature of the observations was the appearance on some magnetically disturbed days of high latitude troughs in the daytime when the F-region density was much higher. Table 2 lists the dates and Kp values for the days when this phenomenon appears to have been detected. Figure 4 gives examples of the differential Doppler records obtained on one of these occasions.

Figure 5 shows the incidence of trough signatures on the differential Doppler records as functions of invariant latitude and GMT (i.e., local time + 5 hours). It can be seen that there is a clear tendency for the trough to move southwards during periods when Kp is high (> 4-) and, moreover, daytime troughs seem confined to winter days (see also Table 2). We are not certain if this is a selection effect, but it did result in there being very large density gradients whenever the phenomenon occurred.

Figure 4 provides an example of successive differential Doppler records gathered late in the afternoon of 28 November 1972, showing the trough signature advancing equatorwards with time. While generally the trough feature remained to the north of the radar on all these days, there were instances when the trough actually passed over the radar site and then was seen to the south as illustrated in Figure 6.

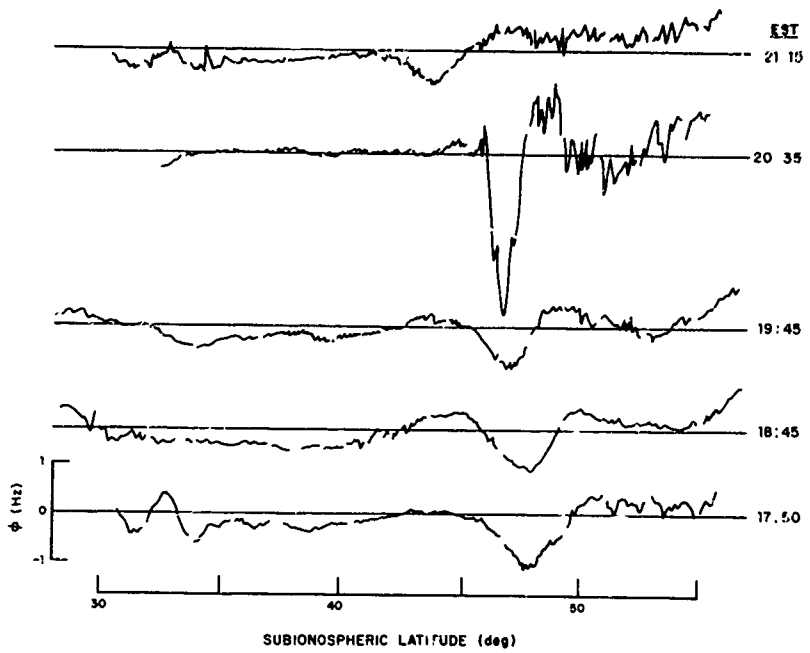


Figure 4. Differential-Doppler curves seen on a sequence of tracks on 28 November 1972. The motion of the trough toward the radar is clearly evident.

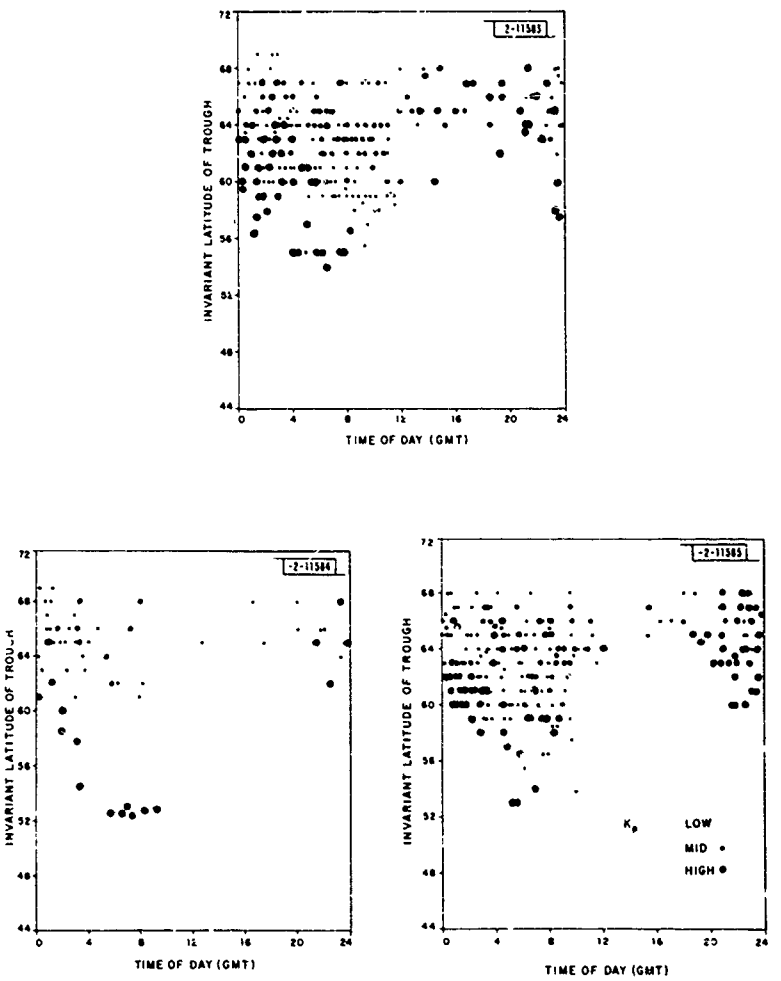


Figure 5. Trough location deduced from the satellite tracking observations vs. time of day for three levels of Kp: (a) winter, (b) summer, (c) equinox.

TABLE 2
Days on which troughs were detected in the daytime.

Date	Kp			
	12-15	15-18	18-21	21-24
17 Dec. 1971	4+	7	7+	5+
12 Oct. 1972	4	5-	5-	5-
28 Nov. 1972	4+	3+	2+	2
1 Feb. 1973	3-	4-	2+	3
21 Feb. 1973	2	5	7+	7-
22 Feb. 1973	4	5+	6+	6

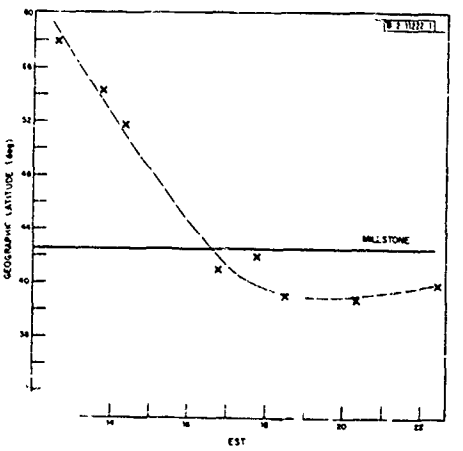


Figure 6. [illegible]ion of the center of the peak Doppler e[illegible]n vs. time on 21 February 1973.

REFRACTION MEASUREMENTS

The measurement of the ionospheric refraction using combined beacon- and skin-tracking proved considerably more difficult than anticipated, owing to the fact that the error null in the beacon-tracker was found to depend upon the polarization of the radiated signal. This changed as the satellite crossed the sky, causing a varying displacement between the apparent nulls in the radar and beacon-tracker monopulse beams[4,5]. To overcome this problem, it proved necessary to employ precise metric radar data gathered for each satellite in many successive passes to solve for its orbit. Unfortunately, this could not be accomplished until an accurate error model had been developed for the radar biases.

Given a precise satellite orbit, the polarization dependence of the UHF beacon-tracker could be calibrated by comparing the ephemeris and apparent positions of the satellite at high elevations at night. The transmitted polarization was determined by measuring the relative amplitude and phase of the right and left circular components of the beacon signal. It then was possible to use this calibration to correct data gathered in the daytime for the effect of the dependence of the error null on received polarization. It should be recognized that the largest effects anticipated (i.e., 60 m. deg. – Table 1) were a small fraction of the 2° receiving beam and hence observing ionospheric refraction proved

extremely difficult to accomplish despite the considerable care employed (which the above description does insufficient justice).

Normal daytime ionospheric refraction usually could not be detected clearly in data for a single pass, owing, in part, to ephemeris errors[4] and, in part, to UHF beacon-tracking errors[5]. However, by averaging many (e.g., 10-20) daytime passes when the ionospheric density was approximately constant (as judged by the local measuremnts of N_{max} F2), the effect could be seen clearly. Figure 7 shows the elevation error vs. elevation obtained by averaging 23 daytime passes. The average value of N_{max} was $7.4 \pm 0.9 \times 10^5$ el/cm^3 and the broken curves represent the expected refraction from ray-tracing studies. The crosses in these plots represent the rms deviations about the mean.

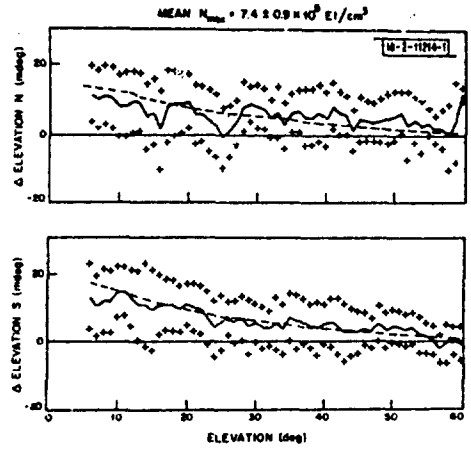


Figure 7. Average residuals between the computed and the observed elevation of the UHF beacon in daytime passes ($f_0F2 > 6.5$ MHz) observed in 23 passes during 13-16 March 1973. Cross marks the rms deviations about the means. Data taken north of Millstone are the top panel, south are the bottom. Broken curve is the refraction computed from ray-tracing studies.

By contrast, the existence of large anomalous refraction when troughs were present in the daytime usually could be inferred from single passes. Figures 8 and 9 provide examples of elevation errors approaching 50 m. deg. that were detected when large latitudinal density gradients were evident (from the differential Doppler records) to the north. While some amplitude scintillation of the beacon signals appears to have been present in each case, this is believed not to have degraded the tracking accuracy[5].

Figure 10 shows elevation residuals for a sequence of passes on 12 October 1972, when the trough appeared to approach the radar during the course of the afternoon. In the pass at 15.45, the trough effects appear to have been masked by the presence of a TID.

PHYSICS OF TROUGH FORMATION

Knudsen[20] appears to have been the first to offer an explanation for the formation of an ionospheric trough on the nightside of the earth. According to this model, magnetic flux tubes are convected from the dayside to the nightside across the polar cap as a result of the impressed magnetospheric electric field. These tubes return to the dayside by subsequently drifting westwards or eastwards. Tubes drifting westwards must travel faster than earth rotation if they are to re-emerge in sunlit regions and some fail to do so, allowing the ionization in them to fall to very low levels as it recombines. In this model, the trough then is the locus of flux tubes which spend a long period in darkness owing to the competing effects of the convection electric field motion and earth rotation. This model has been refined further through numerical studies[21] and a more realistic electric field convection pattern has been proposed by Spiro et al[22] based upon Atmospheric Explorer C data. According to the latter, no special convection electric field

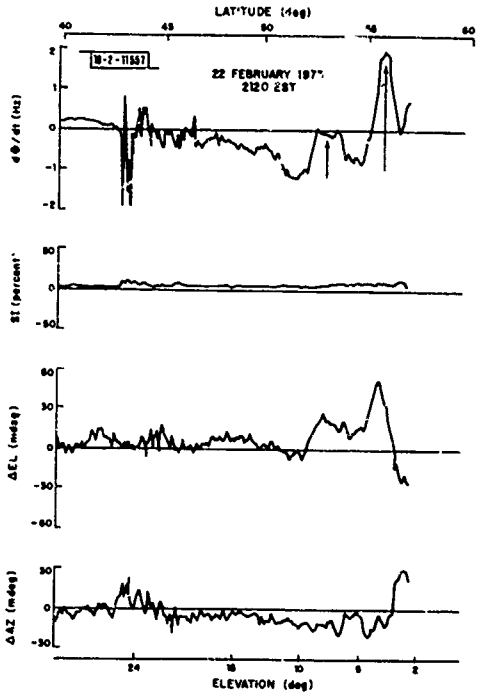


Figure 8. Example of correlated differential-Doppler and refraction error perturbations. The arrows mark the location of Doppler excursions caused by gradients in the high-latitude ionosphere.

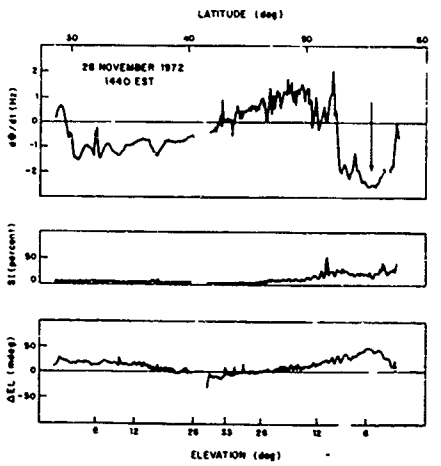


Figure 9. Example of a steep gradient (marked on the differential-Doppler plot by an arrow) and associated anomalously large refraction error.

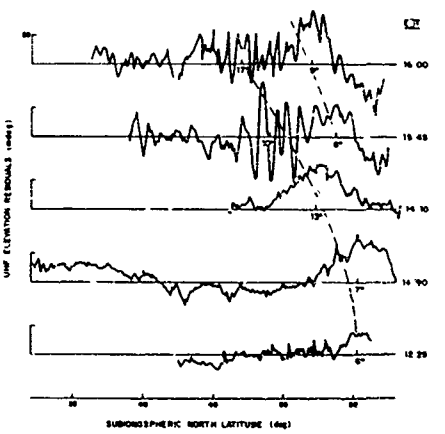


Figure 10. Elevation errors on a sequence of tracks on 12 October 1972. The ticks mark the location of the peak Doppler excursion identified from the differential-Doppler results connected by the broken line. Numbers give the elevation angle for the ray at these points.

pattern is required, only that the electric fields fall reasonably smoothly with magnetic colatitude. There then is always a region on the nightside where convection and corotation exactly balance – creating a stagnation line.

Incoherent scatter studies of the convection electric field recently have been made from Millstone Hill[23,24] which appear to support the explanation offered by Spiro et al[22] for the nighttime trough as it is a region in which the drift velocities appear to be extremely low.

An alternative mechanism for trough formation has been suggested by Banks et al[25] and Schunck et al[26], who have shown that under the influence of a large electric field, O^+ ions will be swept through neutral molecular nitrogen and charge transfer (via $O^+ + N_2 \rightarrow NO^+ + N$) more rapidly than normal. This process creates NO^+ which quickly recombines with free electrons.

Large electric fields (> 100 mV/m) have been observed at subauroral latitudes from satellites[27-29] which could produce the required drift velocities and troughs in electron density have been detected in some cases at satellite altitudes (> 700 km)[29]. These electric fields appear to persist for several hours and usually are seen in the late evening and premidnight sectors. The electric fields are directed northwards and give rise to a westward ion drift and hence would create E-W troughs. Rich et al[29] speculate that these intense fields are generated by the penetration of the partial ring current ions to lower L-shells than the plasma sheet electrons.

We believe that the troughs reported here, which were seen in the afternoon and evening sectors, were produced by this second mechanism. It is difficult to account for the apparent rapid motion in latitude of these troughs by the first mechanism (since it takes several hours for the ionization in a flux tube to decay). A latitudinal change is a natural result if the impressed electric field extends to lower L-shells at later local times,or if the penetration of the ring current increases with time.

Recently, incoherent scatter measurements were made at Millstone Hill on a day (26 February 1979) when a trough formed to the north early in the afternoon and moved to lower latitude later. In the trough, the electron and ion temperatures were found to be very high. The former results from the reduced rate of cooling ($\propto N_e^2$) and the latter appears to be a manifestation of Joule heating The electric field in the trough was of the order of 50 mV/m and directed northwards. These results appear to confirm the suggestion made by Rich[29] and Snow, in addition, that these electric field produced troughs can extend into the afternoon sector.

SUMMARY

The precision of a high latitude defense radar operating at VHF or UHF will be limited by ionospheric refraction. Efforts to model this effect can be inaccurate or misleading when irregular structure associated with the auroral zone is present. One form of such structure is a region of low electron density that forms on the nightside and is known as the trough.

On some very magnetically disturbed days, deep troughs are found in the afternoon and evening sectors when the ambient ion density (to the south) is very high. These troughs appear to give rise to the largest quasi-stationary refraction effects observed during a three-year study at Millstone Hill.

Recent satellite and incoherent scatter observations suggest that these dayside troughs are caused by a large northwards electric field being impressed on the earth's ionosphere, thereby increasing the charge transfer rate for O^+ ions. The origin of this field is thought to be the partial ring current ions which penetrate to lower L-shells than plasma sheet electrons.

ACKNOWLEDGEMENT

The program of propagation study at Millstone Hill was supported by the Department of the Army. The recent incoherent scatter studies were supported by the National Science Foundation.

REFERENCES

1. R. H. Wand, "Ionospheric Limitations on Radar Accuracy at L-Band", Lincoln Laboratory Tech. Note 1975-28. Lexington, Massachusetts (1975). ADA 013 731/5.
2. J. V. Evans and R. H. Wand, "Ionospheric Limitations on the Angular Accuracy of Satellite Tracking at VHF or UHF", Proceedings of AGARD Conference 173, "Radio Systems and the Ionosphere", 7 rue Ancelle 92200, Neuilly sur Seine, France.
3. R. H. Wand and J. V. Evans, "Morphology of Ionospheric Scintillation in the Auroral Zone", in Effect of the Ionosphere on Space Systems and Communications, ed. J. M. Goodman, pp-76-83 (U.S. Government Printing Office, Washington, D.C.) 1975.
4. J. S. Ghiloni, editor, "Millstone Hill Radar Propagation Study: Instrumentation", Technical Report 507, Lincoln Laboratory, M.I.T., Lexington, Massachusetts (1973). AD 775 140/7.
5. J. V. Evans, editor, "Millstone Hill Radar Propagation Study: Radar Calibration", Technical Report 508, Lincoln Laboratory, M.I.T., Lexington, Massachusetts (1973). AD 779 689/9.
6. J. V. Evans, editor, "Millstone Hill Radar Propagation Study: Scientific Results" Technical Report 509 (Parts 1-3), Lincoln Laboratory, M.I.T., Lexington Massachusetts (1973). AD 781 173/7; AD 782 748/8; AD 780 5[illegible]/5
7. J. A. Klobuchar and R. S. Allen, "A First-order Prediction Model of Total-electron Content Group Rate Delay for a Mid-latitude Ionosphere", Report AFCRL-70-0405, Air Force Cambridge Research Laboratories (1973).
8. J. V. Evans and J. M. Holt, IEEE Trans. on Antennas and Propagation, AP-21, 685 (1973).
9. D. B. Muldrew, J. Geophys. Res. 70, 2635 (1965).
10. J. O. Thomas and M. K. Andrews, J. Geophys. Res. 73, 7404 (1968).
11. G. G. Bowman, Planet. Space Sci. 17, 777 (1969).
12. N. J. Miller, J. Geophys. Res. 75, 7175 (1970).
13. M. J. Rycroft and J. O. Thomas, Planet. Space Sci. 18, 65 (1970).
14. Y. Tulunay and J. Sayers, J. Atmos. Terr. Phys. 33, 1737 (1971).
15. D. A. Feinblum and R. J. Horan, "Hilion – A Model of the High Latitude Ionospheric F2 Layer and Statistics of Regular Ionospheric Effects at Fort Churchill", Tech. Rep. U.S. Army Contr. No. DAMC 60-71-C-0005, Bell Telephone Laboratories, Western Electric, Whippany, New Jersey (1973).
16. C. P. Pike, "An Analytical Model of the Main F-Layer Trough", Air Force Surveys in Geophysics No. 348, Rep. AFGL-TR-76-0098, Air Force Geophysics Laboratory, Hanscom Air Force Base, Massachusetts (1976).
17. B. W. Halcrow and J. S. Nisbet, Radio Sci. 12, 815 (1977).
18. M. Mendillo and C. C. Chacko, J. Geophys. Res. 82, 5129 (1977).
19. J. V. Evans, editor, "The Millstone Hill Propagation Study", Technical Note 1969-51, Lincoln Laboratory, M.I.T., Lexington, Massachusetts (1969). AD 701 939.
20. W. C. Knudsen, J. Geophys. Res. 79, 1046 (1974).
21. W. C. Knudsen, P. M. Banks, J. D. Winningham and D. M. Klumpar, J. Geophys. Res. 82, 4784 (1977).
22. R. W. Spiro, R. A. Heelis and W. B. Hanson, J. Geophys. Res. 83, 4255 (1978).
23. J. V. Evans, J. M. Holt and R. H. Wand, J. Geophys. Res. 84, 7059 (1979).
24. J. V. Evans, J. M. Holt, W. L. Oliver and R. H. Wand, J. Geophys. Res. 85, 41 (1980).
25. P. M. Banks, R. W. Schunk and W. J. Raitt, Geophys. Res. Lett. 1, 239 (1974).
26. R. W. Schunk, W. J. Raitt and P. M. Banks, J. Geophys. Res. 80, 3121 (1975).
27. M. Smiddy, M. C. Kelley, W. Burke, F. Rich, R. Sagalyn, B. Shuman, R. Hays and S. Lai, Geophys. Res. Lett. 4, 543 (1977).
28. N. C. Maynard, Geophys. Res. Lett. 5, 617 (1978).
29. F. J. Rich, W. J. Burke, M. C. Kelley and M. Smiddy, J. Geophys. Res. , in press (1980).

The U.S. Government assumes no responsibility for the information presented.

The Combined Effect of Faraday and Squint Rotation for an Electronically Scanned, Satellite-Borne Radar

Gene B. Goldstein
R & D Associates
Marina del Rey, California

ABSTRACT

The average power received at a radar employing a single linear polarization is influenced by the polarization of the field incident upon and scattered by the target. It is well-known that for a satellite-borne radar the ionosphere rotates the polarization of both the incident and scattered fields, thereby affecting the polarization of the field received back at the radar. This is known as the Faraday rotation effect. If the radar employs an electronically scanned antenna, the polarization of both the incident and received fields is also rotated by an amount that depends upon the look direction, which is not fixed. We refer to this scan-induced polarization change as squint rotation. To calculate the power received at such a radar it is necessary to consider the combined effects of Faraday and squint rotation. This is done below for the case of a radar employing a horizontally polarized, electronically scanned antenna.

INTRODUCTION

When a linearly polarized, plane electromagnetic wave is transmitted through an ionized medium in the presence of an external magnetic field, the polarization rotates through an angle Ω_F in the $\bar{E}$-$\bar{H}$ plane. The angle Ω_F is the one-way Faraday rotation angle.

In a satellite-to-ground radar application a single, linearly polarized antenna is frequently employed for both transmission and reception. Both the transmitted and scattered fields pass through the ionosphere in the presence of the earth's magnetic field and will therefore undergo Faraday rotation. The effect of Faraday rotation upon the received signal-to-clutter ratio (S/C) depends upon the magnitude of Ω_F and upon the polarization scattering matrix of the target and background clutter. We will see below that Ω_F is proportional to f^{-2}, where f is the transmitted radar frequency, and that for the example L-band satellite radar considered here, the magnitude of Ω_F is of the order of 10°-15°. For the target and clutter scattering matrices considered in our example, the result is a S/C degradation due to Faraday rotation of 1/2 dB or less.

But an additional source of polarization rotation, viz. squint rotation, occurs when the radar employs electronic scanning. In an electronically scanned antenna, the mainlobe is pointed in different directions by adjusting the phase of the antenna elements instead of physically rotating the antenna face. The direction of the mainlobe establishes the line-of-sight (LOS) of the electronically scanned antenna.

In general, squint rotation occurs whenever the LOS is "squinted" off the boresite axis.* Squint rotation is not an effect of the ionosphere. In fact, we define the squint rotation angle, Ω_{sq}, to be the angle through which polarization incident at the target has rotated in the absence of Faraday rotation, relative to the case when the LOS is directed along the boresite (no squint).

*The boresite axis is defined here to be a line perpendicular to and passing through the center of the face of the antenna.

When the propagation path includes the ionosphere, both Faraday rotation and squint rotation will affect the received power. In general, the combined effect of Faraday and squint rotation upon the received S/C cannot be found by considering them separately. However, in the example illustrated below, an error of less than 1dB is committed by separately calculating the S/C losses for each effect and adding the results (in decibels). The reason for this behavior in the example is that the Faraday losses are never larger than 1/2 dB.

ANALYSIS

Description of Geometry

Figure 1 displays the geometry of interest. In this figure a coordinate system (X,Y,Z) is established with its origin, 0, at the center of the earth. The YZ plane is oriented so that it contains the boresite of the transmitting antenna. The radar is located at the center of the (X_S,Y_S,Z_S) coordinate system, 0_S, and the X_S,Y_S,Z_S axes are, respectively, parallel to the X,Y,Z axes. Note also that the Z and Z_S axes coincide and lie along the radial line drawn from the center of the earth to 0_S.

We also establish a spherical coordinate system centered at 0_S using the azimuth angle ϕ and the nadir angle γ locating the LOS to a point P on the surface of the earth. The unit vectors of this spherical system are denoted by $\vec{e}_R$, $\vec{e}_\phi$, $\vec{e}_\gamma$, and they are also illustrated in Figure 1. The motivation for establishing this spherical system is that in the far field it can be shown (see, e.g., [1]) that the radiated field will possess components only in the ϕ-γ plane; i.e., the field propagates along a spherical wavefront, with the sphere centered at 0_S. The polarization of the field is determined by the orientation of the plane containing the propagation vector (a unit vector directed along the LOS) and the $\vec{E}$ field.

Calculation of Squint Rotation Angle

We are now prepared to evaluate the angle, Ω_{sq}. To this end we shall assume that the transmitted field, $\vec{E}_T$, is horizontally polarized. Thus, when expressed in Cartesian coordinates the field $\vec{E}_T$ may be written as

$$\vec{E}_T = E_{X_S}\,\vec{e}_{X_S} \tag{1}$$

Then, transforming from Cartesian to spherical coordinates and assuming that no other polarization rotation effects have occurred, the field incident at the target, $\vec{E}_I$, becomes

$$\vec{E}_I = E_{I_\phi}\,\vec{e}_\phi + E_{I_\gamma}\,\vec{e}_\gamma = \left[-E_{X_S}\sin\phi\right]\vec{e}_\phi + \left[E_{X_S}\cos\phi\cos\gamma\right]\vec{e}_\gamma \tag{2}$$

In (2) we have used the far-field assumption and neglected the radial ($\vec{e}_R$) component of the incident field.

Now let $\vec{n}_I$ and $\vec{n}_T$ denote unit vectors normal to (respectively) the incident field's plane of polarization and the transmitted field's plane of polarization. These unit vectors may be written as

$$\vec{n}_I = \frac{\vec{E}_I}{|\vec{E}_I|} \times \vec{e}_R \quad ; \quad \vec{n}_T = \frac{\vec{E}_T}{|\vec{E}_T|} \times \vec{e}_R \tag{3}$$

Then Ω_{sq} is defined as the angle between $\vec{n}_I$ and $\vec{n}_T$.

In order to ultimately combine the effects of both squint rotation and Faraday rotation it is necessary for us to preserve the sense [i.e., clockwise (C.W.) vs counter-clockwise (C.C.W.)] of the angles Ω_{sq} and Ω_F. We shall therefore adopt the convention of defining the rotation angle to be positive (+) if it is a C.C.W. rotation when viewed in the direction of propagation. For Ω_{sq}, only angles in the range $-\pi/2 \le \Omega_{sq} \le +\pi/2$ are of any significance. We can preserve the necessary

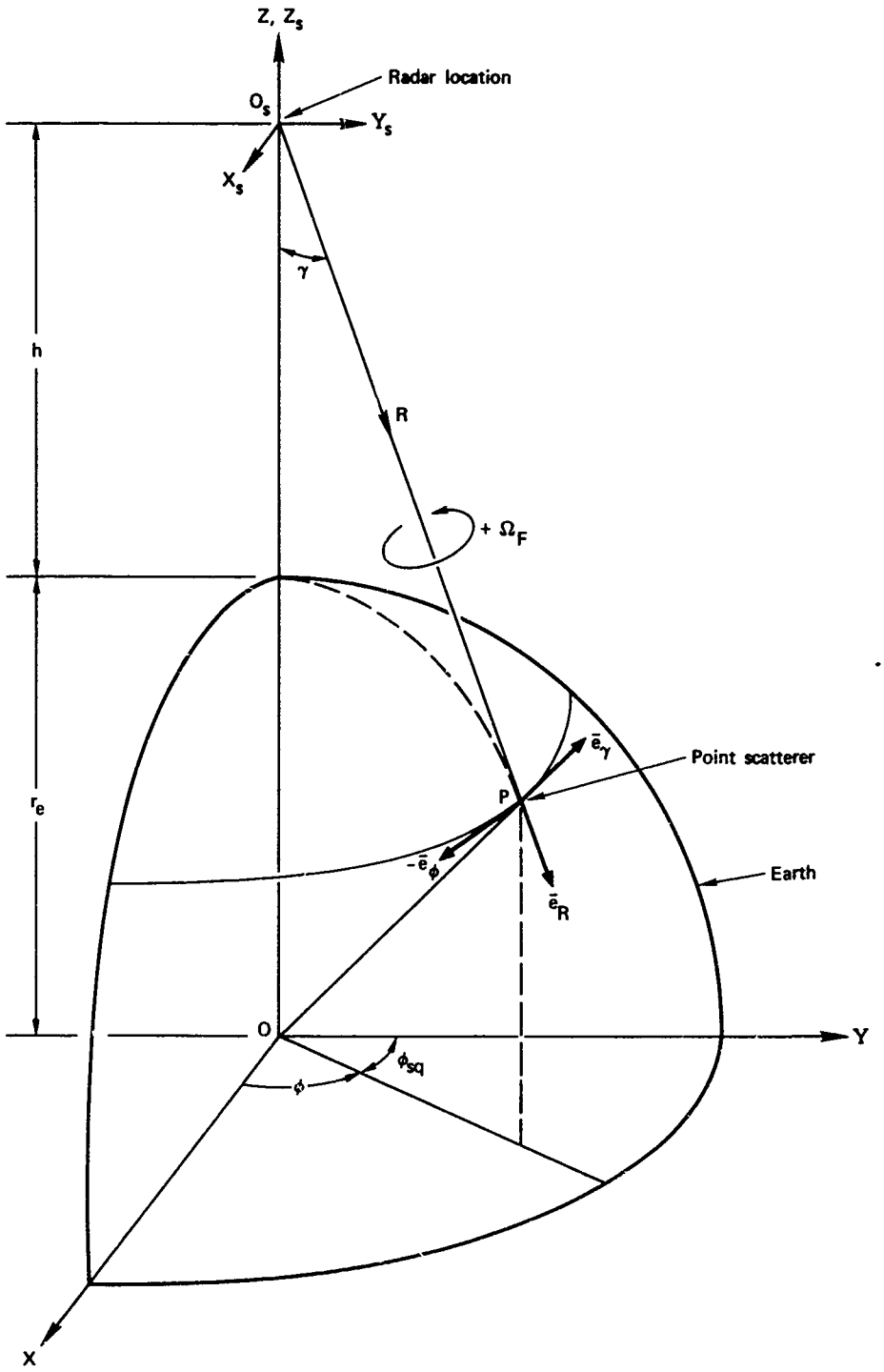


Figure 1. Pertinent geometry.

properties of Ω_{sq} by expressing Ω_{sq} in terms of the tangent of the angle between $\vec{n}_I$ and $\vec{n}_T$. Using (3) the result is

$$\Omega_{sq} = \tan^{-1} \left\{ \cot \phi \cos \gamma \right\} \tag{4}$$

Notice from (4) that $\Omega_{sq} \equiv 0$ whenever $\phi = \pm \pi/2$ (no azimuth squint), irrespective of the value of γ. However, when the LOS is squinted in azimuth, the value of Ω_{sq} decreases as γ departs from nadir. This behavior can also be gleaned intuitively from Figure 1. Therefore, for transmitted horizontal polarization, squint rotation depends not upon the depression angle of the antenna boresite, but upon the depression angle of the LOS. This implies that in applications involving shallow grazing angles at the earth, it is necessary to limit the azimuth squint to a small angular region about the boresite plane in order to minimize Ω_{sq}.

Calculation of Faraday Rotation Angle

The formula we shall use to calculate Ω_F is

$$\Omega_F = \frac{2.362 \times 10^4}{f^2} \int_{h_1}^{h_2} f(h)\, \vec{H}_e(h) \cdot \vec{e}_R\, N_e(h)\, dh \tag{5}$$

This formula is available in a number of standard references on radar propagation [2,3], and the reader is referred to [3] for a discussion of the underlying assumptions and definitions of the quantities involved in the integral. We point out here, however, that we will need the one-way Faraday rotation angle, and many references (including [3]) furnish the result for two-way propagation, which is twice Ω_F. The units of Ω_F in (5) are radians, and Ω_F is positive for C.C.W. rotation viewed in the direction of propagation [4].

Table I below provides representative values for Ω_F when the transmitted frequency is assumed to be f = 1.25 GHz, the radar altitude is 200 nmi and the daytime ionospheric electron density parameters [2,3] are employed in (5). In this table ε is the elevation angle of the LOS measured at the earth (also referred to as the grazing angle), and is related to the nadir angle, γ.

TABLE 1

ONE-WAY FARADAY ROTATION ANGLE, Ω_F (DEGREES)

Frequency = 1.25 GHz Altitude = 200 nmi

ε = elevation angle of line-of-sight (related to γ)

ϕ = azimuth to line-of-sight (Figure 1)

ϕ \ ε	0°	10°	20°	30°
30°	14.2	13.1	11.5	10.2
50°	16.2	14.8	12.8	11.1
60°	16.6	15.2	13 0	11.3
70°	17.0	15.5	13.2	11.5
90°	16.5	15.1	12.9	11.2
110°	14.7	13.6	11.9	10.5
120°	13.4	12.5	11.0	9.9
130°	12.1	11.3	10.2	9.2
150°	8.6	8.4	8.1	7.7

Calculation of Average Power Received

When the Faraday and squint rotation effects are combined, the polarization of the incident field is rotated through the net rotation angle

$$\Omega_1 = \Omega_{sq} + \Omega_F \tag{6}$$

The incident field interacts with a scatterer, producing the scattered field, $\vec{E}_S$ whose polarization is influenced by the polarization scattering matrix, S. We have $\vec{E}_S = S\vec{E}_I$, where

$$S = \begin{bmatrix} s_{HH} & s_{VH} \\ s_{HV} & s_{VV} \end{bmatrix} = \begin{bmatrix} s_{\phi\phi} & s_{\gamma\phi} \\ s_{\phi\gamma} & s_{\gamma\gamma} \end{bmatrix} = s_{\phi\phi} \begin{bmatrix} 1 & a_1 e^{j\alpha_1} \\ a_1 e^{j\alpha_1} & a_2 e^{j\alpha_2} \end{bmatrix} \tag{7}$$

The entries of S are complex quantities (i.e., they have magnitudes and phases) that relate the vector components of the scattered field to those of the incident field. Also, reciprocity guarantees that $s_{\gamma\phi} = s_{\phi\gamma}$.

Upon returning to the radar antenna, the field $\vec{E}_S$ is again rotated by the angle Ω_F, producing the arriving field, $\vec{E}'$. Finally, the received field, $\vec{E}''$, is the component of $\vec{E}'$ aligned with the antenna polarization (in our case, the X_S component of $\vec{E}'$).

The average power received by the radar is proportional to $\overline{|\vec{E}''|^2}$, where we use the overbar to denote averaging over the (assumed random) phase angles, α_1 and α_2. Let $P(\Omega_{sq}, \Omega_F)$ denote the average received power in general and $P_o \triangleq P(0,0)$ denote the average received power in the absence of Faraday and squint rotation. We then obtain the following result:

$$\frac{P(\Omega_{sq},\Omega F)}{P_o} = \overline{T_1^2} + \overline{T_2^2} + 2\,\mathrm{Re}\left\{\overline{T_1\,T_2^*}\right\} \tag{8a}$$

where

$$\overline{T_1^2} = \sin^2\phi \left\{ \cos^2\Omega_1 \cos^2\Omega_F + \left(\frac{\sigma_{VH}}{\sigma_{HH}}\right) \sin^2\Omega_{sq} + \left(\frac{\sigma_{VV}}{\sigma_{HH}}\right) \sin^2\Omega_1 \sin^2\Omega_F \right\} \tag{8b}$$

$$\overline{T_2^2} = \cos^2\phi \cos^2\gamma \left\{ \left(\frac{\sigma_{VV}}{\sigma_{HH}}\right) \sin^2\Omega_1 \cos^2\Omega_F + \left(\frac{\sigma_{VH}}{\sigma_{HH}}\right) \cos^2\Omega_{sq} + \cos^2\Omega_1 \sin^2\Omega_F \right\} \tag{8c}$$

$$\mathrm{Re}\left\{T_1\,T_2^*\right\} = \frac{\sin(2\phi)}{4} \cos\gamma \left\{ \cos^2\Omega_1 \sin(2\Omega_F) - \left(\frac{\sigma_{VH}}{\sigma_{HH}}\right) \sin(2\Omega_{sq}) - \left(\frac{\sigma_{VV}}{\sigma_{HH}}\right) \sin^2\Omega_1 \sin(2\Omega_F) \right\} \tag{8d}$$

$$\frac{\sigma_{VH}}{\sigma_{HH}} = |a_1|^2 \quad , \quad \frac{\sigma_{VV}}{\sigma_{HH}} = |a_2|^2 \tag{8e}$$

Equation (8) is the final result desired. From (8) it is seen that the received power depends upon both Ω_{sq} and Ω_F; not just upon the net rotation, Ω_1. The dependence upon the polarization scattering matrix is also revealed. For example, consider the use of scattering from the target. In (8e) we use σ_{HH} for the radar cross

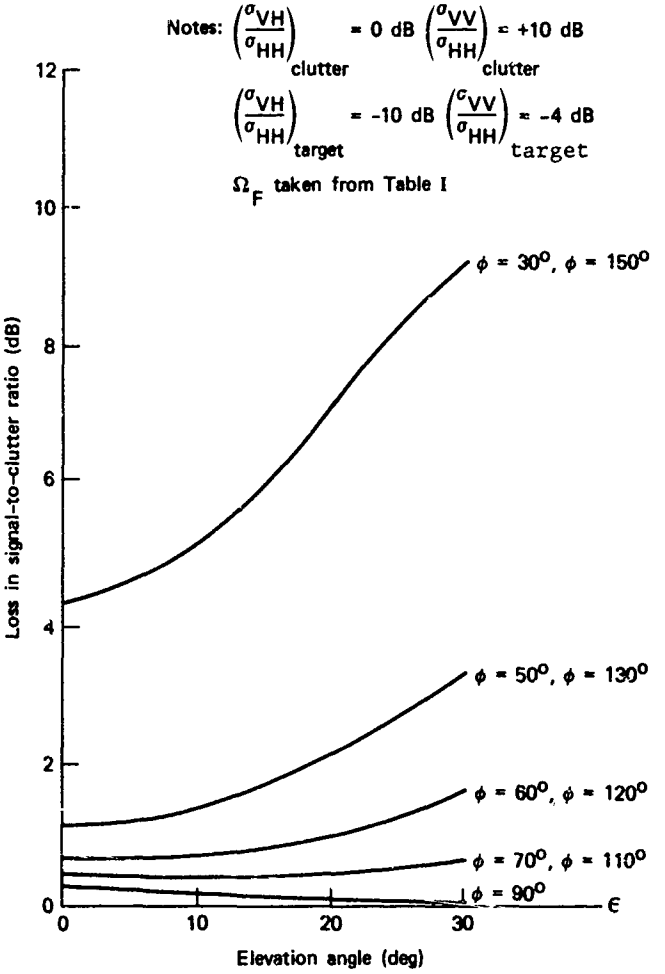


Figure 2. Combined Effect of Faraday and Squint Rotation upon Received S/C.

section of the target for "horizontal polarization", and σ_{VV} and σ_{VH} are used respectively for the target's "vertical" and "cross-polarized" radar cross sections. If the corresponding radar cross sections for clutter are inserted into (8), we obtain the effect of polarization rotations upon the average received clutter power. The combined polarization rotation effect on the received S/C ratio is obtained by using (8) separately for the target parameters and for the clutter parameters and then taking the ratio of the results.

An example of the overall effect of Faraday and squint rotation upon the received S/C ratio is illustrated in Figure 2. The assumed polarization scattering cross sections are shown on the figure. Generally, the scattering cross sections vary with grazing angle, but for simplicity we have assumed constant values which are representative of large ship targets in sea clutter for shallow grazing angles ($0 \leq \varepsilon \leq 10°$).

REFERENCES

1. Ramo, S., and Whinnery, J., Fields and Waves in Modern Radio, pp. 496-498, John Wiley & Sons, New York, 1962.

2. Davies, K., Ionospheric Radio Propagation, pp. 210-213, National Bureau of Standards Monograph 80, Washington, D.C., November 1965.

3. Berkowitz, R.S., Modern Radar, Analysis, Evaluation, and System Design, pp. 318-377, John Wiley & Sons, New York, 1965.

4. Atwater, H., Introduction to Microwave Theory, pp. 206-212, McGraw-Hill, Book Company, New York, 1962.

LIMITATION OF VHF TRACKING CAUSED BY IONOSPHERIC CORRECTION UNCERTAINTIES

S. Pallaschke

European Space Operations Centre
Darmstadt, West-Germany

INTRODUCTION

The European Space Operations Centre (ESOC) at Darmstadt, West-Germany, an establishment of the European Space Agency (ESA) with Headquarters in Paris, France, is in charge of spacecraft control functions including orbit determination. The spacecraft control activity started in spring 1968 and since then 14 satellites have been handled by ESOC.

The satellite orbits can be split up into 3 classes:

- close earth orbits;
- highly eccentric orbits with apogee altitudes up to 250,000 km;
- geostationary orbits including as well the initial phases as transfer- and drift orbits.

A summary of the characteristic parameters of the 14 satellites mentioned above is given in Table I: 'ESOC Supported/Controlled Satellites'.

As time passed the demand for high precision orbit determination increased. However, the demand for an increase in orbit accuracy made it necessary to improve the tracking data handling techniques, primarily in the field of ionospheric corrections when using VHF data. Various models were analysed, two of them with particular emphasis, viz.

W. Becker	The Standard Profile of the Mid-Latitude F Region of the Ionosphere as deduced from Bottomside and Topside Ionograms Space Research XII 1972
R.B. Bent, S.K. Llewellyn, P.E. Schmid	Ionospheric Refraction Corrections in Satellite Tracking Space Research XII 1972

The Bent model has been considered most appropriate to our operational environment and was implemented for the COS-B launch in Aug. 1975. Since then it has been used successfully both for transfer and synchronous orbits as well as for highly eccentric orbiting satellites.

The paper will give an overview of the problems encountered at ESOC with orbit determination when working with VHF tracking measurements.

BENT IONOSPHERIC MODEL

The Bent ionospheric model has been described in

'NASA-GSFC Ionospheric Corrections To Satellite Tracking Data' by
P. Schmid, R. Bent, S. Llewellyn, G. Nesterczuk, S. Rangaswamy
GSFC X-591-73-281 Dec. 1973.

For the sake of completeness a brief summary of the model is given below.

The electron density profile for the F_2 layer (the other layers as F_1, E, D can be neglected for satellite tracking application) is composed of a bi-parabolic bottomside layer and three exponential sections describing the topside layer. The electron density profile is defined by the following parameters:

for the parabolic approximation:

N_m	maximum electron density
h_m	height of maximum electron density
Y_b^m	half thickness of the bottomside part
Y_t^m	half thickness of the topside part

and for the exponential representation:

K_1, K_2, K_3 decay constants for the lower, middle and upper third of the topside layer.

In fact, the model parameters can also be expressed as functions of

- the critical frequency ($f_o F_2$) and
- the maximum usable frequency for a 3000 km single hop earth point-to-point radio communication via the F_2 layer (MUF (3000) F_2).

Both the $f_o F_2$ and MUF (3000) F_2 are functions $\Omega(\phi,\lambda,T)$ of geographic latitude ϕ, longitude λ and time T. This function can be expressed again by a series of products of time dependent functions D(T) and position dependent functions $G(\phi,\lambda)$.

The number of terms used in this harmonic representation covering the whole globe is:

	Number of Terms for		
	D(T)	$G(\phi,\lambda)$	Total
$f_o F_2$	13	76	988
MUF (3000) F_2	9	49	441

After evaluation the computed values are then updated for day-to-day fluctuations using the daily and the 12-months averaged values of the 10.7 cm solar flux.

Once the integrated electron density along the ray path is known the refractive index and consequently the measurement correction can be calculated.

SPSD EXPERIMENT

In the light of the large ionospheric propagation errors of VHF group delay measurements an experimental system applying the real-time correction technique as proposed by Burns and Fremouw (IEEE, Vol. AP-18, No. 6, pages 785 - 790, Dec. 1970) was carried out on COS-B. Basically the Second Phase Shift Difference (SPSD) Technique makes use of the fact that the ionospheric delay is roughly inversely proportional to the square of the signal frequency.

Three frequencies are used within this experiment which are obtained by phase modulating the downlink carrier f_c by a subcarrier f_m in the onboard transmitter, i.e.

$f_u = f_c + f_m$
f_c
$f_l = f_c - f_m$

for COS-B $f_c = 136.95$ Mhz, $f_m = 534.96$ Khz

The SPSD measures

$$\Delta_2 \Phi = (\Phi_u - \Phi_c) - (\Phi_c - \Phi_l)$$

where Φ is the phase of the signal obtained on ground. The phase delay $\Delta\Phi$ for a frequency f can be expressed by

$$\Delta\Phi = -\frac{80.6}{cf} N_t$$

with c as the velocity of light. Thus the total electron content along the path N_t can be derived straight from the SPSD measurement and hence the extra group delay caused by the ionosphere can easily be computed in real time together with the actual ranging measurement.

As mentioned before, this experiment was conducted on COS-B which was launched in Aug. 1975 into a highly eccentric polar orbit. In order not to interfere with the scientific experiments of the satellite, the SPSD could only be recorded inside the van Allen belt, when the gamma ray experiments could not be carried out anyway. For this reason the applicability and hence the end-result was somewhat limited. Furthermore, the solar activity was rather low during 1975. Although the solar activity increased during the later years the COS-B orbit moved to an unfavourable position for SPSD coverage, which is illustrated below:

	August 1975	May 1978	February 1981
perigee height (km)	350	9,980	13,430
apogee height (km)	99,100	89,410	85,980
inclination	90°.3	96°.5	98°.2
arg. of perigee	334°.6	304°.5	277°.0
mean solar flux	80	145	200

Initially, the satellite was tracked by Redu, Belgium (longitude of 5°.1 E, latitude of 50°.0) with ranging and interferometer and by Fairbanks, Alaska (longitude of 147°.5 W, latitude of 65°.0) with ranging only. The current tracking is carried out by Redu with ranging alone where the SPSD ground equipment is installed as well.

Since the installation in Feb. 1976, useful and interesting SPSD measurements have been recorded only up to mid 1978 because of the unfavourable orbit evolution, i.e.

during the interval	21 Feb. 1976 - 30 June 1976	48 passes
	01 July 1976 - 31 Dec. 1976	130 "
	01 Jan. 1977 - 31 Dec. 1977	22 "
	01 Jan. 1978 - 31 Dec. 1978	28 "

Some data have been recorded during the year 1979 and 1980 (20 passes together) but could not be used owing to the low S/N since the satellite could only be acquired at higher altitudes because of the changing perigee location.

The evaluation concentrated on the first 26 months' data and considered primarily the following two aspects:

- comparison of SPSD measurements with the Bent model: long passes with a significant elevation change around sunrise/sunset in order to obtain an increased variation in the ionospheric delay, or short passes with a rapid change in elevation at a period of quiet ionospheric behaviour (around noon) for geometrical analysis;

- stability of measurements. The consistency could be evaluated by comparing night time measurements at higher elevation angles. Contrary to the first aspect the low solar activity was of high interest. The measurement consistency was checked for

  . short term (within the pass over 20 min. for example)
  . long term (within the first quarter of the evaluation period).

The nominal performance of the SPSD measurements was estimated to be about 0°.5 which corresponds to almost 100 m in the COS-B configuration. However, considering the extremely low solar activity at the beginning of the evaluation period the RMS of 100 m is significantly large.

The short and long term calibration tests have verified the nominal performance of 0°.5, in fact a precision of about 0°.4 has been obtained. Problems in calibrating the ground equipment (when changing the polarisation) have been encountered. This is clearly visible on the top left picture on Graph I: Samples of SPSD / Bent Versus Time. The comparison between the SPSD measurements and the Bent model showed for the location Redu in principle a fairly good agreement. In this application the influence of the ionosphere for mid-latitude regions has been computed sufficiently accurate within the Bent model.

Although the evaluation of the SPSD experiment suffered somewhat under unfavourable circumstances (low solar activity, coverage limitation) the validity of the SPSD experiment was demonstrated. The derived accuracy of $0^{o}.4$ would correspond to about 3o - 4o m in case of a GRARR transponder type which is very promising due to the real-time applicability.

TRANSFER ORBIT DETERMINATION

The geosynchronous satellites launched by ESA during the past years were equipped with VHF facilities which were in most of the cases only used during the transfer orbit and the first part of the drift phase. The latter part of the drift phase and the synchronous phase generally use higher frequency bands.

The cases experienced so far have shown that when sufficient time (at least 1.5 revolutions for the transfer orbit and 1 to 2 days for the drift phase) and adequate station coverage is available, the ionospheric uncertainty linked to VHF tracking measurements does not limit the orbit determination. The transfer orbit determination illustrates this.

The typical characteristics of an ESA/ARIANE transfer orbit are:

- transfer orbit	perigee height (km)	2oo
	apogee height (km)	36,ooo
	inclination	$10^{o}.5$
	arg. of perigee	$180^{o}.0$
	longitude of desc. node	$20^{o}.0$ West
	period	1o.5 hours

- ground tracking stations	Station	Longitude	Latitude
	Malindi, Kenya	40.3 E	- 3.0
	Carnarvon, Australia	113.7 E	- 24.9
	Kourou, French Guyana	52.8 W	5.3
	Redu, Belgium	5.1 E	50.0

Up to now apogee motor firing (injection into the drift orbit) always took place at the 2nd or 4th apogee. The nominal tracking sequence for these two cases is as follows:

Injection at perigee 0 (T_i) → T_i + 2.5 hrs. every 15.min. (Malindi, Carnarvon)
T_i + 2.5 hours → Perigee 1 every 3o min. (Malindi, Carnarvon)

in case of apogee motor firing (AMF) at 2nd apogee
Perigee 1 → Apogee 2 every 3o min. (Kourou)

in case of apogee motor firing at 4th apogee
Perigee 1 → Perigee 3 every 6o min. (all 4 stations)
Perigee 3 → Apogee 4 every 3o min. (Kourou, Redu, Malindi)

The 4 ground tracking stations perform VHF ranging (tone ranging system with a minor tone frequency of 1o Hz and a major tone frequency of 2o KHz) with the following overall performance:

ground instrumentation	2o m
onboard instrumentation	1o m

The quoted figures are to be understood as uncertainties within the measurements after calibration. In addition, the uncertainty in the ionospheric delay has to be considered. In order to give an idea of the magnitude, the ionospheric delay for the various stations for a direction towards the equator at an elevation of 30^{o} is tabulated below:

Station	Low Solar Activity	High Solar Activity
Malindi	1,4oo m	4,ooo m
Carnarvon	,oo m	3,8oo m
Kourou	1,4oo m	4,ooo m
Redu	35o m	2,8oo m

Assuming a remaining error of 2o % after modelling the ionospheric effect the following orbit accuracies can be achieved when using tracking measurements up to 1.5 hours prior to the relevant apogee.

Expected Orbit Accuracy (3σ Values)

	Position (km)	Semi-major Axis (m)	Inclination
Assuming low solar activity:			
at 2nd apogee	13.5	150	$0^{o}.010$
at 4th apogee	4.5	25	$0^{o}.006$
Assuming high solar activity:			
at 2nd apogee	50.0	540	$0^{o}.035$
at 4th apogee	20.0	80	$0^{o}.022$

The orbit accuracy required by the attitude determination and the AMF optimisation is very modest, 7o km for the position and 2 km for the semi-major axis. Increasing the ionospheric error to 5o % at high solar activity would still not violate the orbit requirement for an apogee motor firing at the 4th apogee, i.e.

Expected Orbit Accuracy (3σ Values)

	Position (km)	Semi-major Axis (m)	Inclination
Assuming high solar activity with 5o % remaining error:			
at 2nd apogee	120.0	1,300	$0^{o}.085$
at 4th apogee	50.0	190	$0^{o}.055$

As mentioned above the modest requirements on the transfer orbit accuracy in connection with the favourable geometry (sufficient time and ground stations) permit the performance of a successful transfer orbit determination even under the condition of a large remaining error at high solar activity.

It should be mentioned that the experience gained was based on a more inclined transfer orbit type owing to a different launcher (Thor Delta launched at ETR). In one particular case a comparison between the Willman approach (described in 'Advanced Techniques for the Reduction of Geodetic SECOR Observations', STAR N67-15o35, July 1966 by D.C. Brown) and the Bent model has been carried out indicating disadvantages of the Willman model, since additional parameters have to be estimated which could decrease the reliability of the orbit determination.

GEOSTATIONARY ORBITS

Several of the geostationary satellites controlled by ESOC are equipped with tracking facilities of higher frequencies, such as S-band and SHF, in addition to the VHF instruments. Whereas the higher frequency tracking facilities are normally used for the geostationary orbit determination, the VHF capability is used for the transfer and drift orbit determination. Several tracking campaigns have been conducted where both tracking facilities have been used for analysis purposes. Since the ionospheric influence is inversely proportional to the square of the frequency, these tracking campaigns offered good possibilities to evaluate the ionospheric models used within the orbit determination.

The METEOSAT Tracking Exercise

The METEOSAT satellite is positioned at a longitude of 0^{o} and requires for the earth image data processing a precise knowledge of the satellite location. A suitable tracking system (S-band ranging) has been installed at Odenwald, West-Germany (longitude of $9^{o}.0$ E, latitude of $49^{o}.7$) and Kourou, French Guyana (longitude of $52^{o}.8$ W, latitude of $5^{o}.3$). While Odenwald is the commanding station, Kourou houses a land based transponder (LBT). The full ranging sequence consists of alternately

2-way ranging (Odenwald - S/C - Odenwald) and
4-way ranging (Odenwald - S/C - Kourou - S/C - Odenwald).

The overall accuracy of the ranging is

6 m for the 2-way component and
12 m for the 4-way component.

If one compares successive orbit determinations one can derive an estimate for the orbit consistency (satellite position), i.e.

longitude 12o m
latitude 7o m
height 2o m

The distance from the satellite to the two VHF tracking stations Redu, Belgium (longitude of $5^{o}.1$ E, latitude of $50^{o}.0$) and Kourou can be derived with sufficient precision, i.e.

for Redu about 22 m and
for Kourou about 16 m.

In this configuration the ionospheric delay of VHF ranging measurements could be obtained by comparing the VHF value with the computed distance (derived from S-band orbit determination), since other effects such as tropospheric delay could be disregarded. Assuming an overall VHF instrumentation accuracy of about 25 m the ionospheric delay could be deduced with a precision of about 35 m.

One VHF tracking campaign was conducted during the period 24th to 26th April 1979. The comparison between the S-band orbit and the VHF ranging data revealed larger deviations than expected. The differences are given in Table II: VHF / S-Band Tracking Comparison (METEOSAT). If we assume that the differences can be attributed entirely to the ionospheric delay, the performance of the ionospheric model used was not very good. Up to 7o - 8o % of the total delay was the computed correction for the mid-latitude location Redu and only 5o - 7o % for the equatorial station Kourou. Furthermore, the obtained differences did not show a repetitive behaviour over the 3 days so that the orbit determination which relied on VHF data only, converged to a significantly different orbit, i.e.

difference in semi-major axis of 2oo m and
difference in position (27th April) 8 km.

The irregular differences could not be explained fully, the solar flux values were fairly smooth, however, the mean solar flux had already reached a value of 190 x 10^{-22} W/m^{2} Hz.

Unfortunately, the test could not be repeated due to a failure on-board METEOSAT.

The OTS Tracking Exercise

The OTS-2 satellite is positioned at 10^{o} East and should be kept within a margin of $\pm\ 0^{o}.1$. If a reasonable station keeping cycle is to be maintained the orbit must be determined with an accuracy of one order of magnitude better than the positional interval itself. This means $0^{o}.01$ (3σ value) orbit accuracy, which, converted to km, is about 2.5 km (1σ value). The analysis carried out before launch indicated that the OTS-2 prime tracking station Fucino, Italy (longitude of $13^{o}.6$ E, latitude of $42^{o}.0$) with its angular and ranging tracking facilities does not provide enough information for the required positional accuracy. Therefore the Fucino tracking facilities have been complemented by ranging equipment at Villafranca, Spain (longitude of $4^{o}.0$ W, latitude of $40^{o}.4$). The ranging operation is performed within SHF. After some extensive tracking campaigns for calibration purposes, the orbit determination has achieved the following consistency (accuracy):

longitude 1,8oo m
latitude 1,7oo m
height 2oo m

The distance from the satellite to the two VHF tracking stations involved can be derived with the following precision:

for Redu about 3oo m
for Kourou about 24o m.

These figures are not as favourable as the ones obtained for METEOSAT, but the comparison between VHF and SHF data is still meaningful because of the high solar activity these years, which led to an increased ionospheric delay. A VHF tracking test was conducted during the week 23rd to 29th July 193o. In this case the results are comparable to the ones given for METEOSAT, i.e. the ionospheric model was able to correct about 7o - 8o % of the total delay for the station Redu and about 5o - 7o % for the equatorial station Kourou. Tests with varied mean- and daily solar flux have not shown a significant improvement.

SUMMARY

The comparisons between VHF ranging measurements and distances derived from S-band (or SHF) orbit determinations revealed larger uncertainties in the ionospheric modelling, especially in the equatorial region, if one assumes that the difference between the two compared sets can be attributed entirely

to the ionospheric delay.

The errors in modelling the ionospheric delay were up to 20 - 30 % for the mid-latitude station and even 30 - 50 % for the equatorial station.

However, the uncertainty in the ionospheric delay does not cause severe problems when VHF tracking measurements are used for the determination of orbits with sufficiently changing geometry and adequate time and ground station coverage (as it has been demonstrated for the transfer orbit type). The synchronous orbit determination relying on VHF data suffers from the insignificant change in geometry and could easily converge to a biassed solution. If VHF data have to be used, the Second Phase Shift Difference Technique could improve the orbit accuracy significantly.

TABLE I : ESOC Supported / Controlled Satellites

Satellite	Launch Date	Reentry Date	End of Useful Life	Exp. Reentry Date	Mission Objectives	Orbit Type Apogee Height (km)	Incli-nation	Arg. of Per.
ESRO II	May 68	May 71			Cosmic rays, solar X-rays	polar, close earth		
ESRO IA	Oct. 68	June 7o			Auroral & polar cap phenomena, ionosphere	"	"	"
HEOS I	Dec. 68	Oct. 75			Interplanetary medium, bow shock	215,ooo	73°	265°
ESRO IB	Oct. 69	Nov. 69			as ESRO IA	polar, close earth		
HEOS II	Jan. 72	Aug. 75			Polar magnetosphere, interplanetary medium	235,ooo	88°	295°
TD	March 72	Jan. 8o	May 74		Astronomy (UV, X- and gamma rays)	polar, close earth		
ESRO IV	Nov. 72	April 74			Neutral atmosphere, ionosphere, auroral particles	"	"	"
ANS (NL)	Aug. 74	June 77	July 76		Astronomical satellite (ultra-violet rad., X-rays)	"	"	"
COS-B	Aug. 75			Jan. 86	Gamma-ray astronomy	86,ooo	98°	275°
GEOS-1	April 77		July 78	> Dec. 99	Dynamics of the magnetosphere	38,ooo	26°	265°
METEOSAT	Nov. 77				Meteorological application, earth images	geosynchr. at 0° long.		
OTS-2	May 78				Orbital test satellite (European communications)	"	" 10°	"
GEOS-2	July 78				as GEOS-1	"	" 37°	"

GOES-1 (US) was supported by ESOC during the period Nov. 78 to Dec. 79 in connection with the GARP experiment

Remark: The satellites ISEE-2 (launched Oct. 77) and IUE (launched Jan. 78) are not fully supported by ESOC and are therefore not included in this table.

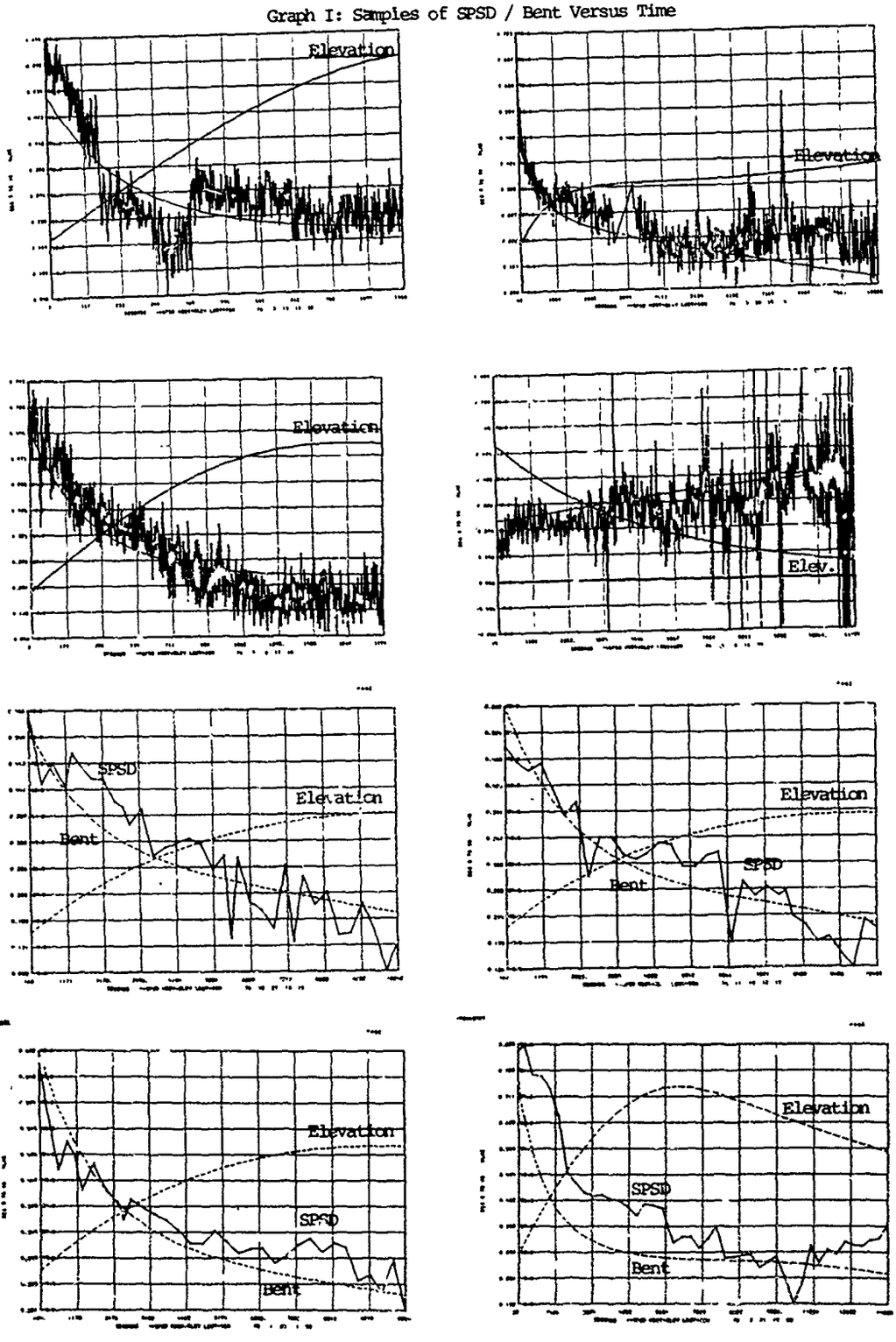
TABLE II : VHF / S-Band Tracking Comparison (METEOSAT)

The table gives the difference between VHF ranging data and distances derived from S-band orbit determination. The value within brackets quotes the corresponding ionospheric delay obtained from the Bent model.

	VHF / S-Band Differences (Meters) on Day:		
Time of the Day	79/04/24	79/04/25	79/04/26
Station Redu:			
00.00		702 (507)	
02.00	442 (447)	367 (434)	214 (485)
04.00	391 (408)	596 (486)	212 (448)
06.00	654 (559)	867 (702)	497 (606)
08.00	888 (786)		
10.00	1250 (1050)		773 (1150)
12.00	1445 (1180)	1630 (1250)	
14.00	1398 (1160)	1226 (1150)	697 (1250)
16.00	1275 (1070)		780 (1140)
18.00	1186 (943)	741 (861)	
20.00	866 (694)	493 (647)	712 (743)
22.00	748 (574)	316 (562)	549 (617)
Station Kourou:			
06.00	1039 (882)	526 (948)	825 (980)
08.00	253 (507)		226 (553)
12.00	2479 (1780)	2024 (1860)	
14.00	3868 (2140)	2302 (2240)	
16.00	4212 (2390)	2442 (2510)	3890 (2570)
18.00	3703 (2370)	2584 (2480)	
20.00	3284 (2060)	3035 (2160)	2611 (2210)
22.00	1935 (1770)	2042 (1860)	

Remark: In fact the Redu measurements refer to 1 hour later on the day 79/04/25 (i.e. 01.00 up to 23.00).

Graph I: Samples of SPSD / Bent Versus Time



TEMPORAL BEHAVIOR OF PULSES AFTER PROPAGATING THROUGH A TURBULENT IONOSPHERE

CHIH-CHUNG YANG AND K. C. YEH
DEPARTMENT OF ELECTRICAL ENGINEERING
UNIVERSITY OF ILLINOIS
URBANA, ILLINOIS 61801

ABSTRACT

A radio signal after propagating through a turbulent ionosphere will suffer distortion owing to dispersion and random scattering. For coarse description of a temporal signal, the temporal moments have been found to be convenient. Past studies have shown that the zeroth moment is related to the total energy in the pulse, the first moment is related to the mean arrival time, and the second moment is related to the mean square pulse width. In this paper, we extend the analysis to the third and fourth moments which are shown to be related to the skewness and kurtosis of the pulse. In digital communications, an ideal pulse, due to propagation effects, can be distorted and stretched to produce a long tail occupying an interval longer than one communication bit. Using the temporal moments an upper bound of the energy content outside of one chip can be estimated.

I. INTRODUCTION

When radio signals propagate through a turbulent ionosphere, such as that often occurs in satellite-earth communication links, they suffer distortion due to frequency dispersion and random scattering. For coarse descriptions of the distorted signal, the temporal moments have been found to be useful [Yeh and Liu, 1977a]. With applications to various communication links in mind these temporal moments have been calculated up to the third moment in a nondispersive medium [Liu and Yeh, 1978] and up to the second moment in a dispersive medium [Yeh and Yang, 1977c; Liu and Yeh, 1977; Yeh and Liu, 1977b; Yeh and Liu, 1979].

In this paper we extend these earlier works. In Section II, the evaluation of the temporal moments is extended to the fourth order in a turbulent ionosphere. Section III contains a discussion of the physical significance of these moments. Quantities such as skewness and kurtosis will then be defined in terms of the temporal moments. These quantities are used to describe the temporal characteristics of the signal intensity distribution. The formulas will then be specified by two sets of parameters that describe the geometry and content of the medium. In Section IV, we introduce some properties that are satisfied by the moments of the probability distribution. These properties are then applied to our signal intensity distribution. An upper bound for the signal energy outside of some time interval from the arrival time in the trailing edge is obtained. The results may be used to determine conditions under which the intersymbol interference can be considerably reduced.

The geometry of the problem is shown in Figure 1. We assume that the ionosphere lies in the region $z \geq 0$ with a homogeneous plasma background and that random irregularities exist only inside the slab between $z=0$ and $z=L$. The positions of the transmitter and the receiver are also shown in this figure. Note that the receiver is always at a position with $z \geq L$.

II. FORMULAS OF TEMPORAL MOMENTS

Assume a plane wave aussian pulse at carrier frequency ω_c with envelope $\exp(-t^2/T_0^2)$ is impressed at $z=0$. In the region $z>0$, this wave propagates in the z

direction in a turbulent ionosphere. By Fourier superposition this wave can be represented by the equation

$$P(z,t) = \int_{-\infty}^{\infty} f(\omega)\ u(z,\omega)\ e^{j[\omega t - k(\omega)z]}d\omega \qquad (1)$$

where $u(z,\omega)$ is the complex amplitude and is in general a random function of position $(\vec{\rho},z)$ and frequency ω. The quantity $f(\omega)$ is the frequency spectrum of the impressed signal and $k(\omega)$ is the wave number.

Define the nth temporal moment by the equation

$$M^{(n)}(z) \equiv \int_{-\infty}^{\infty} \langle P^*(z,t)t^nP(z,t)\rangle dt \qquad n=0,1,2,\ \ldots. \qquad (2)$$

For physical understanding it is more convenient to deal with central temporal moment which is defined by

$$\bar{M}^{(n)}(z) \equiv \int_{-\infty}^{\infty} (t-M^{(1)}/M^{(0)})^n \langle P^*(z,t)P(z,t)\rangle dt$$

$$= \sum_{j=0}^{n} \frac{n!}{j!(n-j)!} (-M^{(1)}/M^{(0)})^{n-j}\ M^{(j)} \qquad n=0,1,2,\ \ldots. \qquad (3)$$

It is evident that $\bar{M}^{(0)}=M^{(0)}$ and $\bar{M}^{(1)}=0$. To calculate the moments, the pulse P represented by (1) is substituted in (2). One immediately recognizes the occurrence of the one-position two-frequency mutual coherence function $\Gamma_2(z,0,\omega_1,\omega_2) \equiv \langle u(z,\vec{\rho},\omega_1)u(z,\vec{\rho},\omega_2)\rangle$. This function is related to the two-position two-frequency mutual coherence function $\Gamma_2(z,\vec{\rho}_1-\vec{\rho}_2,\omega_1,\omega_2)$ which satisfies a certain transport equation in a turbulent ionosphere [Tatarskii, 1971]. Even though the exact analytic solution to this transport equation does not exist at present time the form of (2), after (1) is inserted, suggests that for the evaluation of $M^{(n)}$ we do not need to know Γ_2 but $\partial^m\Gamma_2/\partial\omega_1^m\big|_{\omega_1=\omega_2}$ with m=1,2,...n. Thereby a method is found to calculate $M^{(n)}$ exactly with multiple scattering effects fully taken into account. The drawback of this method is that the algebra involved increases almost geometrically with the order of the moment. This is why in earlier papers only moments up to the 2nd order were computed (e.g. Yeh and Liu, 1977a). The calculations for the central temporal moments are very lengthy. Readers interested in intermediate steps should consult Yang [1981]. In the following, we give only the resulting formulas of the normalized first moment and the normalized second and third central moments. The formula for the fourth central moment is very long but available elsewhere [Yang, 1981].

$$M^{(1)}/M^{(0)} = [z/c - A_2c^{-1}(2Lz-L^2)X^2n^{-4}/4]n^{-1} \qquad (4)$$

$$\bar{M}^{(2)}(z)/M^{(0)} = T_0^2/4 + A_0Lc^{-2}X^2n^{-6}/4 + 4A_4\omega_c^{-2}(3Lz^2-3L^2z+L^3)X^2n^{-8}/3$$
$$+ A_2^2c^{-2}(L^2z^2/4 - L^3z/3 + L^4/8)X^4n^{-10} \qquad (5)$$

$$\bar{M}^{(3)}(z)/M^{(0)} = 3zc^{-1}\omega_c^{-2}X\,n^{-5}/2 - A_2c^{-1}\omega_c^{-2}(2Lz-L^2)X^2(24+15X)n^{-9}/8$$
$$- 6A_6c\omega_c^{-4}(12Lz^3-18L^2z^2+12L^3z-3L^4)X^2n^{-11}$$
$$- A_2A_4c\omega_c^{-2}(12L^2z^3-22L^3z^2+15L^4z-18L^5/5)X^4n^{-13}$$
$$- A_2^3c^{-3}(L^3z^3/4 - L^4z^2/2 + 7L^5z/20 - L^6/12)X^6n^{-15} \qquad (6)$$

In the equations above, $X\equiv\omega_p^2/\omega_c^2$ and $n^2\equiv 1-X$, where ω_p is the background plasma frequency of the ionosphere. Also A_0, A_2, A_4, A_6, A_8 are defined by the equation

$$A_\xi(\vec{\rho}) = A_\xi(\rho) = A_0+A_2\rho^2+A_4\rho^4+A_6\rho^6+A_8\rho^8+.... \quad (7)$$

where $A_\xi(\vec{\rho})$ is given by

$$A_\xi(\vec{\rho}) \equiv \int_{-\infty}^{\infty} B_\xi(\vec{\rho},z)dz \quad (8)$$

and

$$B_\xi(\vec{\rho},z) \equiv <\xi(\vec{\rho}',z')\xi(\vec{\rho}'+\rho,z'+z)> \quad (9)$$

here < > denotes ensemble average. In equation (9), $\xi(\vec{\rho},z) \equiv \Delta N(\vec{\rho},z)/N$, where N is the electron density of the background medium which is assumed to be homogeneous, ΔN is the fluctuating electron density which we have assumed to be an isotropic homogeneous random field.

III. PHYSICAL MEANINGS OF TEMPORAL MOMENTS

The first moment and second central moment have been evaluated and discussed in many papers in the past [Mark, 1972; Yeh and Yang, 1977c; Liu and Yeh, 1977; Yeh and Liu, 1977b; Yeh and Liu, 1979]. The normalized first moment represents the time position of the energy weighted by the intensity distribution. Therefore it is just the arrival time of the signal and we denote it by t_a. From equation (4), it is easy to see that the signal has been delayed when received at the receiver in excess to the free space value z/c since A_2 is always negative.

Now for further calculation, we use the power spectrum for the random turbulence introduced by Shkarofsky [1968]. The two dimensional correlation function $A_\xi(\rho)$ corresponding to this power spectrum is given by

$$A_\xi(\rho) = \frac{2\pi(\kappa_0\sqrt{\rho^2+\ell_0^2})^{(p-2)/2}K_{(p-2)/2}(\kappa_0\sqrt{\rho^2+\ell_0^2})\sigma_\xi^2}{\kappa_0(\kappa_0\ell_0)^{(p-3)/2}K_{(p-3)/2}(\kappa_0\ell_0)} \quad (10)$$

where K is the Hankel function of imaginary argument and $\kappa_0=1/L_0$. The quantities L_0 and ℓ_0 are the outer and inner scales, respectively. The quantity σ_ξ^2 is the variance of electron density fluctuations.

For numerical results, we take L_0=10 kilometers and ℓ_0=10 meters and σ_ξ^2=0.1. Furthermore we use two sets of parameters which are listed as follows:

		Model 1	Model 2
Plasma frequency	f_p	10 MHz	50 MHz
Distance	z	500 km	1000 km
Width of random irregularity slab	L	200 km	500 km

These parameters are picked so that Model 1 corresponds to conditions that can occur naturally in the equatorial ionosphere and Model 2 corresponds to conditions of an ionosphere disturbed by nuclear explosions. In our calculations the width of the impressed pulse T_0 is varied with the carrier frequency f_c to keep the product f_cT_0 constant, i.e. f_cT_0=100 in our case. This is consistent with the requirement, $f_cT_0 >> 1$.

In Figure 2 we present the numerical results for the arrival time. The excess time, t_a-z/v_g, is plotted as a function of the carrier frequency for both Model 1 and Model 2, where $v_g \equiv cn$, the group velocity. This excess time has a very strong dependence on the radio carrier frequency in the form of f_c^{-4}.

The second central moment given by equation (5) has been identified earlier as the mean square pulse width of the signal. Denote the square root of it by τ. The dependence of the normalized mean pulse width $\tau/(T_s/2)$ on the carrier frequency is shown in Figure 3. Note that in Model 1, if the carrier frequency is between 100 MHz and 140 MHz, the term containing A_4, which has f_c^{-6} dependence, is most important. In some earlier papers [Sreenivasiah et al., 1976; Leader, 1979] $A_\xi(\rho)$ in (7) is approximated by a quadratic expression $A_0+A_2\rho^2$. With this approximation, they would miss the A_4 term completely which happens to be dominating under the conditions above. Then their results will be in error when applied to our problem.

Next we consider the central third moment given by (6). Note that A_2 and A_6 are always negative and A_4 is positive. This implies that the third central moment is always positive. Physically, this means that due to propagation effects the pulse becomes asymmetrical with a trailing edge longer than the leading edge. To get a better understanding, let us define the skewness, s, through the following relation:

$$\bar{M}^{(3)}(z)/M^{(0)} = s\tau^3 \tag{11}$$

The value of s is a measure of the extent of the signal asymmetry. Roughly speaking, larger s means bigger relative difference of the stretched lengths between the trailing edge and the leading edge. In Figure 4, we show a distribution with an exponential decay in trailing edge and a Gaussian decay in leading edge. This distribution has skewness s=1.03. If any one clump distribution has s value much smaller than unity, we may conclude that this distribution is less asymmetric than that shown in Figure 4. For our signal intensity distribution, the dependence of skewness on the carrier frequency is shown in Figure 5. Note that if the carrier frequency is above 500 MHz for Model 1 and is above 3 GHz for Model 2, the values of s are much smaller than 1. So we can say in those frequency ranges the received signals must be less asymmetric than the distribution shown in Figure 4, if by some other means we can make sure that the signal is in a single clump.

Because of the length the formula of the fourth central moment given elsewhere [Yang, 1981] is not repeated in this paper. Define the kurtosis, K, by the following relation:

$$\bar{M}^{(4)}/M^{(0)} = \tau^4(K+3) \tag{12}$$

The kurtosis is a dimensionless measure of the distribution concentration extent. For the distribution shown in Figure 4, the value of K is 2.94. If any other one clump distribution has kurtosis value less than 2.94, it must be more concentrated than the distribution shown in Figure 4.

In Figure 6, we show the dependence of K on the carrier frequency for our signal intensity distribution. Note that when the carrier frequency is above 700 MHz for Model 1 and is above 4 GHz for Model 2, the values of K are less than 2.94. This implies that in those frequency ranges the received signal intensity must be more concentrated. So we can conclude that in those frequency ranges the signal must decay faster than the exponential decay on both sides if we assume the received signal is just in a single clump.

IV. DESCRIPTION OF THE SIGNAL TAIL

From the results in previous sections, we have learned that the propagating pulse can be distorted in two ways: pulse width broadening and increasing asymmetry of the pulse. In this latter aspect, we find the distorted signal always has a longer trailing edge. When serious this long tail can impact its use in digital communication because the received signal will not be contained in its original bit; a substantial amount of its energy may extend into the next few bits. As a result errors in the decision making process for the neighboring bits may be committed. This is the reason why it is important to estimate the amount of energy contained outside of the original bit itself in the trailing edge.

Since we are dealing with the normalized signal intensity distribution, this distribution density as a function of time is always positive and the total distribution can be normalized to unity. These properties satisfy the conditions for a function to be a probability distribution density function. So, any theorem about the moments of the probability distribution as a function of random variable can be cited in our research, if these moments are defined in the same ways as the temporal moments. Here let us introduce a property about the moments of the probability distribution [Mises, 1964]. In the following statements, we will not distinguish the probability distribution from the signal intensity distribution, since the results have no difference. This property about the moments can be stated as follows:

> Suppose there exists a distribution I(t) which has an m-step increasing cumulative distribution function $\varepsilon(t)$ and has first 2m-1 order moments $\bar{M}^{(0)}$, $\bar{M}^{(1)}$, $\bar{M}^{(2)}$, ..., $\bar{M}^{(2m-2)}$. If I'(t) is any other distribution with the same 2m-1 moments as above, then the graph of its cumulative distribution function $\varepsilon'(t)$ passes through each "step" and each "riser" of this m-step function $\varepsilon(t)$.

Note that the cumulative distribution function $\varepsilon(t)$ and the distribution density function I(t) is related by

$$\varepsilon(t) = \int_{-\infty}^{t} I(t')dt'$$

As an example, we show a 3-step case in Figure 7. In this figure, for the purpose of our application we denote the positions of the "risers" by t_1, t_2, t_3 and the heights of the respective riser by ε_1, ε_2, ε_3. Note that ε_1, ε_2, ε_3 must be positive and the moments up to fifth order are concerned. Since in the statements of the property we have used the central moments to specify the m-step increasing function, the origin of the horizontal axis is the arrival time of the signal.

In the case of a 3-step function stated above, in order to calculate the positions and heights of "risers", i.e. t_1, t_2, t_3, ε_1, ε_2, ε_3 (6 unknowns) in Figure 7, the moments up to fifth order (6 knowns) are needed. But, if we have the information of only up to fourth moment, we still can construct a 3-step increasing function except that any one of t_1, t_2, t_3, ε_1, ε_2, ε_3, must be left undetermined. This undetermined one can be chosen arbitrarily within some restricted ranges.

Since we are trying to estimate the signal energy outside some time interval from the arrival time in the trailing edge, among t_1, t_2, t_3, ε_1, ε_2, ε_3 we will let t_3 unfixed. Now let us compute t_1, t_2, ε_1, ε_2, ε_3 in terms of t_3, τ, s, K. The starting point of the calculation is the equivalence of those moments between our signal intensity distribution and the distribution with 3-step increasing cumulative distribution function $\varepsilon(t)$. The corresponding distribution density function I(t) of $\varepsilon(t)$ can be expressed by a sum of delta functions as

$$I(t) = \varepsilon_1\delta(t-t_1)+\varepsilon_2\delta(t-t_2)+\varepsilon_3\delta(t-t_3) \tag{13}$$

The calculations are straightforward and the results are as follows:

$$\varepsilon_1 = [\tau^2+t_2t_3]/[(t_3-t_1)(t_2-t_1)] \tag{14}$$

$$\varepsilon_2 = [\tau^2+t_1t_3]/[(t_3-t_2)(t_1-t_2)] \tag{15}$$

$$\varepsilon_3 = [\tau^2+t_1t_2]/[(t_1-t_3)(t_2-t_3)] \tag{16}$$

and

$$t_1 = (p-\sqrt{p^2-4q})/2 \tag{17}$$

$$t_2 = (p+\sqrt{p^2-4q})/2 \tag{18}$$

where

$$p = [\tau^2t_3(K+2)+\tau^3s-t_3^2s\tau]/[\tau^2+t_3\tau s-t_3^2] \tag{19}$$

$$q = \tau^2[\tau^2 s'-t_3\tau s-\tau^2(K+3)+t_3^2]/[\tau^2+t_3\tau s-t_3^2] \qquad (20)$$

We have constructed the 3-step increasing function $\varepsilon(t)$ whose distribution density function $I(t)$ has identical moments up to the fourth order as those of our signal intensity distribution $I'(t)$. Then the graph of $\varepsilon'(t)$ must pass through the third "riser" of $\varepsilon(t)$ at t_3 (see Figure 7). Furthermore, since the height from this intersection point to the top step is the fractional signal energy beyond t_3, ε_3 is therefore an upper bound of this fractional energy. With applications to binary communication in mind let us put t_3 at the boundary of a bit and then compute the fractional signal energy extended into the neighboring bits in the trailing edge. It should be reminded that such computations of t_3 will make sense when t_1 and t_2 are real and ε_1, ε_2, ε_3 are real and positive.

To get some numerical results, we use the parameters given in above two models once again. We plot the curves of ε_3 as a function of the carrier frequency in Figure 8. Two values of t_3 are assigned, one at T_0, the other at $2T_0$. In both models, ε_3 asymptotes to about 0.1 for $t_3=T_0$ and to about 0.005 for $t_3=2T_0$. Also we find when the carrier frequency is larger than 1 GHz in Model 1 and larger than 10 GHz in Model 2, ε_3 is approximately equal to the individual asymptotic value. These two values of the carrier frequency coincide roughly with those beyond which the propagation effects on the pulse width and symmetry can be neglected, as shown in Figures 3 and 5, respectively. So we can conclude that the signal shape is nearly unchanged for the carrier frequency range above those two values in the respective model. Because we have the restrictions on the choice of t_3, we can not complete those curves in Figure 8 when the carrier frequency is below some value.

V. CONCLUSION

We have considered a narrow-banded Gaussian envelope carrier signal being impressed at z=0. After propagating through a turbulent plasma, owing to dispersion and random scattering, this originally symmetric signal is broadened and becomes asymmetric. The trailing edge is longer than the leading edge. From the information we obtained, we can not tell whether the received signal is just in a single clump. But if we can make sure of it by some other means, we compare the skewness and kurtosis values of the average signal intensity distribution with those of the distribution shown in Figure 4 and then get a rough idea about the shape of the received signal.

The results also show the dominating propagation effect between dispersion and random scattering. As the arrival time and pulse width are concerned, the random scattering effect is more important. As far as the signal asymmetry is concerned, dispersion is more important in the high frequency part and random scattering is more important in the low frequency part for cases considered. Finally, for the extent of signal concentration, random scattering effect is dominating.

Although we can not know the exact shape of the received signal intensity distribution, we can find an upper bound for the fractional signal energy beyond some time distance from the arrival time. In a binary communication, this information may help us to predict the errors in a decoding process. But to do it, we need more investigations.

REFERENCES

Leader, J. C. (1978), Intensity fluctuations from partially coherent light propagating through atmospheric turbulence, J. Opt. Soc. Am., 69, pp.73-84.

Liu, C. H., and K. C. Yeh (1977), Propagation of pulsed beam waves through turbulence, cloud, rain, or fog, J. Opt. Soc. Am., 67, No. 9, September.

Liu, C. H., and K. C. Yeh (1978), Pulse propagation in random media, IEEE Trans. on Antennas and Propagation, Vol. AP-26, No. 4, July.

Mark, W. D. (1972), Characterization of stochastic transients and transmission media: The method of power-moments spectra, Journal of Sound and Vibration, 22(3), pp.249-295.

Mises, R. v. (1964), Mathematical Theory of Probability and Statistics, Sec. 4 and 5, Chapter VIII, pp.384-396, Academic Press.

Shkarofsky, I. P. (1968), Generalized turbulence space-correlation and wave-number spectrum-function pairs, Can. J. Phys., Vol. 46, pp.2133-2153.

Sreenivasiah, I., A. Ishimaru, and S. T. Hong (1976), Two-frequency mutual coherence function and pulse propagation in a random medium: An analytic solution to the plane wave case, Radio Sci., 11, 775-778.

Tatarskii, V. I. (1971), The Effects of the Turbulent Atmosphere on Wave Propagation, U.S. Department of Commerce, National Technical Information Service, Springfield, Va.

Yang, C. C. (1981), Temporal Behavior of Pulses after Propagating through a Turbulent Ionosphere, M.S. thesis, Department of Electrical Engineering, University of Illinois.

Yeh, K. C., and C. H. Liu (1977a), An investigation of temporal moments of stochastic waves, Radio Sci., Vol. 12, No. 5, pp.671-680, Sept.-Oct.

Yeh, K. C., and C. H. Liu (1977b), Pulse delay and pulse distortion by random scattering in the ionosphere, AGARD-CP-244, Aspects of Electromagnetic Wave Scattering in Radio Communication, edited by A. N. Ince.

Yeh, K. C., and C. C. Yang (1977c), Mean arrival time and mean pulsewidth of signals propagation through a dispersive and random medium, IEEE Trans. on Antennas and Propagation, Vol. AP-25, No. 5, Sept.

Yeh, K. C., and C. H. Liu (1979), Ionospheric effects on radio communication and ranging pulses, IEEE Trans. on Antennas and Propagation, Vol. AP-27, No. 6, Nov.

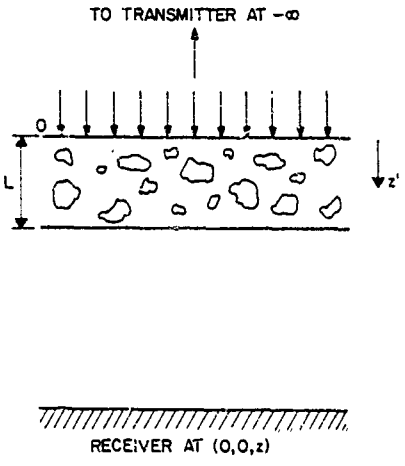


Fig. 1. The geometry of the problem. Note that plasma occupies the whole region $0 \leq z' \leq \infty$, but random irregularities exist only inside the slab $0 \leq z' \leq L$.

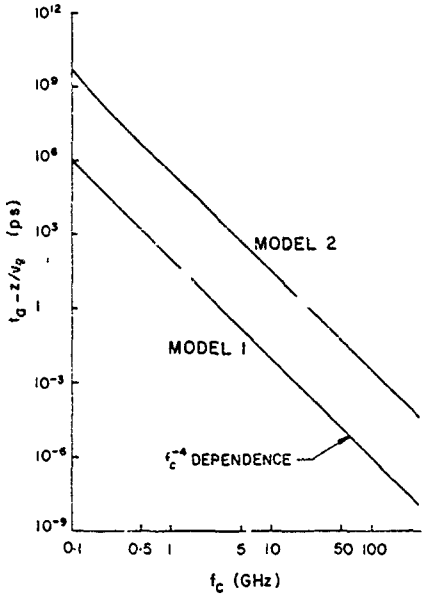


Fig. 2. Excess time $t_a - z/v_g$ as a function of the carrier frequency. Except a small portion of the curve at the low frequency end in Model 2, the two parallel straight lines indicate a f_c^{-4} dependence.

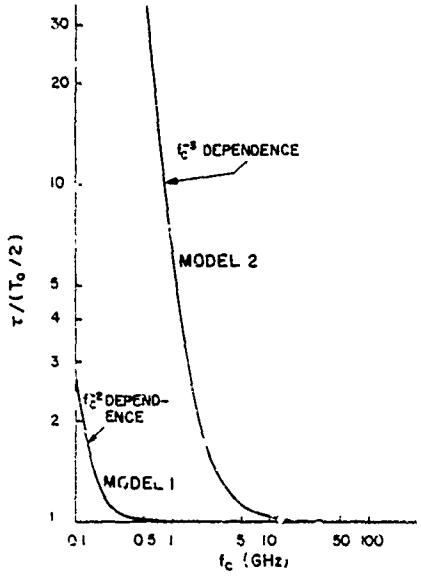


Fig. 3. Normalized mean pulse width $\tau/(T_0/2)$ as a function of the carrier frequency.

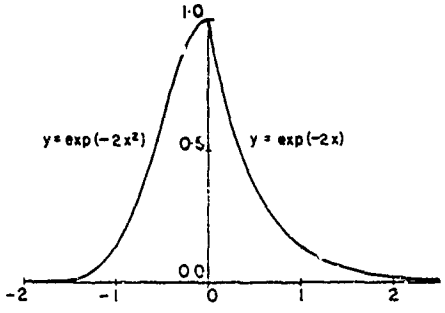


Fig. 4. A distribution with an exponential decay in the trailing edge and an Gaussian decay in the leading edge. It has skewness 1.03 and kurtosis 2.94. Note that we have chosen the origin of the horizontal axis to make $\bar{M}^{(1)}=0$.

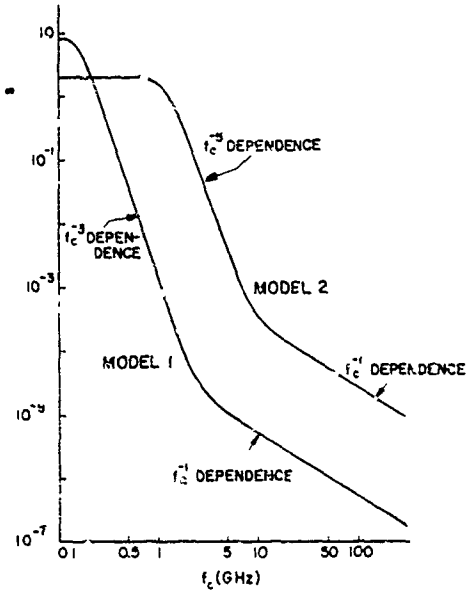


Fig. 5. Skewness s as a function of the carrier frequency f_c. For a definition of s, see equation (11) in the text.

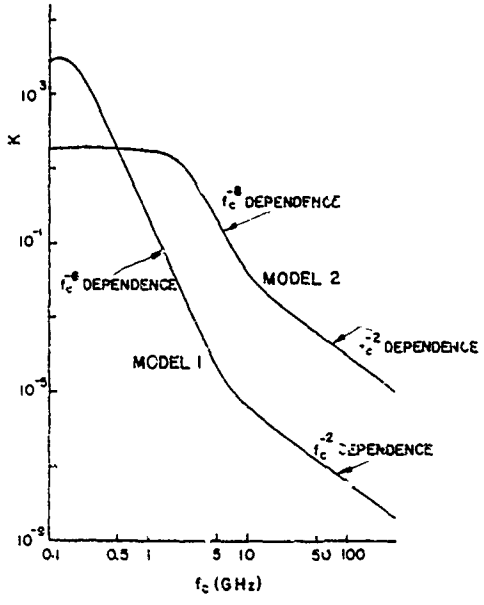


Fig. 6. Kurtosis K as a function of the carrier frequency f_c. For a definition K, see equation (12) in the text.

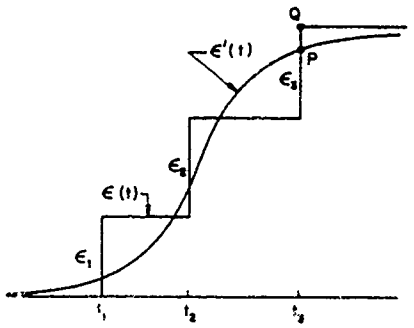


Fig. 7. A 3-step function and a increasing curve whose corresponding distribution density functions have the same moments up to fifth order. ε_i represents the height of the "riser" at t_i, i=1,2,3. Note that $\varepsilon_1+\varepsilon_2+\varepsilon_3=1$. In this case, $\varepsilon'(t)$ passes through each "step" and each "riser" of this 3-step function $\varepsilon(t)$. Since the height between point P and Q is the fractional signal energy beyond t_3, ε_3 is an upper bound of this fractional energy.

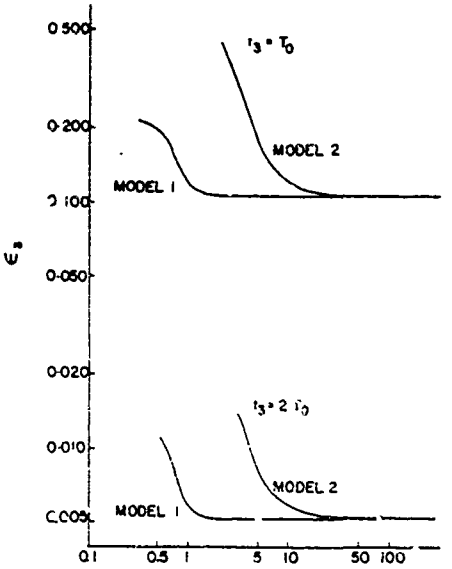


Fig. 8. ε_3, an upper bound of the fractional energy beyond t_3, as a function of the carrier frequency. Two values of t_3 are assigned.

PROPHET - AN EMERGING HF PREDICTION TECHNOLOGY

Robert B. Rose
Naval Ocean Systems Center
San Diego, CA

ABSTRACT

The middle 1970's saw the inception of a new real time hf propagation forecasting concept called PROPHET. Based on mini-computer technology, this system provided point-to-point propagation assessments which reflected the real time state of solar and ionospheric conditions. Over the ensuing years, PROPHET has grown both in capability and utility. Whereas the original PROPHET system was based on five simple algorithms, the newest Advanced PROPHET system features over 15 hf prediction and assment models. This growth, coupled with rapid advances in computer sciences had led to numerous permutations which vary in complexity from mil spec minicomputers to hand-held calculators. This paper presents an overview of PROPHET's growth in capability to the present time, and provides an assessment of its future role in hf communication, surveillance and signal exploitation.

BACKGROUND

In the early 1970's the decrease in solar activity brought on by the sunspot minimum between cycle 20 and 21 and the emergence of communication satellite technology brought about a de-emphasis in the utilization of the hf spectrum between two and 32 MHz as a primary communications medium. Hf equipment was not upgraded, and training of communicators in these types of communications was reduced. Systems using skywave propagation were relegated to a secondary role or neglected altogether. However, the expected continuance of a military signal intelligence mission (SIGINT) and a commercial shortwave broadcast requirement dictated that some competency be maintained in these technologies. Also during this period a Navy program was underway to develop military applications for real time solar/ionospheric sensing from satellite platforms. Paradoxically, the best spectrum to demonstrate the one-to-one cause and effect relationships between solar activity and ionospheric propagation was the hf spectrum. As early as 1972 it was recommended (ref. 1) that "....with a modest investment over the next several years, and using many existing networks, a program can be established that will provide ionospheric dependent systems with:

1. Real time propagation information
2. Solar disturbance forecasts
3. Solar disturbance alerts ...."

A concept was adopted to develop propagation forecast (PROPHET) terminals which would translate real time solar emission data from the SOLRAD and other solar sensing satellite into a real time hf propagation assessment terminal. Conceptually, this system would be able to predict the existence and quality of an hf skywave channel* between any two points on the globe for either quiet or disturbed solar conditions.

In the 1975-1977 time frame, three events took place which came together to provide the needed thrust for new hf technology development. As early as 1976, it was correctly predicted that solar cycle 21 would be much higher than cycle 20

*Skywave channel is defined as the band of usable frequencies simply bounded by the Maximum Usable Frequency (MUF) and the Lowest Usable Frequency (LUF).

(ref. 2), stirring a new interest in commcercial and military use of hf. Second, new concepts in "minimal modeling" led to the development of the MINIMUF hf MUF prediction model (ref. 3). Whereas classical MUF predictions required upwards of 185K bytes of computer code, MINIMUF used 80 BASIC statements. Subsequent accuracy verifications by correlating it with oblique sounder data have shown this simple model to have surprising accuracy (ref 3, 4). The similar approach to hf prediction model development was used for the QLOF model, used to predict the undisturbed Lowest Usable Frequency (LUF) and SIDGRID which relates the change in LUF to solar X-ray flux. Thus "minimal modeling" led to a greatly simplified but accurate prediction capability which provides an assessment of the hf channel between any two points. Third, ADP technology moved from mini's to micro's, providing more computing capability in smaller, more portable packages. These three events stimulated ideas on the feasibility of hf tactical forecasting being performed by simple algorithms and small calculators in the field on a real time basis. In the subsequent four years these ideas have been borne out. The remainder of the paper will describe how simple predictions to improve hf channel effectiveness have been expanded to provide new tools for a broad range of concepts for signal acquisition, denial and exploitation.

SOLRAD-PROPHET

The prototype PROPHET, called SOLRAD-PROPHET, was installed at NAVCOMSTA Stockton, California, in 1976. This was the forerunner of real time geophysically updated ionospheric assessment systems. It was designed to translate a continuous stream of solar emission, magnetospheric and ionospheric data into real time hf propagation assessments. The first test vehicle, shown in figure 1, was originally designed to use inputs from the SOLRAD-HI satellites for a one year developmental test and evaluation of improved hf frequency management and spectrum utilization using PROPHET products.

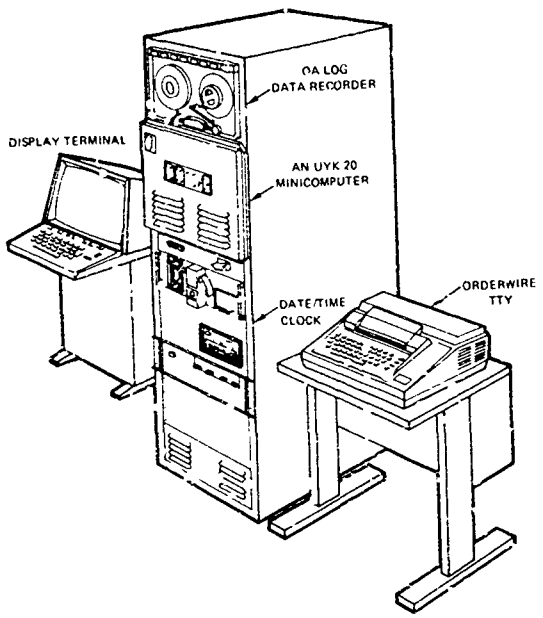


Figure 1. SOLRAD-PROPHET terminal.

Its utility caused it to be retained in operation through October 1980. Over that four year period, it successfully supported communications in EASTPAC, WESTPAC, the Antarctic and recently the Indian Ocean. In those instances where it was used for frequency selection and communications planning it was demonstrated that the real time PROPHET concept could solve simple channel availability questions, successfully mitigate adverse propagation effects, and improve overall frequency management. It was also shown that communicators using PROPHET became more proficient in establishing and maintaining hf circuits (ref. 5).

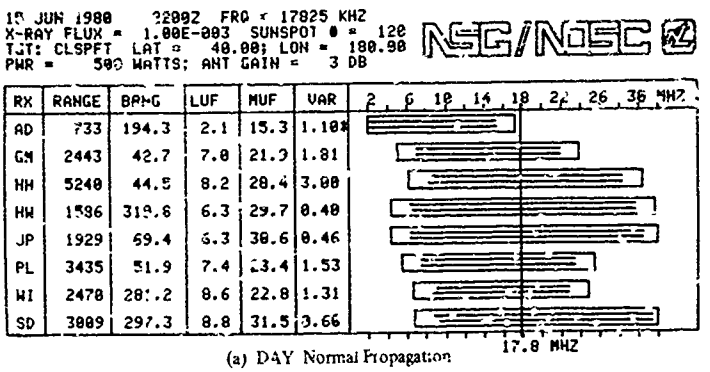
The existence of the SOLRAD-PROPHET capability led to exploration of new applications for this prediction technology, and simplified methods to quickly incorporate it into existing systems throughout the fleet. Concurrently, a steady R&D effort was maintained to assure continued technology advancement. The progress made in each of these areas will be discussed briefly.

PROPAGATION TACTICS

In 1977, it became apparent that the ability to rapidly predict many point-to-point propagation paths in near real time could be advantageous. A concept, called propagation tactics" was formulated which called for exploiting the ability to predict the skywave propagation environment to gain a tactical advantage. The advantage could be in the form of enhanced or degraded geolocation vulnerability, denial of a transmission to an adversary, reduced vulnerability to jamming. These benefits were derived from the predictability of certain skywave idiosyncrasies and were exploitable in addition to optimizing channel effectiveness

CLASSIC PROPHET

Classic PROPHET was the forerunner of desk top, portable non-real time PROPHET terminals. Designed around a 32K, 8 bit microprocessor graphics computer, the system was deployed in 1978 to support Navy HFDF operations. It was structured to allow the Net Control Officer (NCO) to assess propagation conditions from his out-stations to a given target area. Figure 2 illustrates one product which presents the band of frequencies between the MUF and LUF from each DF net station to a specified target area.



(a) DAY Normal Propagation

Figure 2. CLASSIC PROPHET Multi-Station HF Prediction

Concurrently, the development of a new line of bearing (LOB) variance algorithm based on propagation factors provided the HFDF NCO with an assessment of probable HFDF bearing quality for each net station as a function of time of day. An example is shown in figure 3 where it is seen that the new LOB variance is realistically varied by normal day/night propagation influences : Classic PROPHET was successfully tested in 1979. Additionally because of its simplicity of use on readily available desk top graphic computers, it has been successfully employed by numerous other agencies for simple hf point-to-point prediction.

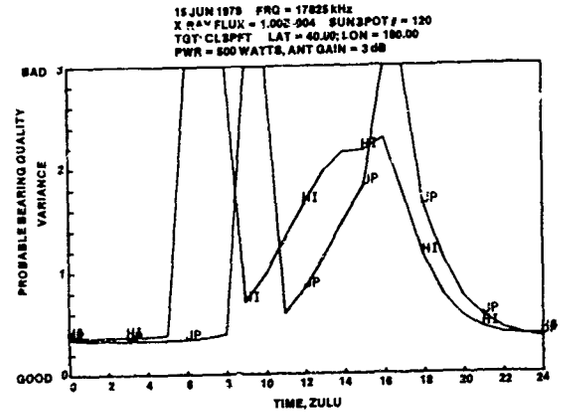


Figure 3. HFDF Line of Bearing Variance as a Function of Time of Day

TACTICAL PREDICTION MODULE

While Classic PROPHET was designed primarily to support HFDF problems, the Tactical Prediction Module (TPM) was designed to address signal security (SIGSEC) and communication security (COMSEC). Designed around a 64K, 16 Bit desk top graphics computer, it was conceived to support hf communications effectiveness in a hostile SIGINT and EW environment. The most current permutation, Version 3 shown in figure 4, represents a third generation refinement designed for a specific communications need and for use anywhere in the world. Figure 5 shows one of its 16 output products. Shown in the 24-hour prediction of the hf channel between two desired points (light area). Embedded are contours of the signal margin(d3) the desired transmitter will have over a jamming transmitter at the receiver site. This system is scheduled for field testing in the spring of 1981.

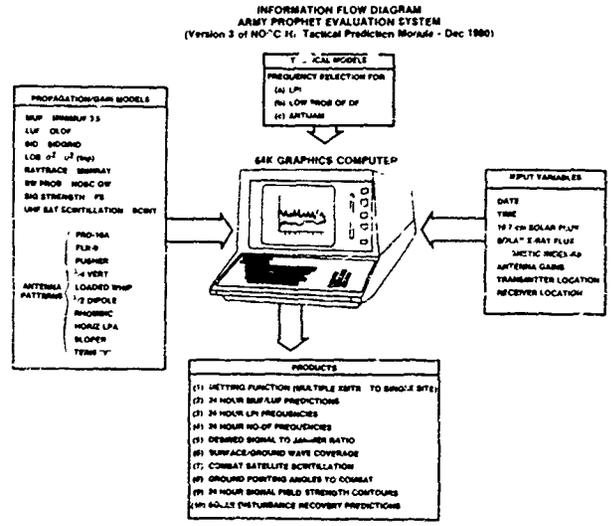
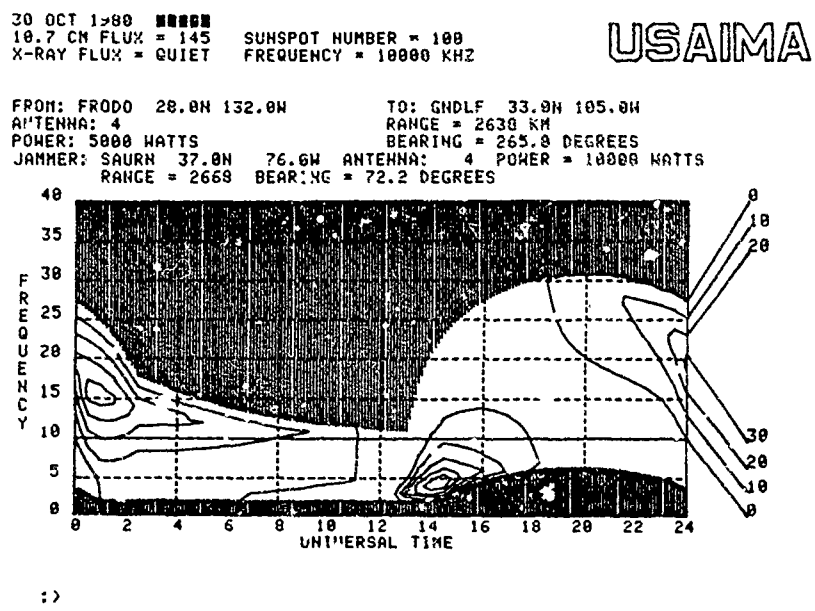


Figure 4. Information Flow Diagram Army Prophet Evaluation System



:)

Figure 5. Desired Signal Over Jammer Signal Contours

ADVANCED PROPHET

While the present versions of the Tactical Prediction Modules contained disturbance assessment models which are designed to accept data from near real time keyboard entry, a concerted developmental effort is continuing on improving fully real time systems which used solar emission data from satellites. The focal point of this effort is the Advanced PROPHET Development System located at NOSC. The objective of Advanced PROPHET is to maintain a fully supported test bed for the 6.1/6.2 environmental forecast technology R&D program maintained by NAVMAT. This system has a dedicated real time hf propagation prediction system which is driven by the NOAA GOES satellite solar sensors and global ionospheric sensors from the Air Force Astrogeophysical Teletype Network. The products always reflect present geophysical conditions as they are. The models are automatically updated with solar X-ray flux, 10.7 cm solar flux, proton flux and geomagnetic indices kp. The 25 propagation prediction models in Advanced PROPHET are maintained to develop, demonstrate and verify new prediction techniques and to test new signal exploitation models for hf, vlf and transionospheric systems. Figure 6 shows a 24-hour field strength prediction useful in assessing the signal strength effectiveness of an hf circuit. The use of a simplified field strength model coupled with a modified noise model allow a 24-hour predictionof MUF/LUF boundaries along with embedded signal to noise contours. While the example shows a benign mid-latitude path, proper modifications from other elements of Advanced PROPHET allow outputs such as this to reflect high latitude absorption or solar flare effects.

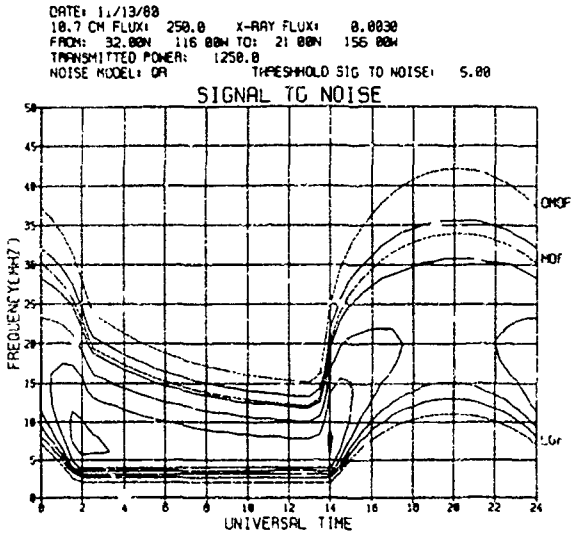


Figure 6. Advanced PROPHET Field Strength Product

Figure 7 demonstrates the ability of Advanced PROPHET to predict the location and extent of the auroral oval as a function of solar activity.

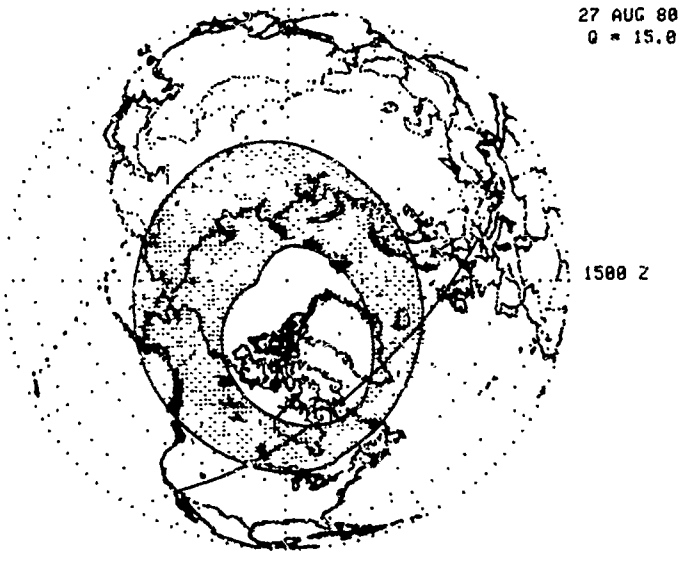


Figure 7. Advanced PROPHET Auroral Oval Location Output

This is particularly important in making an accurate assessment of high latitude propagation conditions, because special auroral zone hf absorption models have to be used when the oval is active. In addition to recent improvements in high latitude prediction, the Advanced PROPHET program is also focusing on equatorial and transionospheric phenomena. Table 1 provides a summary of a representative sample

of the models currently residing in this system. Presently, its capabilities are nowhere duplicated in an operational sense, although model development is influenced by certain operational or new application needs. For example, parts of TPM version 3 were initially developed on Advanced PROPHET, verified, and then simplified to BASIC for use on TPM. It is anticipated this approach will continue to be used into the foreseeable future.

TECHNOLOGY TRANSFER

Because of the simplicity of the models and the success of the original PROPHET system several spinoffs have emerged to satisfy certain short term special needs

FAA-PROPHET - A simplified version of PROPHET was installed at the FAA Flight Service (FSC) in Anchorage, Alaska, to test the utility of PROPHET predictions in the North Pacific. This terminal was based on a LSI-11 microprocessor with a built-in floppy disk and CRT. The system was driven by real time X-ray, 10.7 cm flux and solar proton flux obtained over a data link from the NOSC Real Time Geophysics Laboratory. The objectives of this permutation were to aid in frequency and antenna selection for comunications between Alaska and enroute aircraft in the North Pacific. The test objectives were met.

FOTACS - This is a frequency management and communication planning system. The MINIMUF, QLOF and SIDGRID models are used to provide simplified assessment of hf channel availability. One real time version exists at the NAVCOMSTA Stockton where it replaced the original SOLRAD-PROPHET system in October 1980. Similar systems are scheduled for employment to all Navy NAVCAMS.

COPS-MOD. B - The Communications Planing Support Program (COPS). This system, first devised in 1979, currently supports intra task force communications. Only MINIMUF and QLOF models are used to assess the availability of a skywave channel.

One measure of effectiveness of a system or model is the extent it is used outside its original developmental environment. One example of this type of hf community acceptance has occurred with the MINIMUF model which was published in 1978 in BASIC and Fortran versions. Since its introduction, versions exist which will operate on Tektronix 4051/4052, Wang, HP 9845, Apple II, Commodore Pet, North Star, TRS-80, HP-67 and TI-59 calculators. It is in use by the Army, Navy, Air Force, FAA, commercial short wave broadcasters, and industrial users. It is known to be in use in numerous foreign countries, serving both military and civilian interests.

THE FUTURE OF HF FORECASTING IN EM WARFARE

PROPHET technology is user tested and accepted. Future advances will provide the hf user with even more information which will allow him a greater amount of control over the effects of propagation on his or his adversary's system's performance in either a hostile or benign EM environment. While he can presently directly address channel effectiveness, future systems will concurrently support signal security, exploitation, surveillance and counter surveillance activities. PROPHET will predict the ability of interfering signals to degrade reception in advance. The user will have tools to establish propagationally covert channels or engage in active electromagnetic warfare - both covert and overt. The information to accomplish this will be produced on site or on platform, reflect real time ionospheric and solar conditions and will be structured to allow periodic upgrade as new forecast technology becomes available.

While the future is one of application, hf R&D program objectives include continuing a viable 6.2 development program to improve prediction timeliness and accuracy. New tactics in EW, COMSEC and SIGINT will be refined and tested. Finally, future efforts will continue to assure PROPHET technology is incorporated into new communications architecture.

Model	System	Action	Status
Flare detection	All hf, vlf navigation and comm	hf comm-freq shift reroute traffic	Operational
Flare detection	all hf, vlf nav/comm	hf comm-freq shift reourte traffic	operational
SID GRID	all hf	hf comm freq shift reroute traffic	operational
SPA/vlf	vlf nav Omega	phase correction factor	developed
SPA inversion	all hf, vlf	estimate x-ray flare size (independent of satell.) feed sid grid	in progress
PCA/vlf	vlf navig	phase correction factor for transpolar circuits	developed
PCA/hf	all polar hf	hf comm-advice signal strength loss-freq shift	developed
PCA/vhf	all polar satellite	vhf comm-advice signal loss	developed
QLOF	all hf	hf comm-normal operations, freq management	operational
LOF split	covert hf systems	opt freq selection against known rcvrs	operational
MINIMUF-5	all hf	hf comm-normal ops freq management	operational
15 min update to MINIMUF using auroral E fields	all hf	correct MUF est. (real-time) minimize errors to ± 1 MHz) (feeds MINIMUF)	in progress
RAYTRACE	all hf	hf comm-normal ops. antenna selection	operational
Launch angle multipath using quasi parabolic	all hf	hf comm-normal ops. antenna selection	operational
Polar and auroral ionosphere	all hf vhf sattelite	hf comm-normal ops. polar circuits	in progress
Earth's magnetic field variations (ground)	ASW & any magnetically sensitive	corrections for field changes D_{st} and AE	in progress near completion
Mixing shock front from auroral disturbances	all hf	hf comm-midlatitude (feeds MINIMUF)	in progress
Scintillation grid	vhf/uhf satellite comm	advisory-dB fade probability based on location	operational
Omega correction factors	Omega vlf	correction factors	operational
MF FIELDS	hf	Diurnal MUF/LUF predictions with simplified field strength approximations	Operational
Ionospheric storm	hf	opt freq selection dw to propagation changes	in progress
Ionogram	hf	optimum frequency selection	operational

Table 1. Advanced PROPHET Models

References

1 Rose, R. B., M. P. Bleiweiss and W. K. Moison, DoD and Civilian Systems affected by Solar Activity Emission: SOLRAD Applications Study, Task 1, Naval Electronics Laboratory Center, NELC TR 1839, p 42, 31 August 1972.

2. Hill, J. R.,"Long Term Solar Activity Forecasting Using High Resolution Time Spectral Analysis," Nature, 266, 151-153, March 10, 1977.

3. Levine, P. H., R. B. Rose and J. N. Martin, "MINIMUF-3: A Simplified HFMUF Prediction Algorithm," International Conference on Antennas and Propagation, Part 2 - Propagation, 28-30 November 1978. London IEEE Conference Publication Number 169, 161-167, 1978.

4. Sailors, D. B., "Accuracy of HFMUF Prediction," paper presented at URSI, 12-15 January 1981, Boulder, Colorado.

5. Rose, R. B., P. H. Levine, Satellite Sensing of the Solar-terrestrial Environment and Real Time Radio Propagation Forecasting, (Executive Summary), NOSC Technical Report TR 487, p 37-40, Naval Ocean Systems Center, San Diego, 30 September 1979.

HF PROPAGATION ASSESSMENT STUDIES OVER PATHS IN THE ATLANTIC

D. R. UFFELMAN
E. O. HULBURT CENTER FOR SPACE RESEARCH
NAVAL RESEARCH LABORATORY
WASHINGTON, D.C. 20375

ABSTRACT

In September, 1980, and February, 1981, NRL utilized Barry AN/TRQ-35 oblique sounding equipment which was deployed on board platforms operating in the Atlantic Ocean in order to test a concept to provide a real-time model update with the oblique sounder as a data source. The purpose of this work was to show that very accurate HF channel assessment was possible over paths in the local vicinity of the sounded path. Presented are initial results of two exercises encompassing both short paths (less than 900 km) in the North Atlantic and medium length paths (900-2000 km) in the middle latitude Atlantic. These initial results indicate that by utilizing the oblique sounder to update the MINIMUF 3.5 model developed by NOSC, a significant increase in the accuracy of short term forecasting of the Maximum Usable Frequency (MUF) can be obtained. Over a 24-hour period, RMS errors from this model driven by the 5-day running average of 10.7 cm flux were approximately 4 MHz. Utilizing the oblique sounder as an update tool, this RMS error dropped to near 1 MHz. Further results are required, however, to validate the technique over longer periods of time, different geographies, and larger operational areas.

INTRODUCTION

Workers who are well acquainted with using the HF band (3-30 MHz) for communications know that in a low noise environment a reliable HF channel can be maintained at reasonable data rates provided the proper frequency, bandwidth, and antenna are selected. In fact, if these items are optimized it is reasonable to expect reliabilities approaching satellite links. Unfortunately, because of the nature of the beast, the information bandwidth is not as great as with satellite systems.

In the past, the task has been taken to devise equipment and techniques to improve the system performance under poor channel conditions. Work devoted to this end include improved antenna design, increased transmitter power, modulation and detection techniques commensurate with the medium, diversity techniques, and error correction techniques. These methods generally offer some improvement, but the performance of the channel generally remains below that which is desired. A substantial improvement, however, can be gained by selecting the optimum channel between the transmitter and receiver. Oblique sounding and techniques relying on it play an important role in assisting in this optimization. The work presented herein suggests a novel use of the oblique sounder to provide an update to a model which determines the characteristics of the HF channel between two points. It is shown that, at least in the initial work, significant improvement in the RMS error between actual channel data and the model occurs employing this model update scheme.

BACKGROUND

General

NRL initiated an examination of the potential for updating various ionospheric propagation model calculations for Navy uses when presented with the possibility that the Air Force might launch an advanced technology topside ionospheric sounder as part of its environmental assessment program. In order to determine the success of each model update procedure, it was necessary to directly assess the HF channel in order to have a basis for comparison. In the process of designing experiments to acquire the type of data needed to test the update hypothesis, NRL

attempted to take advantage of various scheduled operations which employ oblique sounders to do HF channel assessment. The equipment currently being utilized in these operations is the Barry AN/TRQ-35 Tactical Frequency Management System (TFMS). It is from these data that a direct assessment of the channel can be made and results using topside update may be compared.

In the process of doing the experimental design it was decided to also attempt the update directly from the oblique sounder network. In order to test this idea, at least two simultaneous sounder links had to be operating such that one path could be used as a source for the update and the second path, an arbitrary unknown path in practical applications, could be the path against which the updated model calculation was compared. After several opportunities to collect appropriate data were cancelled, NRL finally participated in the NATO Teamwork '80 exercise in September of 1980. Since topside sounder coverage was unavailable directly overhead in the area of Teamwork '80, the primary objective of this exercise was to obtain the type of oblique sounder data against which the model update could be performed using the oblique sounder data directly. The model update using oblique sounder data as a source, therefore, is the subject of this paper.

The Equipment

The data obtained from the TRQ-35 ("Turkey-35") receiver is in the form of Polaroid photographs of the oblique sounding ionogram. For short paths, the sounder network is set up to operate between 2 and 15 MHz, and for long paths the limits are 2-30 MHz. A sounding requires four minutes and forty seconds to complete. The equipment has the capability to record as many as three different stations, with each transmitting in a different five minute time slice of a fifteen minute period. Figure 1 shows the sounder receiver display and the important parameters which are extracted from that data [7]. Three different parameters are scaled from these data. The first is the cutoff point for ionospheric propagation between the transmitter and receiver known as the maximum usable frequency (MUF). In the example shown here, the maximum usable frequency is 7.6 MHz. The second parameter scaled is the band of frequencies in which optimum transmission might occur. This is called the FOT band. The FOT band is defined as that region in the data where the signal strength is high and no multi-path is evident. In the example shown here, the FOT band occurs between approximately 3.7 MHz and 6.1 MHz with the best frequencies being near 6.1 MHz where absorption should be the lowest. The last parameter scaled is the lowest usable frequency or LUF. The LUF is the frequency where the low cutoff occurs and is typically determined by D-region absorption. In this case the lowest usable frequency appears to be slightly above 2 MHz.

The data shown in Figure 1 also contains a histogram at the top of the display indicating the signal strength at each of the received frequencies. The histogram actually is a plot of the receiver AGC. Since this oblique sounding equipment is user oriented and not research instrumentation, the full amount of information normally available by properly thresholding the receiver such that the noise floor is slightly evident, is not available here. However, for the communications problem being addressed with the instrument, it is only important to ascertain the MUF, FOT, and LUF. Hence, excepting the fact that the manpower overhead is high to record a complete set of data, the equipment is quite usable and the data is easy to scale.

DISCUSSION

The essence of the model update technique is illustrated by Figure 2. A model calculation of the diurnal variation of the maximum usable frequency for a circuit is performed using the appropriate parameters which drive the model. This step is illustrated by the line labeled "model application". The actual data is plotted against this computation and in most cases the model has a bias with trends tending to be agreeable. The next step is to extract a reading of the appropriate parameter (MUF for this case) at one point in time from the oblique sounder. The model is then forced to fit this point by varying the relevant driving parameters. This step is indicated by the center plot of the figure labeled "known path". The model that is currently being examined, MINIMUF 3.5, is fitted by varying the driving parameter which is a sunspot number. This sunspot number derived from the force fit is then used to generate like computations over other paths and these are compared directly with sounder data to determine how well the technique worked. This is illustrated in the figure by the plot labeled "unknown path".

One objective of this work is to ascertain the spatial and temporal perishability of a model update. This can be satisfied by testing the performance of various models over a large enough data base. The current model being tested is MINIMUF 3.5, but others are envisioned for testing in this manner such as IONCAP and HFMUFES. It is expected that the model update scheme will work well over a local area. If this can be verified, a single oblique sounder path in the mid-latitude Atlantic, for example, could be used to update the appropriate model and this updated model could be used to obtain accurate forecasts of the maximum usable frequencies of other links operating in

that area. Experimentally, of course, this is determined by actually having sounder measurements over paths other than the reference path. Experimental results will be shown in the following sections which indicate the initial success of this technique.

RESULTS

Mt. Whitney Operations; September 18-19, 1980

In the past, obtaining the type of data required to validate this technique has been difficult. Data is required from two or more oblique sounder paths operating simultaneously in a local area. One path is used to update the model and other paths can be used to check the success of the update. NRL's first opportunity to obtain data of this kind occurred in September, 1980. An NRL representative went on board the USS Mt. Whitney to obtain oblique sounder data from a net which was being used to support the NATO Teamwork 80 exercises. Figure 3 is a map of the experimental setup in the September 18-19 time frame. A sounder transmitter was located at Soc Buchan, Scotland, labeled T_1; Kolsaas, Norway, labeled T_2; and Orland, Norway, labeled T_3. The USS Mt. Whitney, on-board which the receiver was located, was anchored off the coast of Norway. This is denoted by R. Path lengths are 830 km (T_1-R), 340 km (T_2-R), and 1044 km (T_3-R). Since NRL could have only one representative on-board to gather data and the data had to be obtained in the form of Polaroid photos, it was extremely difficult to get a continuous set of data for the full exercise period. However, on the 18th and 19th of September, the technician worked a straight 24-hour period in order to obtain one complete day of data. Partial days are also in the data base for the period 9 September - 23 September. The existence of one complete diurnal cycle was the determining factor in the selection of the sample of data on which the update technique was first tested.

The Unupdated Model

To determine the improvement gained by employing this technique, the model was run in its standard recommended mode [4] whereby the five day running average of 10.7 centimeter flux was used to drive the model. This model calculation was compared with the actual maximum usable frequency as scaled from the data and the RMS error was computed. Figure 4 shows this comparison. The vertical lines indicate the difference between the model and the actual maximum usable frequency as measured every 15 minutes by the sounder over the Soc Buchan to Mt. Whitney path (T_1-R). Note that the difference between the model and the actual data has an RMS error of 3.82 MHz. This is in good agreement with the advertised accuracy of MINIMUF 3.5 [4]. Figure 5 is the same type of calculation, but for Kolsaas, Norway to the Mt. Whitney. The calculated RMS error here is 4.03 MHz.

The Updated Model

Next, the update was attempted using the Kolsaas path as the source path. The model was forced to fit the measured MUF on that path at 0600Z. From the forced fit, a sunspot number was extracted and used to drive the model for the Soc Buchan to Mt. Whitney path. The results from these updates are shown in the next two figures. Figure 6 indicates the model update calculation for the Soc Buchan path whereby the Kolsaas path was used to derive the applicable sunspot number. In this case, note the marked improvement in the model fit to the data. The RMS error was dropped to 1.64 MHz with the largest portion of that error occurring in the evening when scattering phenomena led to increased scaled maximum usable frequencies.

Figure 7 indicates the improvement in the Kolsaas path model calculation. Here the RMS error is 2.87 MHz and again, most of that error is due to the scatter propagation which occurred.

Since the MINIMUF model has no capability to predict scatter, the data was rescaled to remove the scatter. Figure 8, which is an example of the actual data, indicates how this could be done. First note the "extended nose" of the ionogram which is attributed to scattering phenomena. Working back in the ionogram, one can see the standard portion of the ionogram where o-x splitting occurs. Hence, it was relatively simple to identify the point where the standard MUF exists.

The next two figures show the result of the update with the scatter removed. Figure 9 is the longer Soc Buchan to Mt. Whitney path. The calculated RMS error has dropped to .85 MHz indicating an extremely good fit to the actual data. Figure 10 is the comparison of the Kolsaas to Mt. Whitney path. It also shows a .85 MHz RMS error, again indicating a good fit.

An additional subtlety should be noted in this analysis. There was an approximate two hour shift to the left of the MINIMUF model calculation relative to the actual data. This was removed.

At this point no attempt is made to explain this shift. However, the problem is simply handled by matching the model sunrise with the data sunrise and then performing the fit from that point. It has been indicated that September-October MINIMUF model calculations at high latitudes over short paths have shown this shift in the past and at present this is not understood. It is suspected, however, that at high latitudes near equinox, there is a seasonal effect which is responsible for this result (2).

Table I below is a summary of the data examined so far indicating the improvement in the forecast using the model update for the September 18-19 time frame with the scatter removed. A number of different computations were attempted. Situation #1 is the case where the MINIMUF model was driven using the actual five day running average of 10.7 centimeter flux. The RMS error for each path is calculated. It was noticed during the analysis that the 10.7 centimeter flux and the sunspot number which NRL had access to did not convert by the algorithm in MINIMUF. Hence, situation #2 was generated using the five day running average of sunspot number. The third case used the one day sunspot number since February 18 was more active than the previous four days. The RMS error declined further. As mentioned earlier, the MINIMUF calculation was shifted relative to the actual data. Hence, MINIMUF was shifted to remove this error and situation #4 reflects that calculation with the same driving parameters as situation #1. There is only about a .4 MHz RMS error difference due to the shift. Likewise, situation #5 is comparable to situation #2 and situation #6 is comparable to situation #3. Finally, situations 7-14 indicate the results using a reference path to perform the update. For example, situation #7 utilizes a shifted MINIMUF, where the model was forced to fit on the Soc Buchan path at 0600Z. The resulting sunspot number was then used to derive the Soc Buchan path, and the Kolsaas path. RMS errors are indicated. Several different update times are investigated where situations 8-10 are updates using the Soc Buchan path at 0800Z, 1000Z and 1200Z respectively.

In order to obtain an indication of the dependence of the technique on path differences, the same update scenario was applied using Kolsaas as the reference path. Situations 11-14 show the RMS errors using that path as the reference with fits derived at 0600Z, 0800Z, 1000Z and 1200Z respectively.

TABLE I

IMPROVEMENT IN FORECAST USING UPDATE

Accuracy (RMS Error, MHz) of Update Technique
For Sept 18-19, 1980 Data With "Nose Extension" Removed

No.	SITUATION	FIT TO RAW DATA		SSN
		SOC BUCHAN	KOLSAAS	USED
1	Actual MINIMUF Using 5 day 10.7 cm Flux	3.36	2.76	112
2	Actual MINIMUF Using 5 day SSN	2.65	2.22	149
3	Actual MINIMUF Using 1 day SSN	1.98	1.72	197
4	Shifted MINIMUF Using 5 day 10.7 cm Flux	2.97	2.43	112
5	Shifted MINIMUF Using 5 day SSN	2.07	1.75	149
6	Shifted MINIMUF Using 1 day SSN	.99	.95	197
7	Shifted, 0600Z Fit at Soc Buchan	.66	.63	227
8	Shifted, 0800Z Fit at Soc Buchan	.75	.67	249
9	Shifted, 1000Z Fit at Soc Buchan	.75	.67	249
10	Shifted, 1200Z Fit at Soc Buchan	.75	.67	249
11	Shifted, 0600Z Fit at Kolsaas	.85	.85	205
12	Shifted, 0800Z Fit at Kolsaas	.64	.62	240
13	Shifted, 1000Z Fit at Kolsaas	.66	.63	227
14	Shifted, 1200Z Fit at Kolsaas	.66	.63	227

To further validate this technique another operation of opportunity in the Atlantic was drawn upon in late February, 1981. The path lengths were 1000-2000 km and these were at mid-latitude. Preliminary results indicate that the update technique worked almost as well as the example shown herein with RMS error dropping to near 1 MHz for a 24 hour period. Those results will be reported in a future document.

SUMMARY

In the near future the U. S. Navy will be deploying a number of oblique ionospheric sounders to be used to improve HF long haul communications. In addition, a great deal of work is being done

with the PROPHET prediction system to determine effective use of propagation tactics as well as assess management based on the condition of the propagation environment. It is NRL's intent to examine the coupling of these two instruments into a system which will provide greatly increased propagation assessment and forecasting accuracies. Initial results indicate this is a reasonable idea and significant improvement may be gained by marrying oblique sounding and PROPHET-type systems. Although this technique has not been validated over a large variety of links under various conditions, initial indications are that this should be done to verify the possibility that large increases in short term forecasting accuracy might be effected.

CONCLUSION

RMS errors near 1 MHz between a model calculation and the actual maximum usable frequency for an HF circuit indicates an extremely useful technique for making accurate decisions regarding tactics which take into account the effect of the propagation medium. The model update work represented herein should be verified over various geographies at various times of the year. At the same time, other models used by the community such as IONCAP and HFMUFES should also be examined in this context. It is NRL's intent to do this as follow-on work. Verification of this technique on a general basis could lead to a major advance in the use of computerized assessment and forecast to employ propagation tactics and to effect greatly increased efficiencies in the use of frequency assets and channel reliabilities in the future.

ACKNOWLEDGEMENTS

Any technical endeavor requires the input and assistance of numerous people. The work represented here is no exception and several key individuals are acknowledged. First, Navy CAPT Henry Orejuella is gratefully acknowledged for initially supporting this idea and providing funding. Mr. Bill Juchsch at CINCLANTFLT and Mr. Ray Rozanski have also been key supporters. The data could not have been taken if it were not for the tireless efforts of Mr. Tom Priddy who rode the Mt. Whitney in September. Mr. Larry Harnish has been indispensable in scaling the data and writing analysis software. Finally, Mrs. Gailyn Nocente has been extremely patient with the author through numerous revisions written in what amounts to hieroglyphics.

REFERENCES

1. Jelly, Doris H., "Ionospheric Sounding as an Aid to HF Communications"; CRC Report #1225; Ottawa, Canada; December, 1971.

2. Rose, R. B., private communication.

3. Rose, R. B., "PROPHET-An Emerging HF Prediction Technology", IES-81; Washington, D. C.; April, 1981.

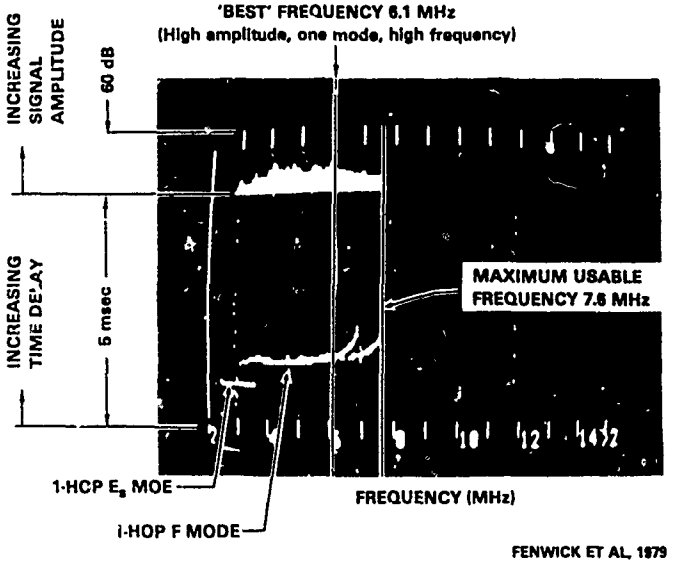
4. Rose, R. B. and J. N. Martin, "MINIMUF-3: A Simplified HF MUF Prediction Algorithm", NOSC TR 186; San Diego, California; 19 February, 1978.

5. Rose, R. B. and J. N. Martin, "MINIMUF-3.5", NOSC TR 201; San Diego, California; 26 October, 1978.

6. Rose, R. B., "Further Verification of the MINIMUF-3.5 HF MUF Prediction Algorithm for 1. Frequencies Above 32 MHz and 2. Path Lengths of Less Than 250 nmi"; NOSC TN 758; San Diego, California; 20 September, 1979.

7. Fenwick, Robert B. and T. J. Woodhouse, "Real Time Adaptive HF Frequency Management", Special Topics in HF Propagation, AGARD CP-263; November, 1979.

ANNOTATED SOUNDING FROM THE CHIRP SOUNDER



FENWICK ET AL, 1979

Figure #1

NRL APPROACH

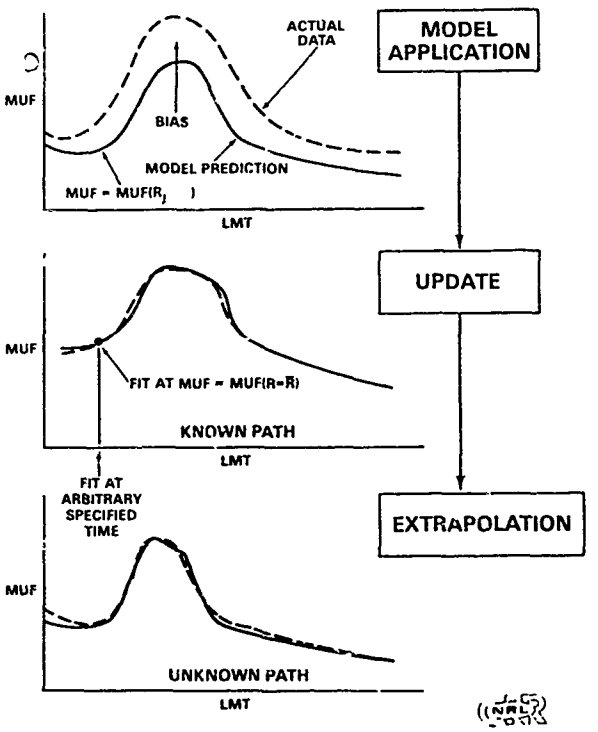


Figure #2

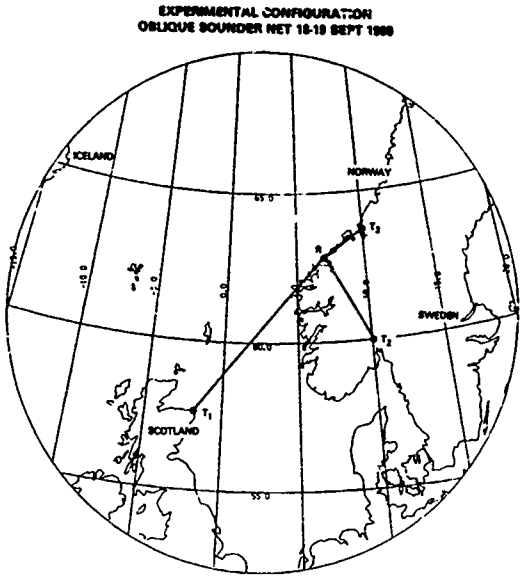


Figure #3

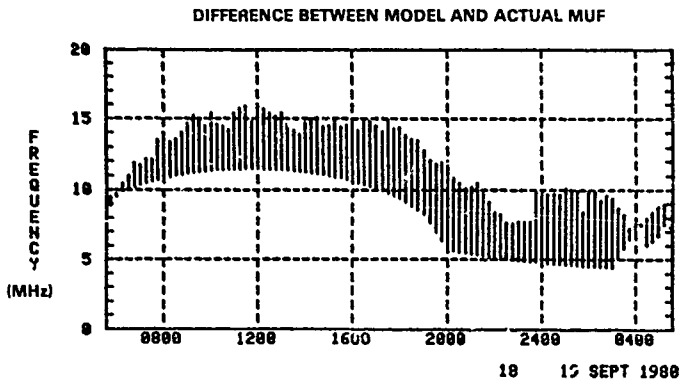


Figure #4

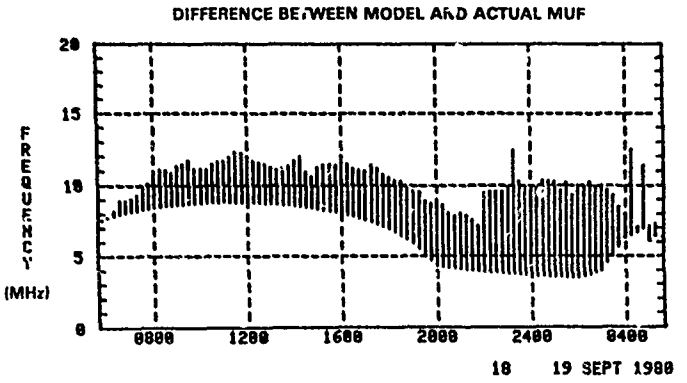


Figure #5

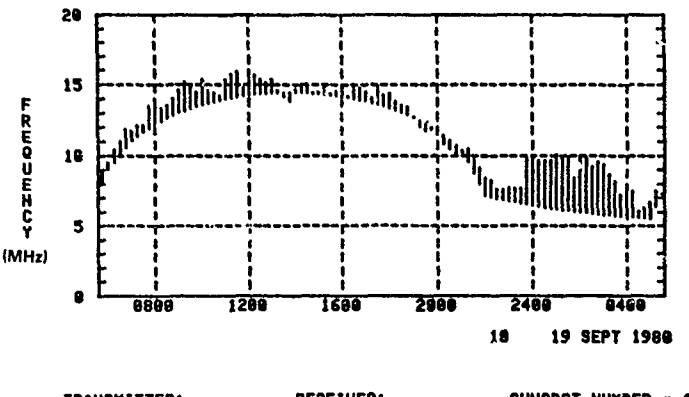
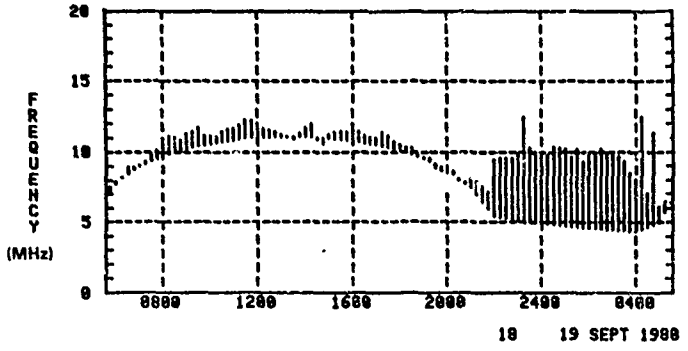


Figure #6

DIFFERENCE BETWEEN MODEL AND ACTUAL MUF



TRANSMITTER:
KOLSAAS, NORWAY
60.0N,10.3E

RECEIVER:
USS MOUNT WHITNEY
62.95N,0.34E

SUNSPOT NUMBER = 205
10.7 CM FLUX = 250

R.M.S. ERROR = 2.07MHZ

MINIMUF TIME DELAYED 2 HRS.

Figure #7

OBLIQUE SOUNDINGS SHOWING EXTENDED MUF
(NOSE EXTENSION)

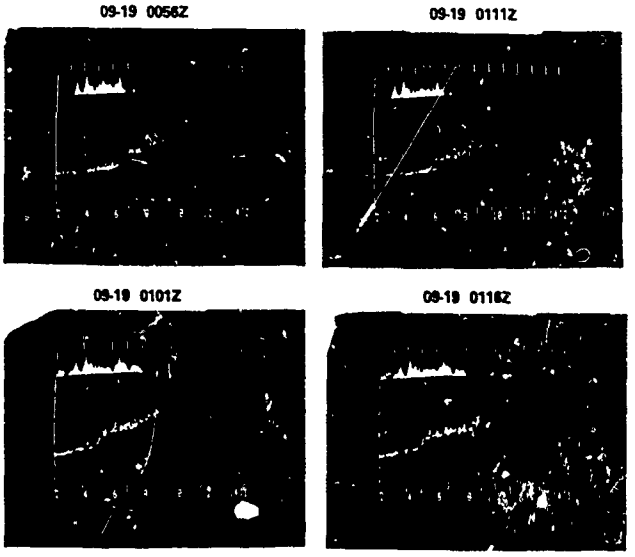
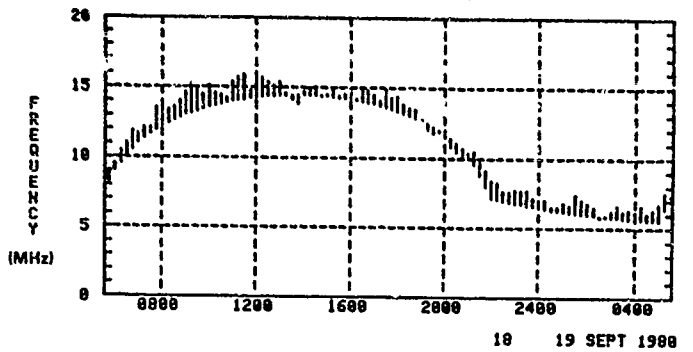


Figure #8

DIFFERENCE BETWEEN UPDATED MODEL AND ACTUAL MUF
(NO NOSE EXTENSION)



TRANSMITTER:
SOC BUCHAN SCOT
57.3N, 1.5W

RECEIVER:
USS MOUNT WHITNEY
62.95N,8.34E

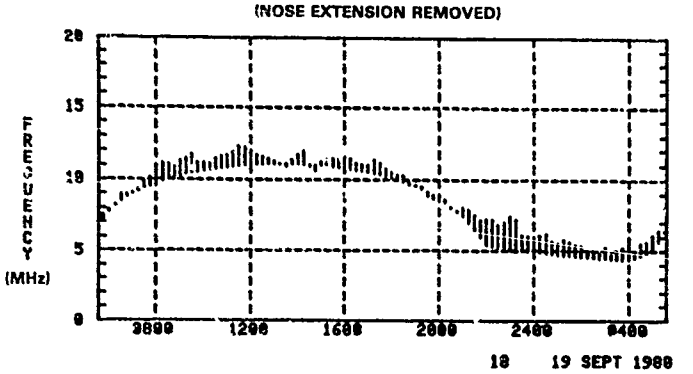
SUNSPOT NUMBER = 205
10.7 CM FLUX = 258

R.M.S. ERROR = 0.85MHZ

MINIMUF TIME DELAYED 2 HRS.

Figure #9

DIFFERENCE BETWEEN UPDATED MODEL AND ACTUAL MUF
(NOSE EXTENSION REMOVED)



TRANSMITTER:
KOLSAAS, NORWAY
68.8N,16.3E

RECEIVER:
USS MOUNT WHITNEY
62.95N,8.34E

SUNSPOT NUMBER = 205
10.7 CM FLUX = 258

R.M.S. ERROR = 0.85MHZ

MINIMUF TIME DELAYED 2 HRS.

Figure #10

A GLOBAL MODEL FOR WIDEBAND HF SKYWAVE PROPAGATION

A. Málaga

SIGNATRON, INC.
Lexington, MA 02173

1. INTRODUCTION

The HF channel is a dispersive and time varying channel whose characteristics depend on the operating frequency, range, geographical location, time-of-day, season and sun-spot number. Existing HF propagation models such as the ITS-78/HFMUFES model [1] are mainly concerned with the prediction of the optimum operating frequency (FOT) as well as the HF skywave signal strength and to a lesser extent with the multipath structure of the received HF signal. This type of information is suitable for narrowband voice and low data rate transmission HF systems which perform optimally near the maximum usable-frequency (MUF) where the multipath is minimal. However, modern communication techniques which employ digital signalling, wider bandwidths and adaptive methods to combine multiple skywave returns coherently can perform better under strong multipath conditions. Hence in addition to the received signal strength, information about the number of returns, their relative delays, delay dispersion, Doppler shift and Doppler smears is required in order to design and properly plan a point-to-point HF communications link. In this paper we describe an HF propagation model suitable for the prediction of all of these parameters as a function of operating frequency, range, geographical location, time-of-day, season and sun-spot number. The diurnal, seasonal, geographical and solar-cycle variations in the various parameters are obtained from their relationship to diurnal, seasonal, geographical and solar-cycle variations in the critical frequencies and heights of the D, E, F1 and F2 layers.

HF propagation prediction models such as ITS-78/HFMUFES [1] make use of worldwide predictions of monthly median values of the critical frequency of the F2 layer (f_0F2) published by NOAA [2] to predict the monthly median value of the received signal level. Hour-to-hour and day-to-day variations are obtained from distributions of signal strength about the monthly median value. These distributions have been derived from a large data base gathered under a wide range of conditions, and show a dependence on geomagnetic latitude, season, local time and path length. However a similar data base for the multipath and fade rate (Doppler) parameters is not available. On the other hand, it is well known [3] that the delays and Doppler shifts of the various skywave returns depend primarily on the heights and critical frequencies of the E, F1 and F2 layers, while their strength (amplitude) depends primarily on the absorption characteristics of the D layer. Hence a worldwide HF propagation prediction model can be established from a model which describes the diurnal, seasonal, geographical and solar cycle variations of the D, E, F1 and F2 layers.

The diurnal, seasonal and geographical dynamics of the D, E, and F1 layers can be modeled quite well [4] as a function of a single parameter, cos χ (cosine of the instantaneous solar zenith angle), proportional to the solar intensity. However the same is not true of the F2 layer which exhibits diurnal, seasonal and equatorial anomalies [5] when compared to expectations based on cos χ variations. For example, the F2 layer critical frequency, f_0F2, can be higher at mid-day in winter than in summer (seasonal anomaly), and on a given day can peak in late afternoon rather than at mid-day (diurnal anomaly). Unlike the E and F1 layers, the critical frequency of the F2 layer is not maximum at the equator but rather it exhibits a trough (equatorial anomaly).

Although f_0F2 cannot be modeled as a function of the instantaneous cos χ, Levine, et al [6] have proposed to model the diurnal and seasonal variations of f_0F2 as a lagged response to the instantaneous solar intensity cos χ. Allowing the lag time constant to be long ($\simeq$ 10 hours) in summer and short ($\simeq$ 1 hour) in winter at middle and equator-

ial latitudes, they are able to reproduce the seasonal anomalies. Furthermore by choosing the lag time constant during the day to be a simple monotonic function of the mid-day solar zenith angle, and constant at night (2 hours) independent of season or geographical location, they are able to reproduce the diurnal behavior of f_0F2. We make use of this F2 layer dynamics model and the Chapman model (cos χ) description of the D, E and F1 layers dynamics in order to predict the diurnal, seasonal and geographical variations of the multipath skywave propagation parameters.

In the remainder of this paper we discuss the relationship between the various parameters of the HF channel and the ionospheric layer parameters and their variations.

2. HF CHANNEL CHARACTERIZATION

At frequencies in the HF band (3-30 MHz), the composite received signal at distances beyond the horizon consists of a signal which propagates along the ground (ground wave) and a number of ionospherically reflected signals (skywaves). When propagation is over land, the ground wave is stronger up to distances of 80-100 km during the day time hours and 30-50 km at night. When propagation is over sea-water the ground wave predominates up to distances of 200-250 km during the day and 100-150 km at night.

Because the ionosphere is an inhomogeneous medium, i.e., it consists of three main layers varying in altitude from 100 km to 400 km and known as the E, F1 and F2 layers, the skywave return consists of several signals each of which has been reflected at a different height. The number of signals reflected from each layer depends on whether the operating frequency is above or below the maximum-usable-frequency (MUF) of the layer, i.e., the maximum frequency for which one can expect a return from that layer. If the operating frequency is below the MUF of a layer but close to it, we can expect two returns from this layer known as the high ray and low ray. In addition each of these returns may be decomposed into two magnetoionic components, known as the ordinary and extraordinary rays, which follow somewhat different paths. The high ray returns are usually more heavily attenuated than the low ray returns while the magnetoionic components may be indistinguishable from each other because they have nearly equal delays. Thus one may often observe only a single return from a layer. Because each return is reflected at a different height, they are delayed relative to each other as well as relative to the ground wave if present. Aside from this type of multipath, one may also observe multipath due to skywaves which have been reflected from the ionosphere once (one-hop) and skywaves which have been reflected twice and three times (two-hop and three-hop skywaves). On very long paths (> 3000 km) one may only receive two-hop and three-hop returns. The relative delays between the various one-hop returns may vary from tens to hundreds of microseconds while the relative delays between the ground wave, one-hop, two-hop and three-hop skywave returns when present may vary from hundreds of microseconds to a few milliseconds.

In addition to multipath time dispersion, HF signals also exhibit frequency dispersion also referred to as fading. The fading is due to the dynamic nature of the ionosphere. The ionization levels of the E, F1 and F2 layers change constantly with time of the day, month, season and from year-to-year. The heights of the E, F1 and F2 layers also change from hour-to-hour, month, season, etc. Because of diurnal changes in the heights of the layers and the electron content in each layer, each of the ionospheric returns exhibits a Doppler shift. Those signals reflected from the higher layers (F1 and F2) exhibit larger Doppler shifts than those reflected from the E-layer, and two-hop returns have larger Doppler shifts than one-hop returns. During the morning hours when the height of the layer decreases, the Doppler shifts are positive while in the afternoon and evening the Doppler shifts become negative as the height of the layers increases. The relative Doppler shifts between the various returns due to diurnal changes in the ionosphere vary from .01 Hz to about one Hz. Larger Doppler shifts may occur during periods when ionospheric disturbances (TID's) are observed.

During 'quiet' times when there is little solar, magnetic and weather activity the ionosphere continues to fluctuate randomly in addition to its regular diurnal and seasonal changes. These random fluctuations are caused by irregularities in the ionization profile of the ionosphere and result in frequency spreading (smearing) of each of the ionospheric returns. In general, the higher the signal is reflected in the ionosphere the more frequency spreading it exhibits because it must travel through more irregularities. Frequency spreading of the individual returns may vary from .01 Hz to a few tenths of one Hz at mid latitudes and higher over auroral and equatorial paths.

The preceding description of HF propagation phenomena indicates that we may represent the HF channel as a linear random time varying channel whose time varying transfer function (response to a complex sinusoidal excitation) is given by

$$G(f,t) = \sum_{i=1}^{N} A_i c_i(t) \exp\{j2\pi[\nu_i t - \tau_i f]\} \quad (1)$$

where N is the number of returns including the ground wave if present, τ_i is the delay of the ith return, ν_i is the Doppler shift, A_i is its amplitude. The $c_i(t)$ are normalized random time-varying multiplicative factors whose statistics describe the short term fading of the individual returns and their autocorrelation determines the coherence time and frequency spreading of the various returns. The amplitudes, A_i, delays, τ_i, and Doppler shifts, ν_i, vary slowly with time, i.e., exhibit diurnal and seasonal variations.

As an example we have plotted in Figures 1, 2 and 3 the delays (relative to the ground wave delay), Doppler shifts and attenuations of all the possible returns (modes) for a 1300 km link whose mid-path point is located at a latitude of 40° north and its longitude is 100° west. The local time at mid-path is 16 hrs, the month is December and the sun-spot number is 150. The delays, Doppler shifts and attenuation (path loss) are plotted as a function of frequency. Each trace corresponds to the relative delay (Figure 1), Doppler shift (Figure 2) or attenuation (Figure 3) of a particular mode. The number in front of the mode label (E, F1 or F2 mode) denotes the number of hops between transmitter and receiver while the subscript 0 or X indicates whether it is an 'ordinary' or 'extraordinary' mode. Modes 50 dB weaker than the strongest mode are not shown. Some of the traces are double valued, i.e., at each frequency there are two possible values of delay, Doppler shift and attenuation. One value of delay is that of the low ray while the other is that of the high ray; similarly for the Doppler shift and path loss traces. These curves show that at some frequencies (8-14 MHz), the strongest mode ($2F2_0$ or $1F2_0$) is not necessarily the first mode to arrive ($1E_0$). Furthermore, the delays of some of the modes (F1 and F2 modes) exhibit a strong frequency dependence. This frequency dependent delay results in delay dispersion of the individual modes when the bandwidth of the transmitted signal is large enough.

Wideband channel representation

When the bandwidth of interest is sufficiently large, say 10 KHz, the delays τ_i in the time-varying transfer function should be more appropriately modeled as

$$\tau_i = \zeta_i + b_i f + \delta_i \quad (2)$$

where the ζ_i are the average delays at the center frequency, the b_i are the slopes of the delay traces in Figure 1 (at the center frequency), which determine the amount of delay dispersion, and the δ_i are zero-mean random constants which account for random delay effects due to the same irregularities which cause Doppler smear. This representation of the delay assumes that the delay of each return varies linearly with frequency within the bandwidth of interest.

The curves of Figures 1, 2 and 3 apply to a particular link, for a particular time of day, day, month and sun-spot number. The delays, Doppler shifts, attenuations and the random gains $c_i(t)$ vary from hour-to-hour, day-to-day, seasonally, etc. Hence it is important to be able to predict the extent of these variations. In particular, the diurnal variations are directly responsible for the Doppler shifts as well as changes in the MUF (19.9 MHz in Figures 1, 2 and 3). Thus in the remainder of this paper we concentrate on establishing the relationship between the delay, Doppler shift and attenuation of the various modes and the diurnal, seasonal and geographic behavior of the ionosphere. The random gains also exhibit diurnal and seasonal variations; however we postpone a discussion of these to future work.

3. RELATIONSHIP TO IONOSPHERIC LAYER PARAMETERS

Without going into the details of ionospheric propagation which are well known [4,7], we now proceed to establish the relationship between the delays, Doppler shifts and amplitudes of the individual returns and the ionospheric layer heights and critical frequencies.

Delays

If we neglect the earth's magnetic field, the delay of a given ionospheric return can be calculated from

$$\tau_i = \frac{P'(\theta_i)}{c} \quad (3)$$

where $P'(\theta_i)$ is the virtual or equivalent path of a return whose angle of incidence on the ionosphere, measured from the vertical, is θ_i. The virtual or equivalent path is the path that a wave incident at an angle θ_i would follow from transmitter to receiver if its propagation velocity were constant and equal to the speed of light in free space c.

From the geometry of Figure 4, it can be seen that the virtual path P' for a K-hop return is given by

$$P'(\theta_i) = \frac{KD_K}{\sin \theta_i} \qquad (4)$$

where D_K is the horizontal or flat-earth range for 1/K of the path. This distance is related to the path great circle distance, D, over a round earth with radius of curvature R_e by

$$D_K = 2R_e \sin \frac{D}{2R_e K} \qquad (5)$$

These expressions for the virtual path of a multi-hop return assume that the ionospheric layers have similar heights and critical frequencies at the K reflection points. This is not true in practice especially for East-West propagation paths so that (4) and (5) are first order approximations. For most cases of interest, these approximations will yield sufficiently accurate results except in the case in which the path straddles the day-night boundary.

Angle of incidence

In order to completely specify the delay of a given ionospheric return we need still determine its angle of incidence θ_i. If we neglect the earth's magnetic field, then the angle of incidence of a return can be obtained from the following relationship between the horizontal range D_K and the electron density height profile $N(\chi,h)$ of the ionospheric layers [7]:

$$D_K = 2 \tan \theta_i \int_0^{h_r} \left[1 - \frac{a\, N(\chi,h)}{(\cos \theta_i f)^2}\right]^{-1/2} dh \qquad (6)$$

where f is the operating frequency in MHz, $N(\chi,h)$ is the electron density in electrons/cubic cm, which varies with height h and solar zenith angle χ, and $a = 80.6 \times 10^{-6}$. The height h_r is the true reflection height measured above the horizontal range D_K. Reflection occurs at that height at which the group velocity of the wave is zero. The group velocity v_g of a wave of frequency f incident on the ionosphere at an angle θ_i is given by [7]

$$v_g = c \left[1 - \frac{a\, N(\chi,h)}{(\cos \theta_i f)^2}\right]^{1/2} . \qquad (7)$$

Closed form solutions of the true reflection height, h_r, and the horizontal range D_K as a function of the angle of incidence θ_i can be found for a number of ionosperic profiles [7]. The simplest type of profile which preserves the multi-layer structure of the ionosphere and which allows a closed form solution of (6) and (7) is one in which the electron density profiles of the E, F1 and F2 layers are assumed to vary parabolically with height. The critical frequency of each layer is proportional to the square root of the peak electron density of the layer, that is

$$\begin{aligned} f_0E &= \sqrt{aN(\chi,h_mE)} \\ f_0F1 &= \sqrt{aN(\chi,h_mF1)} \\ f_0F2 &= \sqrt{aN(\chi,h_mF2)} \end{aligned} \qquad (8)$$

where f_0E, F_0F1 and f_0F2 are the critical frequencies of the E, F1 and F2 layers, respectively, and h_mE, h_mF1, h_mF2 are the heights of the maxima of the layers.

For a fixed horizontal range D_K and ionization profile $N(\chi,h)$ there is a number of angles of incidence θ_i which satisfy (6). Each solution corresponds to a return.

Doppler shift

The Doppler shift in Hz of a multi-hop return is proportional to the frequency of the transmitted signal, and the time rate of change of the phase path length, $P(\theta_i)$, that is

$$\nu_i = -\frac{f \cos \theta_i}{c} \frac{dP}{dt} \qquad (9)$$

The phase path length is the integral of the refractive index over the path so that

$$P = 2K \int_0^{h_r} \left[1 - \frac{a\, N(\chi,h)}{(\cos \theta_i f)^2}\right]^{1/2} dh \qquad (10)$$

where h_r is the true reflection height, K is the number of hops and $N(\chi,h)$ is the ionization profile which changes slowly in time. In particular the critical frequencies and the heights of the maxima of the layers exhibit diurnal variations which are directly responsible for the Doppler shifts of the skywave returns. A more specific discussion of the diurnal variations is given in Section 4.

Attenuation

The attenuation of the individual returns is due to three main effects: (1) spherical spreading; (2) non-deviative absorption in the D-region; and (3) deviative absorption in the E and F regions.

Spherical spreading losses are usually expressed in terms of an effective distance S_e. The received power has the usual inverse dependence on S_e^2. From physical

considerations it can be shown [3] that the power loss due to spreading should be proportional to the rate of change of the horizontal ground range, D_K, with respect to the angle of incidence θ_i and the number of hops K. Thus for the geometry of Figure 4,

$$S_e^2 = K^2 D_K \cot \theta_i \left|\frac{dD_K}{d\theta_i}\right| \qquad (11)$$

The spreading loss in dB is then given by

$$L_S = 20 \log \left(\frac{4\pi S_e}{\lambda}\right) \qquad (12)$$

where λ is the wavelength of the incident wave. The analytical dependence of the horizontal range, D_K, on the angle of incidence for a particular profile may be obtained from Equation (6).

Attenuation due to non-deviative absorption in the D-layer is the predominant loss mechanism. During the day time, the D-region absorption loss may be calculated (in dB) from the semi-empirical expression [1]

$$L_{ND} = K \alpha_{ND} W \operatorname{Sec} \theta_i (\cos \chi)^{.75} \qquad (13)$$

$$\alpha_{ND} = \frac{286(1+.0087Y)(1+.005R)}{\nu_D^2/4\pi^2 + (f \pm f_H)^2}$$

where K is the number of hops, Y is the latitude in degrees, R is the smoothed 13-month sun-spot number, W is the winter anomaly factor (discussed in [1]), χ is the solar zenith angle, $\nu_D/2\pi$ is the electron-neutral collision frequency in the D-region (~ 3.1 MHz) and f_H is the gyrofrequency at D layer heights (~ 1.3 MHz). The upper sign in the denominator of (13) applies to 'ordinary' rays while the lower sign applies to the 'extraordinary' rays. The justification and development of this formula can be found elsewhere [1] and will not be discussed further. Other methods for calculating the absorption loss are given by other authors [8,9] and may be found to be more accurate depending on the circumstances.

At night the non-deviative absorption is more accurately predicted by [1]

$$L_{ND} = \chi \frac{2.86\ (1+.0087Y)(1+.005R)}{\nu_D^2/4\pi^2 + (f \pm f_H)^2} \quad dB \qquad (14)$$

When the frequency of an ionospheric return is close to the critical frequency of the layer from which it is reflected, deviative absorption in that layer can contribute significantly to the attenuation of that return (especially high ray returns). If we assume that the collision frequency in the layer is constant, then the deviative absorption in dB for the ordinary rays can be calculated from [4,7]

$$L_D = 4.343\ \nu(P'-P)/c \quad dB \qquad (15)$$

where P' is the virtual path given by (4), P is the phase path given by (10), c is the speed of light and ν is the collision frequency (in sec^{-1}) in the reflecting layer. The collision frequency decreases with height so that deviative absorption is more significant for modes reflected from the E layer while modes reflected from the F2 layer are barely affected by it. Typical values of the collision frequency at different altitudes may be found in [4]. This author also gives the following formula for the deviative absorption of 'extraordinary' modes in dB:

$$L_D = 4.343\ \frac{\nu}{c}\ (P'-P)\ \frac{f}{f-f_H} \qquad (16)$$

where f is the frequency and f_H is the gyrofrequency. The factor $f/(f-f_H)$ is greater than unity so that deviative absorption is greater for 'extraordinary' modes than for 'ordinary' modes.

In addition to these three loss mechanisms, multi-hop modes are also affected by losses at each reflection from the ground. Methods of computing this loss may be found in [1].

Magnetic field effects

The main effect of the earth's magnetic field is to split a wave incident on the ionosphere into two waves called the ordinary and extraordinary waves. However in the foregoing discussion of delays (or virtual path) and Doppler shifts (or phase path) we neglected the effect of the magnetic field. The expressions given for the virtual and phase paths, ignoring magnetic fields effects, are accepted approximations for the ordinary wave [4,7]. For the extraordinary wave, however, the critical frequencies of the E, F1 and F2 layers are shifted relative to those for the ordinary wave, and it is essential to take this into account.

Denoting the critical frequencies of the ordinary and extraordinary waves by the subscripts o and x, we have that the critical frequency of the E layer for the extraordinary mode, f_xE, is given by [7]

$$f_x E = \frac{1}{2}\left\{f_H E + \left(f_H E^2 + 4 f_0 E^2\right)^{1/2}\right\} \quad (17)$$

where f_0E is the critical frequency of the E layer for the ordinary mode and f_HE is the gyrofrequency associated with earth's magnetic field at the height of the maximum of the E layer. Similar expressions apply to the critical frequencies of the F1 and F2 layers.

4. DIURNAL, SEASONAL AND GEOGRAPHICAL VARIATIONS

Having established the relationship between the macroscopic parameters of HF propagation, i.e., delay, Doppler shift, and attenuation, and the relevant ionospheric parameters, we are now ready to discuss their diurnal, seasonal, geographical and solar cycle variations. These variations result directly from diurnal, seasonal and solar cycle variations in the critical frequencies and heights of the maxima of the D, E, F1 and F2 layer.

The diurnal, seasonal and geographical variations of the D, E and F1 layers are modeled quite well [4] as a function of the instantaneous solar zenith angle χ. If we define χ_N as the local noon value of the solar zenith angle, then the diurnal variations in the solar zenith angle are well approximated by

$$\cos\chi = \cos\chi_N \cos\left[\frac{\pi(T-T_N)}{\Delta T}\right] \quad (18)$$

where ΔT is the daytime duration and T_N is the time of local noon. The parameters T_N, ΔT and $\cos\chi_N$ vary geographically and seasonally. Following Levine et al [6], the time of local noon (relative to mid-day) is well approximated by (to within 1 minute)

$$T_N = 12 + \left[\frac{X}{150}\right]_{Dec} + 0.13\sin A + .156\sin(2A) \quad (19)$$

where

A = 0.0172(D+10)
D = 30.4 (M-1) + D1
D1 = day (1,31)
M = month (1,12)
X = longitude west of Greenwich (0,360°)
and $[\ ]_{Dec}$ denotes decimal part of .

The local noon value of the solar zenith angle is approximately given by [6]

$$\chi_N = \frac{\pi}{2}\left(\frac{Y}{90^0}\right) + .409\cos A \quad (20)$$

where Y is the latitude in degrees (-90,90°). The duration of the daytime in hours may be approximated by [6]

$$\Delta T = \frac{24}{\pi}\cos^{-1}\left[\frac{\sin(.409\cos A)\sin Y - 0.26}{\cos(.409\cos A)\cos Y}\right] \quad (21)$$

where the factor -0.26 approximately represents the difference between sunrise (or sunset) at the surface of the earth and at ionospheric heights.

Diurnal, seasonal, geographical and solar cycle variations in the critical frequencies of the E and F1 layers are well approximated by the Chapman layer behavior so that [10]

$$f_0E = .9\left[(180 + 1.44R)\cos\chi\right]^{.25}$$
$$f_0F1 = (4.3 + .01R)(\cos\chi)^{.25} \quad (22)$$

where R is the 13 month smoothed sun-spot number. At nighttime when $\cos\chi = 0$, there is some residual ionization in the E-layer so that f_0E reaches a minimum of 0.5 MHz while the F1 layer disappears.

The heights of the maxima of the E and F1 layers also follow the Chapman layer behavior, i.e.,

$$h_mE = 105. - 14.\ln(\cos\chi + .117)$$
$$h_mF1 = 165. - 18.\ln(\cos\chi + .047) \quad (23)$$

where the factors .117 and .047 ensure that the heights of the E and F1 layers do not exceed their maximum nightime heights of 135 and 220 km respectively. Their minimum noontime heights have been given as constants, 105 and 165 km respectively. However there is evidence that they vary with latitude, with lower heights observed at equatorial latitudes [4].

Equations (18) through (23) describe the diurnal, seasonal, geographical and solar cycle behavior of the E and F1 layers. The corresponding behavior of the F2 layer cannot alone be described in terms of the simple Chapman layer because the physics of this layer is believed to involve an interaction of photo-chemical processes

which are too complex to be modelled as cos χ variations. In fact, the F2 layer exhibits diurnal, seasonal, and equatorial anomalies. The theory of the processes which control the diurnal and seasonal variations of electron density in the F2 layer has been summarized by Risbeth [5]. This theory has been used by Levine, et al [6] to model diurnal variations in f_0F2 as the response of a dynamic system "driven" by a function of cos χ. Their model has proved to be just as accurate as other models [1] which employ world maps of f_0F2 in predicting the MUF for point-to-point Hf links. They express the variations in f_0F2 as

$$f_0F2 = (8 + .032R)(\cos \chi_e + .00887)^{.25} \quad (24)$$

where $\cos \chi_e$ is a function of the instantaneous solar zenith angle cos χ which takes into account relaxation processes.

At nightime, there is no solar induced ionization so that f_0E decays exponentially and [6]

$$(\cos \chi_e)_{NIGHT} = (\cos \chi_e)_{SUNSET} \exp[-(T-T_S)/t_N] \quad (25)$$

where $T-T_S$ is the time elapsed since sunset, and t_n is a nightime relaxation time which is taken to be a constant (2 hours) independent of season and geographical location.

During the day $\cos \chi_e$ is given by the sum of a term proportional to the solar induced ionization, i.e., proportional to cos χ, and a term which accounts for recombination-like electron-loss chemical processes. Thus during the day [6]

$$(\cos \chi_e)_{DAY} = \frac{\cos \chi_N}{1 + \beta^2} \left\{\cos\left[\frac{\pi(T-T_N)}{\Delta T}\right] + \beta e^{-(T-T_{DW})/t_D} + \beta \sin\left[\frac{\pi(T-T_N)}{\Delta T}\right]\right\} \quad (26)$$

where $\beta = \pi t_D/\Delta T$, while T_N and χ_N are the local noon time and solar zenith angle, ΔT is the daytime duration, T_{DW} is the local dawn time ($T_{DW} = T_N - \Delta T/2$) and t_D is a relaxation time (in hours) which depends on geographical location and season. Specifically, Levine et al [6] assume the relaxation time to be a function of the actual noon-time solar zenith angle χ_N which depends on geographical location and season:

$$t_D = \text{Max}\left\{9.7 (\cos \chi_N)^{9.6}, 0.1\right\} \quad (27)$$

Thus during the summer at equatorial and moderate latitudes $t_D = 9.7$ hours whereas in the winter $t_D \ll 9.7$ so that (27) coupled with (26) reproduce the diurnal and seasonal anomalies of the F2 region.

Equations (18) through (26) have been used to plot the diurnal and seasonal variations of f_0E, f_0F1 and f_0F2 in Figures 5 and 6 for a high sun-spot year (R = 150). These variations are seen to follow accepted behavior [4,10].

The same behavior of f_0F2 can be extended to predict the height of the maximum of the F2 layer according to

$$h_mF2 = h_0F2 - 25 \ln (\cos \chi_e + .00033) \quad (28)$$

where the minimum daytime height of the maximum of the F2 layer, h_0F2, is highly dependent on latitude. It approximately varies according to [4]

$$h_0F2 = 230 - 50 \ln (\cos y) - .3y \cos A \quad (29)$$

where y is the latitude in radians and A is a seasonal parameter defined after equation (19).

These expressions have been used to plot in Figure 7 the diurnal behavior of the heights of the maxima of the E, F1 and F2 layers during the month of June. Unlike the E and F1 layers which basically disappear at night, the F2 layer remains and its peak continues to rise reaching a maximum height at 5 hours (i.e., sunrise). After sunrise, the height of the peak of the layer decreases reaching a minimum sometime in the afternoon depending on the time of year. The seasonal variations in the heights of the layers are not as pronounced.

5. DISCUSSION

The dependence of the virtual and phase path of skywave returns on the critical frequencies and heights of the maxima of the E, F1 and F2 layers can be used to determine the diurnal, seasonal, geographical and solar cycle variations in the number of modes, their delays, Doppler shifts and relative strength. As an example we have plotted in Figures 8 and 9 the diurnal variations (month of December and high sun-spot number -150) in the maximum

delay spread and Doppler shift spread between the various modes (up to 40 dB weaker than the strongest mode) received over a 1300 km path at a latitude of 40° north and longitude 100° west. The delay spread and Doppler shift spread are plotted as a function of the operating frequency so that the well known diurnal variations in the MUF, where the delay spread and Doppler spread reduce to zero, can also be observed. These curves show that at a fixed time of day the delay and Doppler spreads increase as the operating frequency decreases below the MUF, reaching a maximum at some frequency. This frequency coincides with the maximum frequency at which 2-hop and 3-hop F2 mode propagation is possible. Below this frequency both the multipath and Doppler spreads decrease. The frequency at which the multipath and Doppler spreads reach a maximum is proportional to the MUF; in fact it is approximately .55 MUF. The multipath spread is in general greatest in the early afternoon hours when ionization in the F2 region reaches a maximum. On the other hand the Doppler spread is greater at sunrise and sunset when changes in the ionospheric layers occur at a faster rate, and lowest at noon and at night when the layers are in near equilibrium. The seasonal, geographical and solar cycle variations of these parameters at a fixed time of day are less pronounced. Nonetheless, use can be made of the inclusion of these effects in the model to generate long term distributions as a function of frequency and distance.

Summarizing, we have described in this paper an HF propagation model suitable for calculating the various parameters needed to plan and predict the performance of wideband HF communications systems. The model takes into account the diurnal, seasonal, geographical and solar cycle variability of HF propagation in a simple but fairly accurate manner. Some additional refinements are required, however, to reflect more accurately the characteristics of the ionosphere near the equator and the auroral oval. Hence the model is more accurate at mid-latitudes.

References

[1] A. Barghausen, J.W. Finney, L.L. Proctor, L.D. Schultz, "Predicting Long Term Operational Parameters of HF Skywave Telecommunication Systems", ESSA Technical Report ERL 110-ITS-78, May 1969.

[2] W.B. Jones, D.L. Obitts, "Global Representation of Annual Solar Cycle Variation of f_0F2 Monthly Median 1954-1958", OT/ITS Research Report No. 3, October 1970.

[3] L.W. Pickering, "The Calculation of Ionospheric Doppler Spread on HF Communication Channels", IEEE Trans. Communications, V. COM-23, No. 5, pp 526-537, May 1975.

[4] Y.L. Al'pert, Radio Wave Propagation and the Ionosphere, New York: Consultants Bureau. 1963.

[5] H. Risbeth, "A Review of Ionospheric F-Region Theory", Proc. IEEE, V55, No. 1, pp 16-34, January 1967.

[6] P.H. Levine, R.B. Rose, J.N. Martin, "A Simplified HF MUF Prediction Algoroithm", IEE Conference on Antennas and Propagation, Publication No. 169, Pt. 2, pp 161-167, November 1978.

[7] K.G. Budden, Radio Waves in the Ionosphere, Cambridge: University Press, 1961.

[8] K. Rawer. "Calculation of Sky-wave Field Strength", Wireless Engineer, pp 287-300, November 1952.

[9] P.L. George, et al, "A New Method of Predicting Ionospheric Absorption", Telecommunication Journal, V. 41, pp 308-312, 1974.

[10] K. Davies, Ionospheric Radio Waves, Waltham, MA: Blaisdell, 1969.

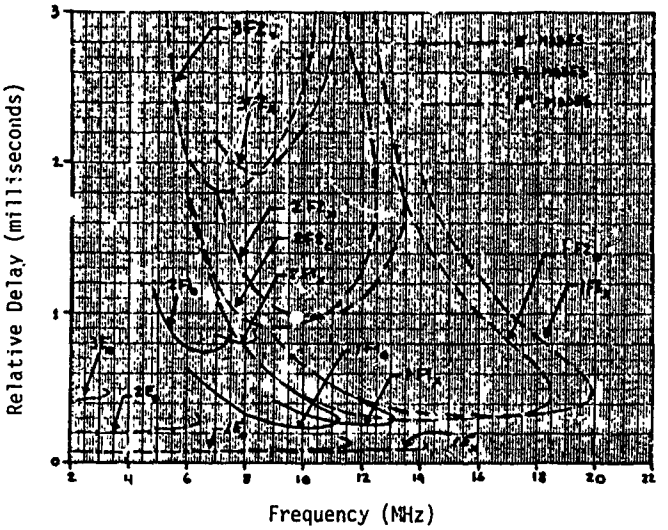


Figure 1: Mode delays relative to the ground wave delay as a function of frequency. Path length = 1300 km; time of day = 16 hours; month = December, SSN = 150; Latitude = 40°N; Longitude = 100° W.

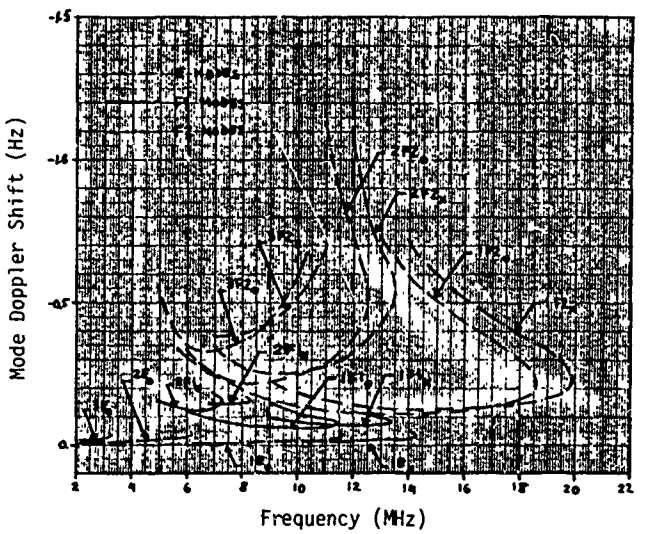


Figure 2: Mode Doppler shifts as a function of frequency. Path length = 1300 km; time of day = 16 hours; month = December; SSN = 150; Latitude = 40°N; Longitude = 100°W.

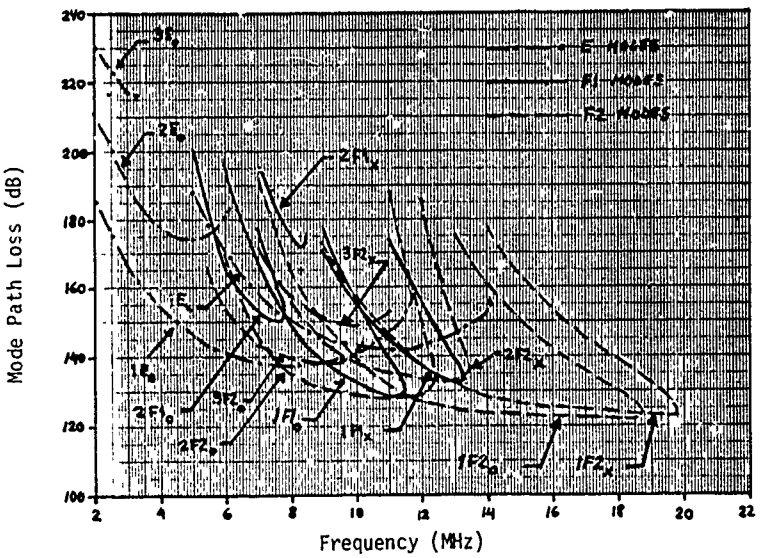


Figure 3: Mode attenuation as a function of frequency. Path length = 1300 km; time of day = 16 hours; month = December; SSN = 150; Latitude = 40°N; Longitude = 100°W.

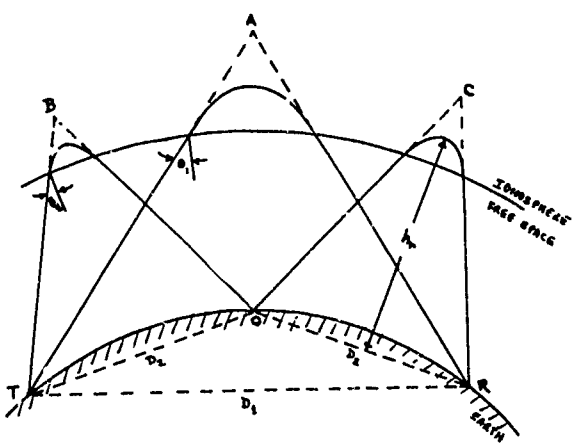


Figure 4: One-hop and two-hop path geometry.

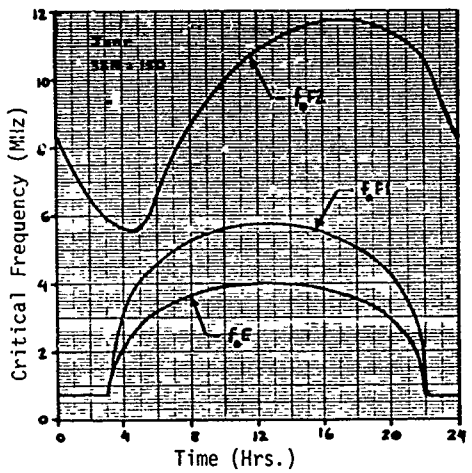


Figure 5: Diurnal variations of the critical frequencies of the E, F1 and F2 layers.

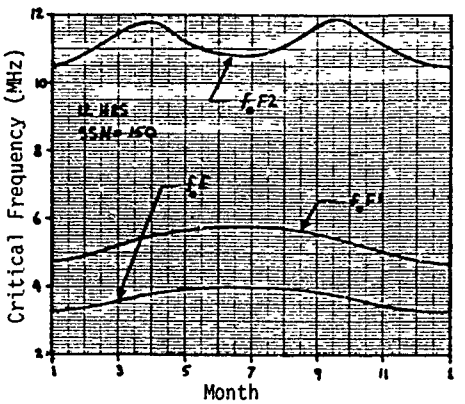


Figure 6: Seasonal variations of the critical frequencies of the E, F1 and F2 layers.

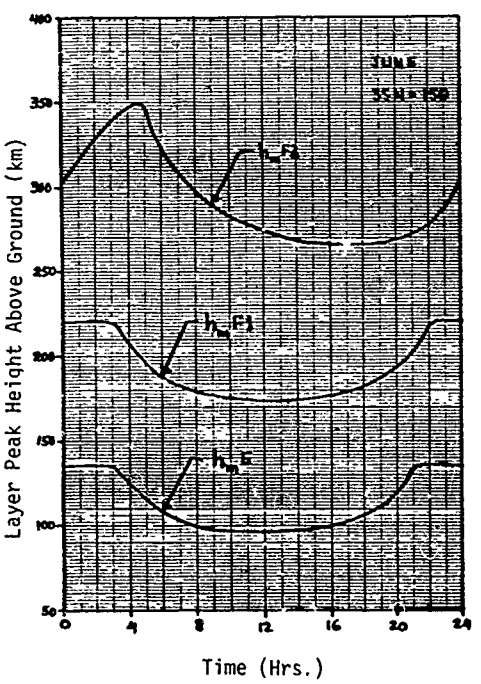


Figure 7: Diurnal variations of the heights of the maxima of the E, F1 and F2 layers.

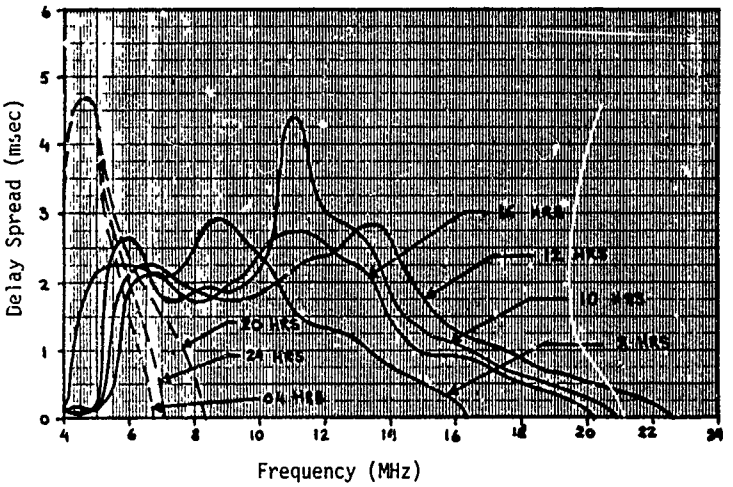


Figure 8: Diurnal variations in delay spread as a function of frequency. Path length = 1300 km; month = December; SSN = 150; latitude = 40°N; longitude = 100°W.

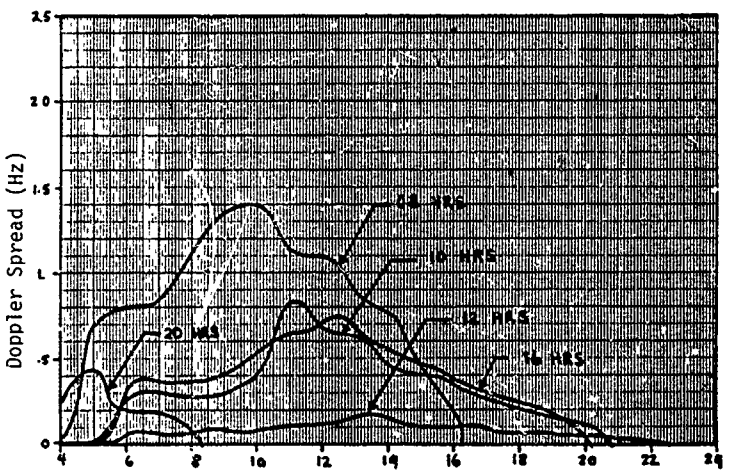


Figure 9: Diurnal variations in the Doppler shift spread as a function of frequency. Path length = 1300 km; month = December; SSN = 150; latitude = 40°N; longitude = 100°W.

THE USE OF THE INTERNATIONAL REFERENCE IONOSPHERE (IRI-78) MODELS FOR CALCULATING LONG PATH FIELDS AT VLF

F. P. Snyder
J. A. Ferguson
Naval Ocean Systems Center
San Diego, CA 92152

INTRODUCTION

This report presents comparisons of measured very low frequency (VLF) long path propagation data with corresponding theoretically computed values. The theoretical values were computed using full-wave VLF waveguide mode theory and D-region electron density profiles presented by Rawer et al (1978), and designated by COSPAR and URSI as the International Reference Ionosphere - 78 (IRI-78). The "International Reference Ionosphere" project, a joint URSI and COSPAR effort since 1969, is intended to summarize the most important features of the terrestrial ionosphere. The IRI is intended to combine approved experimental results, consisting of in situ rocket observations combined with radio propagation experiments between the ground and the rocket, so as to provide a useful ionospheric reference with no dependence on theoretical assumptions. This is in contrast to theoretical models derived from ground-to-ground propagation experiments. The IRI profiles are obtained from computer models composed of a set of independent functions, each valid in a particular height range, that attempt to represent averages of measured ionospheric parameters.

A number of ionospheric D-region models have been constructed on the basis of radio propagation data alone. They are subject to a number of objections. For example, it is impossible to show that the electron density distributions obtained by using inversion techniques are uniquely determined by the propagation data employed. Even if profiles cannot be properly established by inversion of such measurements, propagation experiments do offer useful data for checking experimentally obtained profiles. For example, Singer, et al (1980) computed radiowave absorption in the MF and HF ranges using an early version of the IRI, and compared these data with ionospheric absorption measurements made in Europe with the pulse and continuous wave methods. They found that the D-region portion of the earlier IRI model should be revised, particularly with respect to the solar cycle variation. Similar comparisons with VLF long path propagation data have not been reported.

EXPERIMENTAL DATA

Recently, a set of exponentially varying D-region electron density profiles were recommended for VLF long path propagation studies by Morfitt (1977). These profiles were obtained by fitting multifrequency long path VLF propagation data to theoretically computed data on a trial and error basis. Some of the VLF propagation data used in the Morfitt (1977) analysis are used in this report. The characteristics of the propagation paths discussed in this report are presented in Table 1.

Table 1. Propagation Path Characteristics

Propagation Path	Frequency (kHz)	Sunspot Number (R)	Season
Day Conditions			
NPM (Hawaii) - San Francisco	19.9,24.0	20	summer
San Diego	16.6	20	summer
Wake Island	16.6,19.8	20	summer
Sounder (Hawaii) - San Diego	9.34,10.897 15.567,21.794 28.021	20	winter
Night Conditions			
NPM (Hawaii) - Seattle	23.4	120	winter
Sounder (Hawaii) - San Diego	9.34,10.897 15.567,21.794 28.021	20	winter
Sounder (Hawaii) - Ontario	9.34,10.897 14.010,15.567 17.124,21.794 24.908,26.454 28.021,31.134	120	winter

From Table 1 we see that the bulk of the daytime propagation data is for the summer season and for low sunspot numbers, while for night conditions the bulk of the data is for the winter season and high sunspot number. With exception of the daytime path NPM(Hawaii) to Wake Island, which is an east-to-west propagation path, all of the propagation paths considered are west-to-east. The NPM station is a communications broadcast station and the Sounder is a multifrequency experimental transmitter located on the island of Hawaii. All of the signal amplitudes were recorded aboard an airplane flying between the transmitter and the path endpoint.

THEORETICAL CALCULATIONS

Long path VLF propagation is conveniently treated in terms of waveguide theory wherein the VLF field is written as a summation of terms called modes. For a time harmonic source, the VLF modesum for the vertical electric field may be written as

$$E_v(d) = \frac{K(P,f)}{\sqrt{\sin d/a}} \sum_n A_n G_n e^{-ik_o S_n d} \qquad (1)$$

where $K(P,f)$ is a complex constant dependent on transmitter power (P) and frequency (f), d is the distance from the transmitter on a homogeneous, smooth earth of radius a, A_n is the excitation factor for mode n, normalized to unity for flat earth with perfectly conducting boundaries, k_o is the free space wave number, S_n is the propagation factor, and G_n represents a height-gain function for mode n, normalized to unity at the ground. A similar expression can be written for the other electric and magnetic components of the VLF field. The magnitude of the vertical electric field mode sum of equation (1) is considered for this report.

The height-gain function (G_n) of equation (1) represents the variation of the modal function as a function of height. In addition to the magnitude of the vertical electric field component, some properties of the height-gain functions are also considered. In particular, by considering the ionosphere as a lossy dielectric, the Joule heating losses can be computed as a function of height, both for individual modes as well as the total field. For selected multifrequency sounder data, a comparison is made between the total Joule heating losses for the IRI-78 profiles and the exponentially varying profiles recommended by Morfitt (1977). Specifically, we will compare among the various profiles the altitude at which the heating loss is a maximum as well as the altitudes were the heating loss drops to 1/e and to 1/10 of the maximum value.

COMPARISONS BETWEEN IRI-78 AND MEASURED DATA

Comparisons of the measured VLF field strengths with the field strengths computed using the IRI-78 profile are presented in this section, first for the daytime and then for the nighttime. For both cases, the same collision frequency profile was used. The collision frequency profile, given by $\nu = 1.816 \times 10^{11} \exp(-.15h)$, where h is the height above the surface, is the same collision frequency profile used by Morfitt (1977). The IRI-78 electron density model is given in the Appendix. The field strengths were computed using the technique reported by Ferguson and Snyder (1980).

Daytime

Figure 1 shows the measured and computed VLF field strengths for the NPM to San Francisco propagation path at 19.8 and 24.0 kHz. For this low sunspot number, summer case, there is excellent agreement between the computed and experimentally measured results. The signal amplitudes differ by only a few dB, except at the null locations. Furthermore, the null locations are in agreement to within 100 to 200 km. The NPM to San Diego 16.6 kHz comparison is shown in Figure 2. This propagation path is nearly identical to the NPM to San Francisco path. Here again the agreement between the measured and computed field strengths is excellent.

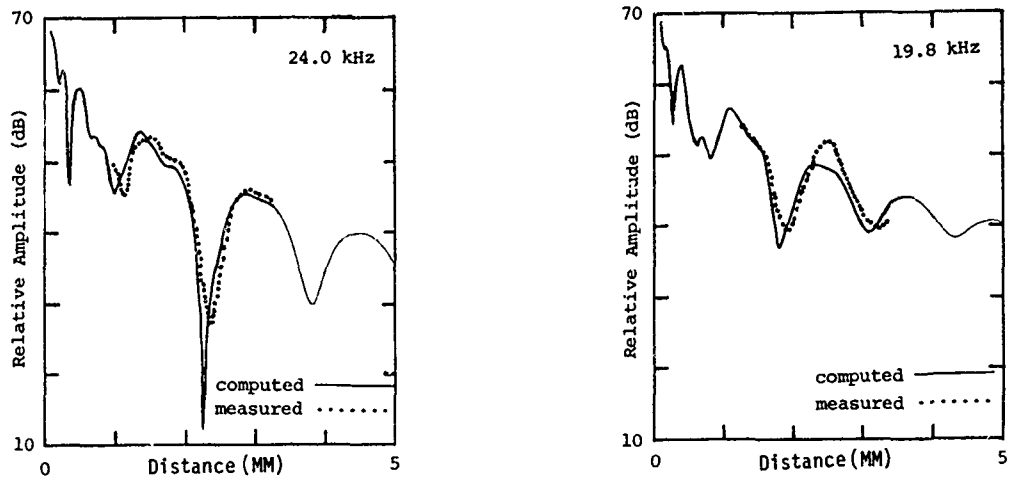


Figure 1 Computed and Measured Signal Strengths
NPM to San Francisco - Day

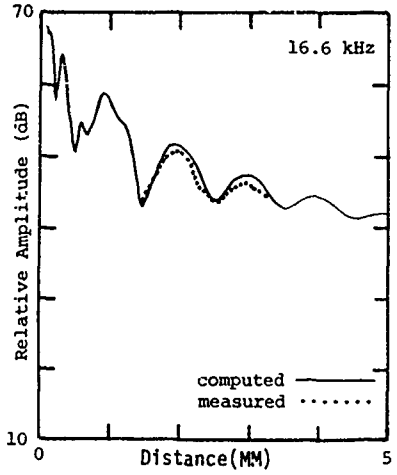


Figure 2 Computed and Measured Signal Strengths
NPM to San Diego - Day

The electron density profiles for the NPM to San Francisco path are shown in Figure 3. Both the IRI-78 and the exponentially varying profile are included. The exponential profile is given by Morfitt (1977) as one with reference height (h') of 70 km and gradient (β) of 0.5 km^{-1}. The β,h' notation is adopted from Wait and Spies (1964) who assumed that $\omega_p^2/\nu = 2.5 \times 10^5 \exp(\beta(z-h'))$, where z is the height above ground in kilometers, ω_p is the angular plasma frequency, and ν is the electron collision frequency given previously. The ratio, ω_p^2/ν, actually represents an ionospheric conductivity parameter. A plot of this parameter for both profiles is shown in Figure 3b. The conductivity parameters differ considerably throughout the height range presented. However, they are nearly the same in the vicinity of the reference height for the exponential profile. It is this similarity in the height range near the reference height that results in nearly identical mode sums for the two profiles.

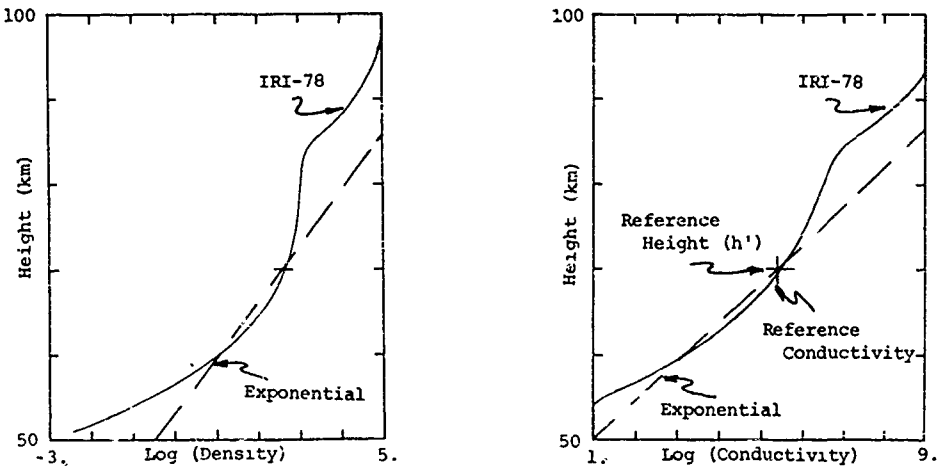


Figure 3 Electron Density and Conductivity Parameters
NPM to San Francisco - Day

Figure 4 shows the results for the NPM to Wake propagation path, the only path from east-to-west. The agreement between the computed and measured field strengths for two lowest frequencies, 16.6 and 19.8 kHz is very good, while the computed amplitude for 26.1 kHz is approximately 5 dB lower than the measured values. However, the location of the signal minima are in good agreement in all three cases.

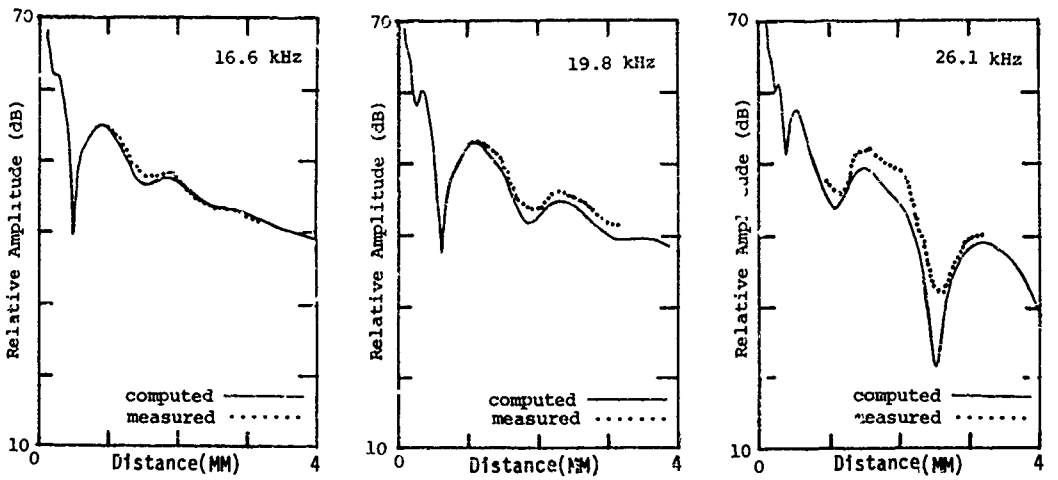


Figure 4 Computed and Measured Signal Strengths
NPM to Wake Island - Day

The computed and measured amplitudes for the multifrequency sounder to San Diego propagation path are shown in Figure 5. For this case, VLF signals were measured on two flights of the airplane on consecutive days, one flying from the San Diego area towards the sounder location and the other flying in the opposite direction. The season for this case is winter rather than summer, although the sunspot number is the same for this case as for the data in Figure 1. Note that there are some differences between the two sets of measured data and that the computed results are in better agreement with data from one of the days than with the other. The difference between the two sets of measured amplitudes is an indication of the daily variability of the VLF signal strengths as a function of distance as well as an indication of how closely one can expect to match experimental and theoretical VLF field strengths using as simple an ionospheric profile as the IRI model. The best agreement between computed and measured amplitudes is for the lower frequencies and there is relatively poor agreement for the higher frequencies. It is interesting to observe that the parameters of the exponentially varying profiles recommended by Morfitt (1977) are frequency dependent. This difficulty of achieving a good fit over an entire frequency range with a single profile is similar to the results reported by Singer et al (1980).

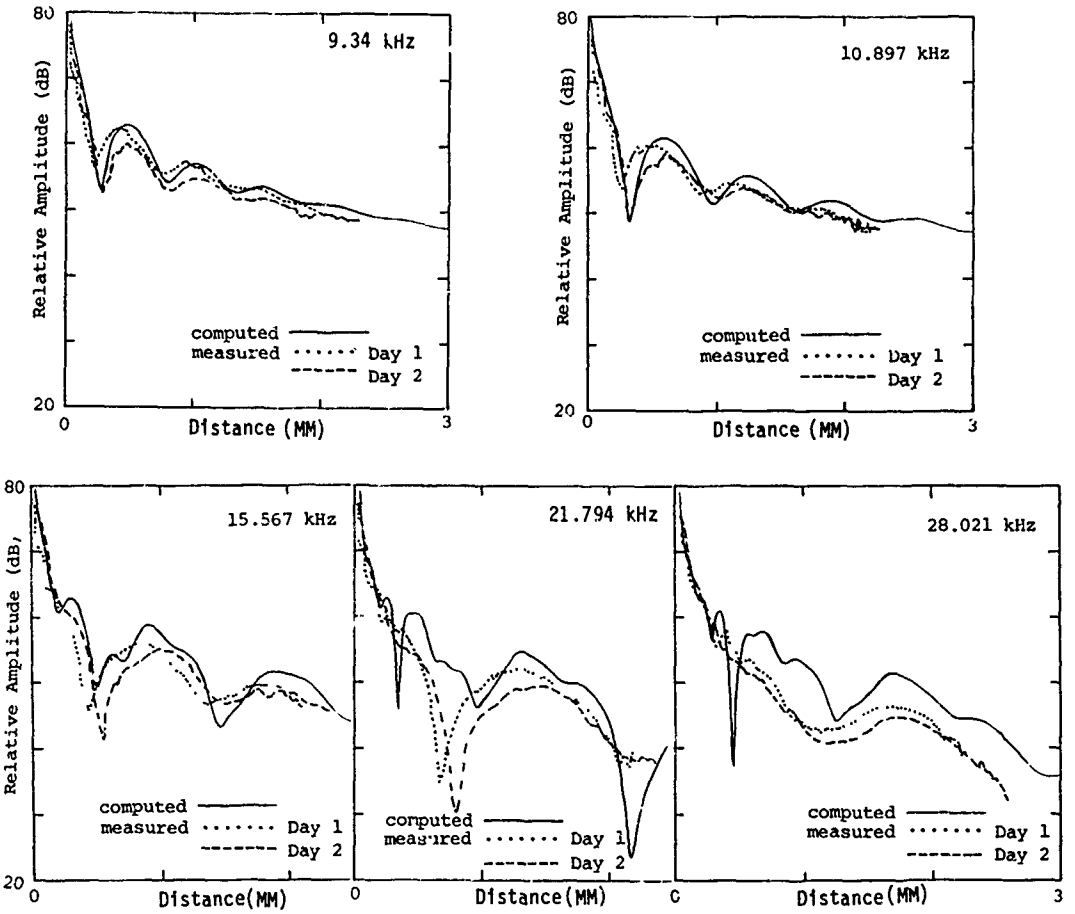


Figure 5 Computed and Measured Signal Strength
Sounder to San Diego - Day

A comparison of the computed Joule heating functions for the three profiles used in the multifrequency sounder comparisons is presented in Figure 6. Shown in the figure is the height where the heating maximizes for the three profiles as a function of frequency. Also shown are the heights where the heating drops to 1/e and 1/10 of the maximum values. The height of the maximum heating for the IRI-78 profile is intermediate to the corresponding heights for the two exponential profiles. The frequency dependence for the height of the maximum Joule heating for

the two exponential profiles results from the frequency dependence in the profile parameters. Note that such a frequency dependence does not occur for the IRI-78 profile. The frequency dependence for the upper and lower limits of the heating is a result of the variation in field amplitude due to the mode structure rather than the variation in profile height due to the variation in profile parameters.

Figure 6 indicates the influential regions of the ionospheres used in the calculations. In all three cases, a region of about eight to ten km above the maximum of the heating is significant, whereas there is a difference between the apparently significant lower regions of the ionosphere for the two exponential profiles as compared to the IRI-78 profile. Refering to Figure 4, we see that in general the computed amplitude is larger than the measured amplitude. Although not shown on the figures, the computed amplitudes using the exponential profiles are more in agreement with the measured amplitudes, indicating a higher attenuation, which is consistent with the larger heating loss indicated in Figure 6.

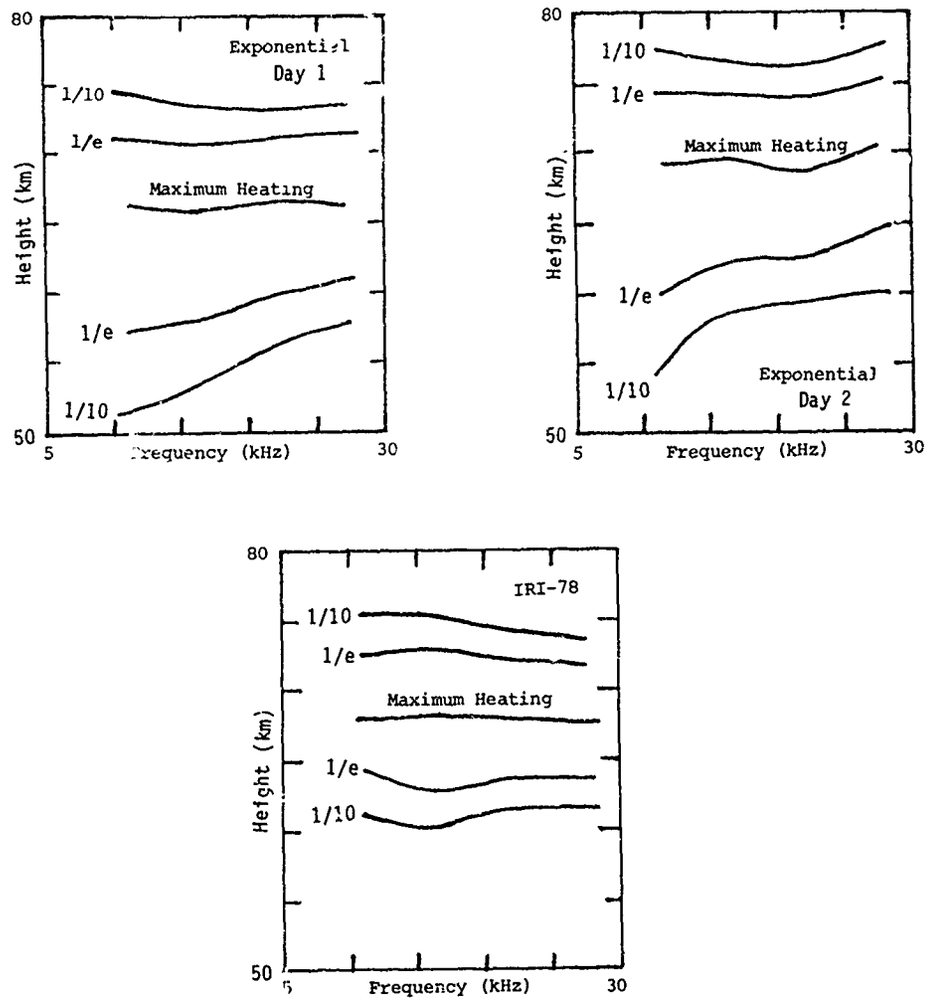


Figure 6 Significant Altitudes for the Joule Heating Function
Sounder to San Diego - Day

Nighttime

The IRI-78 model for nighttime was examined for three propagation paths as shown in Table 1. One of the propagation paths was from the single frequency transmitter NPM at 23.4 kHz to Seattle. Figure 7 shows the measured and computed field strengths for this path. For this high sunspot number, winter condition, the computed field strengths are much larger, by 6 dB or more, than the measured signal strengths. Although difficult to establish with just this single set of data, it does appear that the modal interference null locations are predicted using the IRI-78 model at approximately the proper distances from the transmitter. Although not shown, the ionospheric conductivity parameters for the IRI-78 profile and the exponential profile recommended by Morfitt (1977) were examined and found to be nearly the same in the near vicinity of the reference height for the exponential profile. This similarity might well account for the proper location of the modal interference nulls in the computed field strengths.

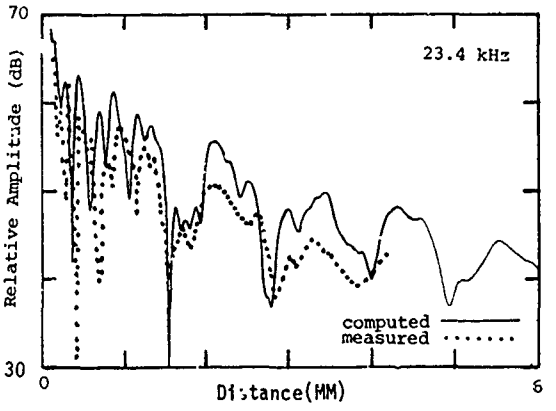


Figure 7 Computed and Measured Signal Strength
NPM to Seattle - Night

Two sets of nighttime data for the multifrequency sounder are indicated in Table 1. One set, on the propagation path to San Diego, is for a low sunspot number, winter condition, while the other set, on the propagation path to Ontario, is for a high sunspot number, winter condition. The data for the Ontario path were obtained within a month of the data for the NPM to Seattle path. The data for the San Diego path, however, were obtained at least five years later than the Ontario or Seattle path data.

Comparisons of the measured and computed field strengths for three frequencies for the San Diego path are shown in Figure 8 and for the Ontario path in Figure 9. A striking feature of both Figures 8 and 9 is the almost total lack of similarity between the computed and measured field strengths. In general, there is no similarity in null locations or the amplitudes. The measured data are typically 10 to 20 dB lower than the computed amplitudes. The measured amplitude data for the two propagation paths are quite dissimilar, probably due to the large difference in sunspot number. The computed amplitudes, however, are very nearly identical, due to the nearly identical profiles for the IRI-78 model. In fact, the only difference in the computed profiles for the IRI-78 is in the electron densities at altitudes in excess of 100 km.

An examination of the Joule heating for the IRI-78 profile, shown in Figure 10, clearly indicates the most important altitude region for this profile is between 80 and 90 km. Thus, there is very little sensitivity to profile variations above 90 km. The corresponding Joule heating for the recommended exponential profile is also shown in Figure 10. The heating for the exponential profile is clearly distributed over a much larger altitude range than for the IRI-78 profile.

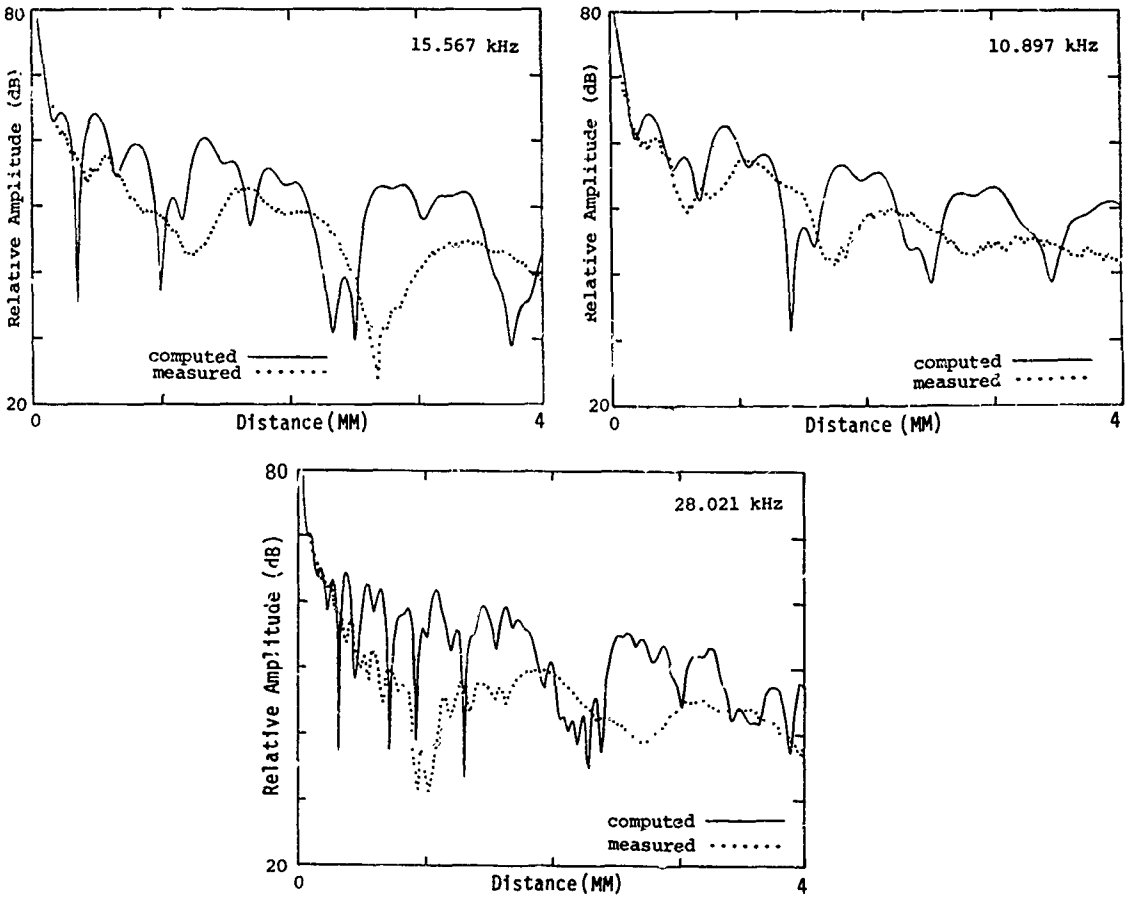


Figure 8 Computed and Measured Signal Strengths
Sounder to San Diego - Night

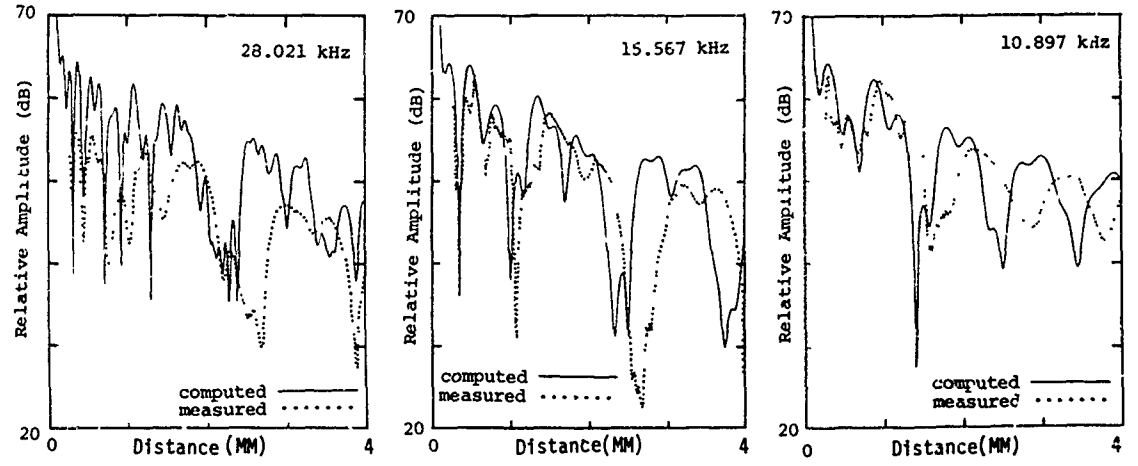


Figure 9 Computed and Measured Signal Strengths
Sounder to Ontario - Night

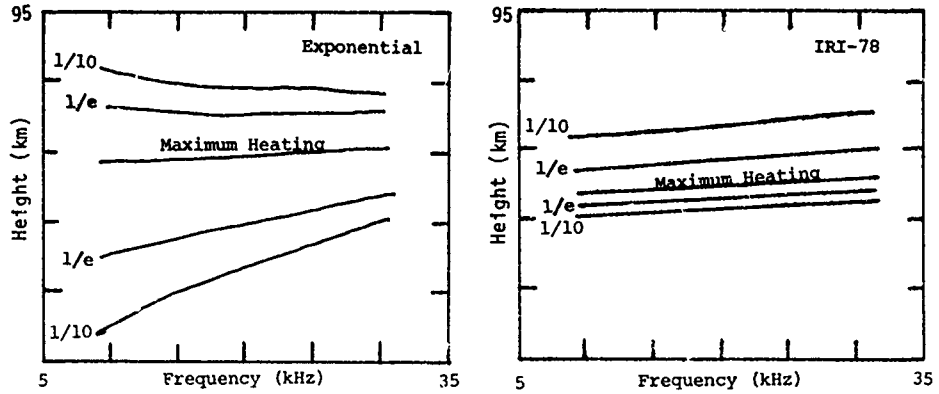


Figure 10 Significant Altitudes for the Joule Heating Function Sounder to Ontario - Night

DISCUSSION

In view of the results presented in the foregoing sections, no attempt was made to determine an improved set of parameters for the daytime IRI-78 profiles. Some limited data have been examined (although not presented) which indicate, that the IRI-78 D-region model may be inadequate for long path VLF field strength calculations under conditions of high sunspot number. It should also be recalled that the simple model of the IRI-78 profiles is inadequate to provide a good fit to daytime VLF propagation data over a large frequency band as has been indicated in the multifrequency sounder to San Diego analysis.

The IRI-78 profile model has been shown to be inadequate for computing long path VLF signals for nighttime conditions, at least for the winter season. Measured data for nighttime conditions for two different sunspot numbers indicate a sunspot dependence in the ionospheric profile. The IRI-78 model does not reflect a similar dependence. Numerous modifications to the parameters of the IRI-78 nighttime model were examined in an attempt to find a reasonable set of parameters for VLF long path nighttime propagation modeling. The procedure adopted was to start with the given parameters for the IRI-78 and modify an individual parameter in small steps attempting to approach a better agreement between the computed and measured field strengths. The results of such modification in every case lead to a set of parameters which produced electron densities similar to the exponential electron densities of Morfitt (1977). Therefore, the IRI-78 nighttime D-region electron densities are found to be generally too high, a result just the opposite of that reported by Singer, et al (1980).

REFERENCES

Ferguson, J. A. and F. P. Snyder (1980), Approximate VLF/LF waveguide mode conversion model, Naval Ocean Systems Center Technical Document 400.

Morfitt, D. G. (1977), Effective electron density distributions describing VLF/LF propagation data, Naval Ocean Systems Center Technical Report 141.

Rawer, K., D. Bilitza and S. Ramakrishnan (1978), Goals and status of the international Reference Ionosphere, Rev. Geophys. and Space Phys., v. 16, no. 2, pp. 177-181.

Singer, W., J. Taubenheim and J. Bremmer (1980), A test of the IRI lower ionosphere models by comparison with radio propagation data, J. Atmos. Terr. Phys., v. 42, pp. 241-248.

Wait, J. R. and K. P. Spies (1964), ristics of the earth-ionosphere waveguide for VLF radiowaves, NBS Technical No. 00.

APPENDIX

The IRI-78 electron density model for the D-region is as follows:

Input parameters: HME, NME, HMD, NMD, F1, F2, F3

Compute: HDX = HMD + F2

$$N(h) = \begin{cases} NME \times \exp(-D1 \times (HME-h)^{k}) & HME > h > HDX \\ NME \times \exp(F1X - F1^{2}X^{2}/2 + PX^{3}) & HDX > h \end{cases}$$

$$P = \begin{cases} (F2 \times F1^{2}/2 - F1 + 1/F2)/(F2 \times F2) & h > HMD \\ (F3 \times F1^{2}/2 - F1 - 1/F3)/(F3 \times F3) & h < HMD \end{cases}$$

$$X = h - HMD$$

Continuity of N(h) and its derivative dN/dh at h = HDX(NDX,DNDX) gives

$$k = -DNDX \times (HME-HDX)/(NDX \times Ln(NDX/NME))$$

$$D1 = DNDX/[(NDX \times k \times (HME-HDX)^{k-1})]$$

The input parameters F1, F2, and F3 have the meaning:

F1 = [d(lnN)/dh] h = HMD

F2(F3) is the difference between HMD and the height where the electron density increases (decreases) to NMD×e (NMD/e).

Parameter values for propagation paths of Table 1.

	HME (km)	NME (m^{-3})	HMD (km)	NMD (m^{-3})	F1 (km^{-1})	F2 (km)	F3 (km)
Day Path							
NPM - San Francisco	110	1.44×10^{11}	81	1.12×10^{9}	0.04	4.6	-11.5
- San Diego	↓	1.46	↓	1.14	0.03	↓	↓
- Wake		1.50		1.15	0.03		
Sounder - San Diego		1.20		4.33	0.03		
Night Path							
NPM - Seattle	105	3.52×10^{9}	88	4.0×10^{8}	0.05	4.5	-4.0
Sounder - San Diego	↓	1.93	↓	↓	↓	↓	↓
- Ontario		3.52					

A SEMI-EMPIRICAL MODEL OF IONOSPHERIC ELECTRON DENSITY

K. A. Pfitzer
S. J. Scotti
W. P. Olson

McDonnell Douglas Astronautics Company
Huntington Beach, California 92647

INTRODUCTION

A model of ionospheric electron density has been developed. It was intended that it be global, computationally efficient, and that it could be used to quantitatively describe ionospheric features that influence radio communications. It was designed to be useful as a predictive tool by having input parameters that can be measured in real time. We report here on its construction and recent updates which attempt to tie the behavior of the F_2 layer to measurable parameters. Model features include: the structure of the E, F_1, and F_2 layers, the equatorial anomaly, the winter anomaly, polar cap structure, semiannual and seasonal dependences, and the mid-latitude trough. The input parameters to the model are chosen such that satellite data can be used with the model to specify N_e in real time.

The predictive capability of the model stems from its dependence on inputs from other quantitative environmental models. For example, the location of the mid-latitude trough is largely controlled by the locations (L values) of the plasmapause and the auroral oval. In fact, several gross ionospheric features are controlled by the geomagnetic field. The MDAC magnetic field model provides the necessary magnetic inputs. The model also relies heavily on the MDAC model of upper atmospheric neutral density, N. Several ionospheric structures and temporal dependences tie directly to variations in atmospheric density. For example, the semiannual variation in N_e is produced almost entirely by the semiannual variation in N as described by the MDAC neutral density model. The structures of the E and F_1 layers are also determined by calculating the absorption of solar electromagnetic radiation in the neutral density model. A short summary of the features of the MDAC atmospheric density model used in this study is given below.

2.1 Atmospheric Density Model

The MDAC atmospheric density model includes as inputs both the solar ultraviolet and charged particle energy sources. The model is semi-empirical--it is based on available satellite data and our present understanding of both the UV and corpuscular energy sources. It offers a global description of the atmosphere's density, taking into account the dependence of atmospheric density on the solar flux constant. The UV and corpuscular effects are computed in separate coordinate systems and their contributions added.

Ever since an accelerometer was first flown on a polar satellite (Bruce, 1968), it has been apparent that there are at least two density bulges in the lower thermosphere during geomagnetically quiet times. The low-latitude bulge is usually attributed to heating by solar UV radiation (although other energy sources may contribute importantly), but the high-latitude bulge is produced by particles precipitating into the upper atmosphere through the dayside cusps.

The solar UV heating source is best described in geographic coordinates. Thus, inputs to the atmospheric computer subroutine must include the universal time and the time of year. Local time is also entered as a function of the coordinates of the point where the atmospheric density is to be determined. The dayside cusp particles are constrained to precipitate into the atmosphere along magnetic field lines. The cusp intersection with the atmosphere is nominally 15 degrees below the magnetic dipole axis with its longitude center on the magnetic noon meridian plane as defined by the dipole axis and the sun-earth line. The extent of this intersection in longitude is about 12 hours.

The extent of the cusp in latitude is several degrees (3-5). The region of the atmosphere actually heated by these particles is, of course, much larger because the impact energy is spread out by winds, thermal conduction, and possibly gravity waves. The corpuscular contribution is found in geomagnetic coordinates and combined with the UV contribution.

This model of the density of the neutral upper atmosphere is functionally very simple and computationally very fast. Yet, it describes most of the known variations in the atmosphere; latitude, longitude, diurnal, seasonal, semi-annual, altitude, solar cycle, and variations in particle precipitation. Thus it can be used to predict the density of the atmosphere as a function of time, position, solar and magnetospheric parameters. Since it includes the heating effect of particles precipitating into the high latitude regions, it is especially valuable in the study and prediction of the density of the polar atmosphere. Furthermore, since this high latitude effect is parameterized, it can be used to predict increases in the high latitude density when the solar wind particle flux increases and the cusp location is observed to change. The model is analytic and thus differentiable.

The three-dimensional mercator projection in Figure 1 shows the atmospheric density at 400 km during summer solstice at a universal time of 1600 hours. (The density peak produced by solar UV radiation is located to the north of the equator and at about 1400 hours local time.) The atmospheric heating peak caused by charged particle precipitation is most intense on the subsolar magnetic longitude (containing the magnetic dipole). Since the total density is dependent on the particle heating effect as well as the UV heating, the northern particle heating peak which is in sunlight is much more pronounced than the southern particle heating peak which is far into darkness during the summer solstice night.

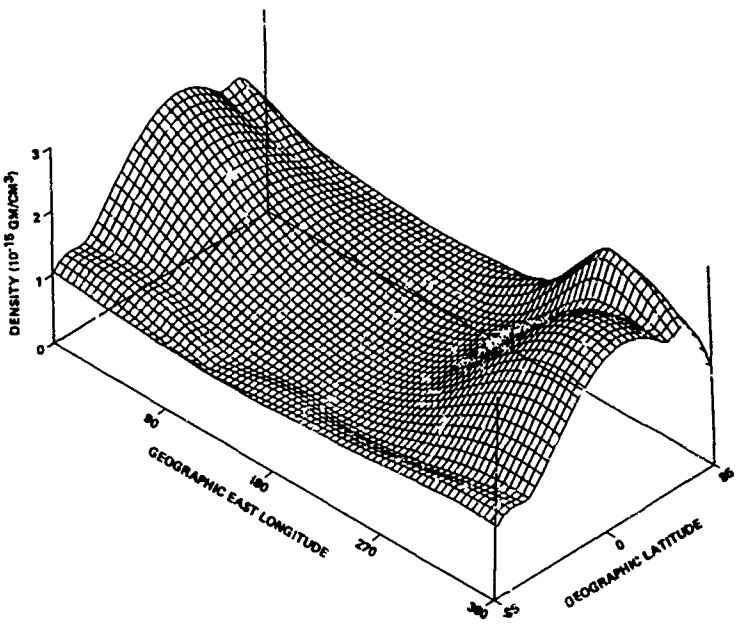


Figure 1. Density at 400 km near summer solstice at a universal time of 1600 hours ($F = \bar{F} = 115$).

2.2 Layer Formation

In addition to the inputs from magnetic field and atmospheric neutral density models, the model of N_e makes use of our knowledge of ionospheric layer formation and anomalies. The formulation of the N_e model is discussed below

Let S be the energy flux at location ℓ and S + dS the flux at location $\ell + d\ell$. Let σ be the absorption cross section of the atoms of the gas and N their number density (the neutral density of the upper atmosphere). The energy absorbed, dS, in a cylinder of unit cross section and axis

parallel to the direction of the incident beam is given by

$$dS = S\sigma N\, d\ell$$

$$\frac{dS}{S} = \sigma \int_{\infty}^{\ell_0} N d\ell = -\tau$$

where τ is the optical thickness of the atmosphere at location ℓ_0 along the path to the sun. As $S \to S_\infty$ (its value above the atmosphere) $\tau \to 0$.

Thus $$S = S_\infty \exp(-\tau),$$

and the energy absorbed per unit volume is

$$\frac{dS}{d\ell} = N\sigma S = N\sigma S_\infty \exp(-\tau).$$

If η is the ion/electron pair production efficiency per unit energy, then q (the number if ion/electron pairs produced per unit volume) is given by

$$q(\ell) = N\sigma\eta S_\infty \exp(-\tau).$$

The above equation gives the production of electrons. The two predominant loss mechanisms for ion/electron pairs are recombination and attachment. For recombination the loss equation is

$$\frac{dN_e}{dt} = -\alpha\, N_e\, N^+$$

where N_e is the electron density, N^+ the positive ion density, and α the recombination coefficient.

Thus, at equilibrium

$$\frac{dN_e}{dt} = q - \alpha\, N_e\, N^+ = 0.$$

Then, assuming that $N^+ \approx N_e$,

$$N\sigma\eta\, S_\infty \exp(-\tau) - \alpha\, N_e^2 = 0.$$

Therefore, $N_e = \left[\frac{N\sigma\eta\, S_\infty}{\alpha} \exp(-\tau)\right]^{1/2}$.

For attachment the loss equation is

$$\frac{dN_e}{dt} = -b\, N_e\, N$$

In discussing attachment processes in the ionosphere, it is generally assumed that the number of neutrals is much greater than the ions so that when ionization occurs, the neutral density does not change significantly. Thus

$$\frac{dN_e}{dt} = -\beta N_e$$

is generally used, where β is the attachment coefficient.

Then, at equilibrium

$$\frac{dN_e}{dt} = q - \beta N_e = 0 \text{ , and}$$

$$N_e = \frac{N\sigma n_\infty}{\beta} \exp(-\tau) .$$

In general, α and β may vary with height because the reactions usually involve three bodies instead of two.

2.2.1 The E and F_1 Layers. It has been observed that the E and the F_1 layers of the ionosphere can be described quite well using the recombination equation. Thus, using simple layer theory for the E and F_1 layers

$$N_e^E = [N S_E \alpha_E \exp(-\tau_E)]^{1/}]$$

$$N_e^{F_1} = [N S_{F_1} \alpha_{F_1} \exp(-\tau_{F_1})]^{1/2}$$

where N_e^E and $N_e^{F_1}$ are the electron concentrations for the E and F_1 layers, S^E and S^{F_1} are the solar UV flux affecting the E and F_1 layers, α_E and α_{F_1} are the recombination coefficients, N the neutral density, and τ_E is given by

$$\tau_E = \int_\infty^{\ell_o} \sigma_E N \, d\ell \approx \sigma_E \int_\infty^{\ell_o} N \, d\ell$$

where $T \equiv \int_\infty^{\ell_o} N \, d\ell$ is the total atmospheric cross section between the observation point and the sun.

The MDAC atmospheric density model (described in 2.1) is used in the evaluation of the above integral.

Then $$\tau_E = \sigma_E T$$

and $$\tau_{F_1} = \sigma_{F_1} T$$

where σ_E and σ_{F_1} are the absorption coefficients for the wavelengths affecting the E and F_1 layers.

These equations give an approximate description of the observed structures in the E and F_1 layers. Changes in T depend on the neutral density and the location of the sun. Thus, by using the MDAC neutral density model (which contains most of the large scale variations observed in the neutral atmosphere) and a simple sun-position program, the temporal variations observed in the E and F_1 layers are accurately described.

The F_1 layer is essentially turned off at local sunset but the E layer, while diminished from its daytime strength, persists throughout the night. Thus

$$N_e^E \text{ (night)} = [N S_E' \alpha_E \exp(-\tau_E')]^{1/2}$$

where S_E' is the intensity of scattered light and

$$\tau_E' = \sigma_E \int_{\infty}^{\ell_o} N \, d\ell$$

where $\int_{\infty}^{\ell_o} N \, d\ell$ is the optical thickness (integrated vertically - not toward the sun,.

Profiles of N_e at noon and midnight, as represented by this simple model, are shown at the equator and mid latitudes in Figure 2.

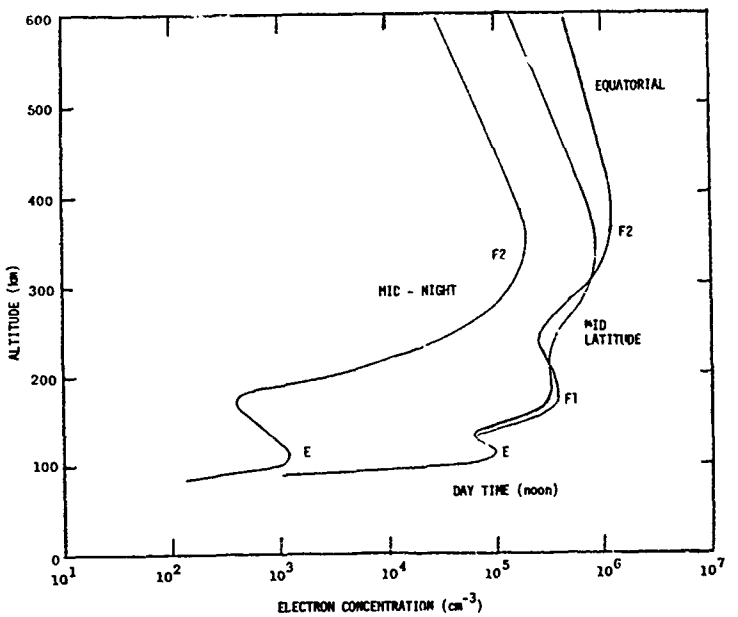


Figure 2. Electron density profiles using a single spectral line for the E and F_1 layers

In Figure 2 we note the rather large unrealistic dip between the E and F_1, and F_1 and F_2 layers. This large decrease in electron density between the layers is due to the use of a single wavelength to describe each layer. This single wavelength description is, of course, an over simplification. The electromagnetic spectrum from the sun is quite complex and several wavelength bands are responsible for the layer formations.

The above analysis was then expanded to include several wavelength bands in the study of layer formation. The only time consuming calculation is the integral which calculates the optical thickness, τ_1. The calculated value of τ can be used for wavelengths having different absorption and ionization coefficients and source strengths, thus making a "multiband" wavelength analysis without extensive increases in computer time possible.

Although a large number of wavelengths can easily be used, it was found that 2 terms for E layer and 3 terms for the F_1 layer provide an excellent representation of altitude dependence. The model for the E and F_1 layers currently incorporates the following wavelength bands in the parameter list.

E Layer

Term 1	1026 °A
2	977 °A

F_1 Layer

Term 1	10-170 °A
2	170-796 °A
3	796-911 °A

The absorption coefficients associated with each of the bands has been adjusted to produce the electron density peak at the correct altitude and the production coefficient is adjusted to give the correct magnitude in intensity at the peak. As the incident intensities of the various wavelength bands change with time, the amplitude as well as the shape of the layers change. By monitoring the energy in the five wavelength bands with satellite instrumentation, a real time predictive capability is built into the model. The slow variation of the atmospheric density model (caused by the slower heating effects) is coupled with the rapid response to changes in the ionizing radiation.

Figure 3 is a sample altitude profile plot which combines the effects of the various wavelengths to produce a smoothly varying density profile.

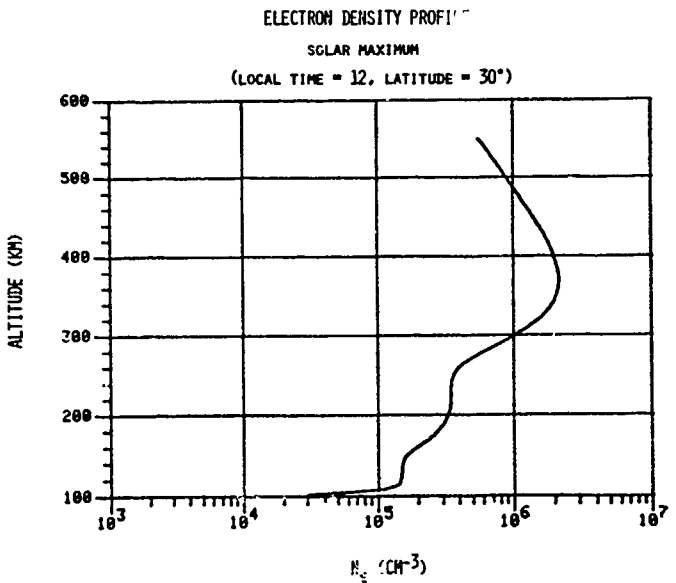


Figure 3. Electron density profiles using two spectral lines for the E layer and three spectral lines for the F_1 layer

2.2.2 The F_2 Layer. The structure and variability in the F_2 layer is not so directly controlled by neutral atmosphere number density. Attempts to treat the F_2 layer using U.V. generated ionization and electron losses which are a function only of density have ended in failure. It is generally agreed that the bottom of the F_2 layer is defined as the region where the electron loss rates change dramatically (from the order of seconds to the order of hours). Several studies were undertaken with neutral constituents and electron density dependent attachment and recombination coefficients. In none of these attempts was it possible to directly reproduce the observed seasonal, semi-annual or solar cycle dependencies.

Our first F_2 model contained an empirical function which arbitrarily (adjusted to provide a best fit at some point in time) defined a loss term varying with height and an ionization term depending only on the ambient density. Further empirical observations suggested that the normal F_2 layer was limited to the region equatorward of the latitude of the plasmasphere. Thus, a function was also introduced with limits the F_2 layer (except for the winter anomaly peak) to L values $\lesssim$ 4-6. (The L value can roughly be thought of as the equatorial extent of magnetic field lines measured from the center of the earth as given in earth radii). The winter and equatorial anomalies were also included as ad hoc functions.

This early function fitting for the F_2 layer proved to be only partially successful and much of the predictive ability for the F_2 region was lost. Thus we have recently attempted to include additional physics into the study of the F_2 layer. This reduces computational speed somewhat but does improve our ability to produce an F_2 layer with predictive capability.

F_2 Region: The F_2 region is formed by three distinct processes. The bottomside layer is a balance between photochemical production and loss, the peak is a balance between diffusion along field lines and chemical loss, and the topside layer is purely diffusive. The difficulty of modeling this region has been overcome by simplifying the physics of the three processes to the extent necessary to obtain an analytic solution to the continuity equation. The transition between bottomside and topside layers is then accomplished without any ad hoc assumptions and the layer's response to conditions other than those used to set free parameters will be correct.

Daytime ion chemistry: The major F_2 ion is $O^+(^4S)$. Using Torr and Torr (1978) to determine the ion reaction rates, and Jacchia (1977) to provide typical chemical concentrations; the production and loss rates of this ion have been estimated. At F_2 altitudes where production is important, photo-ionization is the dominant source of O^+. Assuming an optically thin atmosphere, this production is expressed as

$$P(O^+) = S[O] = S[O]_{Z_b} \, e^{-\left(\frac{Z-Z_b}{H_o}\right)}$$

where Z_b is the bottom boundary of the F_2 region and H_o is the atomic oxygen scale height. The loss of O^+ is through the reactions:

$$O^+ + N_2 \xrightarrow{K_1} NO^+ + O$$

$$O^+ + O_2 \xrightarrow{K_2} O_2^+ + O$$

where the K's are the temperature dependent rate coefficients and the product ions are rapidly destroyed. Because the N_2 and O_2 scale heights are nearly equal, the loss of O^+ is taken to be

$$L\,[O^+] = \{K_1[N_2]_{Z_b} + K_2\,[O_2]_{Z_b}\}\, e^{-\left(\frac{Z-Z_b}{H^*}\right)}\,[O^+]$$

where H* is an average scale height determined by forcing L to be exact at the top boundary.

Continuity equation: Letting n be O^+ concentration, the ion continuity equation is given by

$$\sin^2 I \, \frac{\partial}{\partial Z}\left\{D_a \, n \left(\frac{1}{n}\,\frac{\partial n}{\partial Z} + \frac{1}{2\,H_p}\right)\right\} + P(O^+) - L\,n = 0$$

where D_a is the ambipolar diffusion coefficient, H_p is the plasma scale height, and I is the dip

angle. D_2, P, and L have separate exponential dependences in the vertical; otherwise, the coefficients in the equation are constant. The analytic solution to this equation contains modified Bessel Functions, and integrals involving these functions. In the model, we approximate this solution by substituting a simple approximation for the Bessel functions, and using 6 point Gaussian quadratures in the integrals. The bottom boundary condition (at 200 km) is n given by photochemical equilibrium. The top boundary (taken at 100 km or at the magnetic equator intercept of the field line) has the condition that the diffusion velocity is zero.

The MSIS composition and neutral temperature model(Hedin, et al., 1977) is used to determine the scale heights and densities of O, N_2, and O_2. This chemistry model causes variations with local and universal time, latitude, longitude, season, and solar and magnetic activities in ion density. The coefficient S in production of O^+ can be tied directly to EUV flux measurements; however, at present a value of S is used which is tied linearly to the 10.7 cm index. A correction to S has been added for large zenith angles where the atmosphere is not optically thin. The value of H_p is most uncertain since electron and ion temperatures are more difficult to model that ion density. Parametric analyses show that the major effect of different H_p is in the topside layer. A value for H_p was chosen that varies with magnetic latitude only. It has a maximum at mid-latitudes, and minima at the equator and the poles as observations of electron temperature suggest.

The results using the diffusive effect and the MSIS model (to include some of the chemical effects) have been encouraging. The model produces the correct altitude dependence; the winter anomaly; semi-annual, anual, and solar cycle variations. Figure 4 shows the E, F, and F_2 altitude profiles for solar minimum and solar maximum. Figure 5 is a plot of the critical frequency over a solar cycle at an altitude of 310 km, near the F_2 peak. The variations which are a combination of the annual, semi-annual, and solar cycle variations agree well with observations.

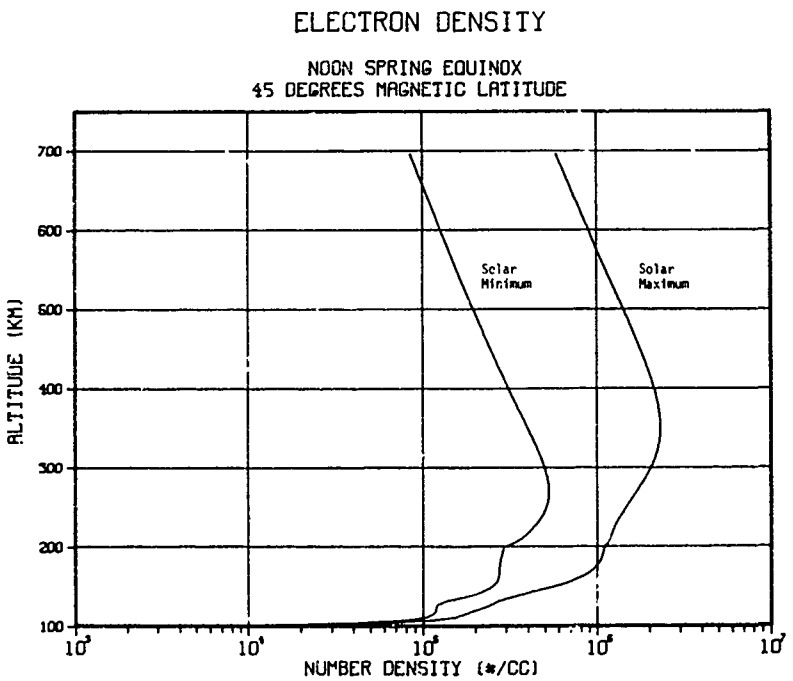


Figure 4. Electron density profiles using the F_2 diffusion equations

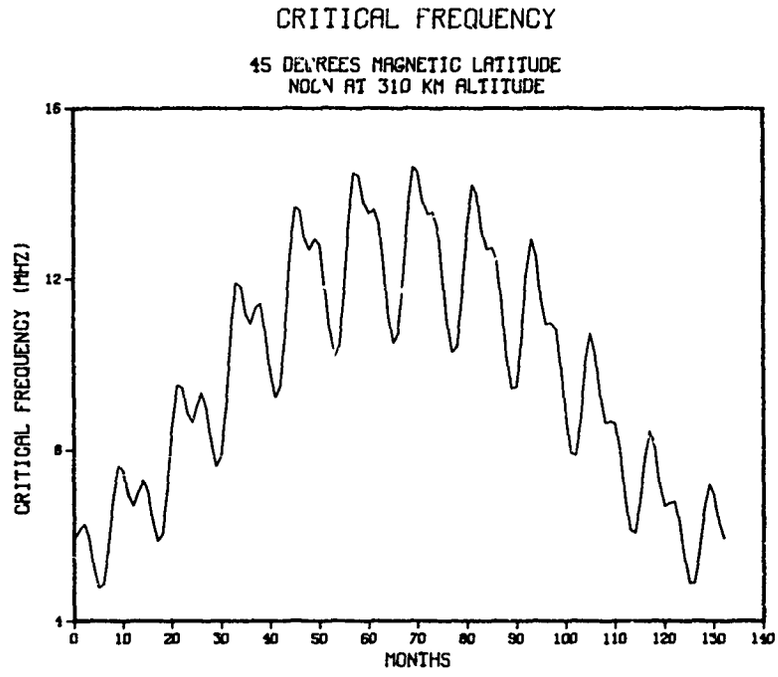


Figure 5. Critical frequency over a solar cycle near the F_2 peak

REFERENCES

Bruce, R. W., Upper Atmosphere Density Determined from a Low-g Accelerometer on Satellite 1967-50 B, Rep. TOR-0158 (3110-01)-16, Aerospace Corporation, El Segundo, California, 1968.

Hedin, A. E., C. A. Reber, G. P. Newton, N. W. Spencer, H. C. Brinton, H. G. Mayr, and W. E. Potter, A Global Thermospheric Model Based on Mass Spectrometer and Incoherent Scatter Data, MSIS 2. Composition, J. Geophys. Res., 82, 2148, 1977.

Hedin, A. E., J. E. Sala, J. V. Evans, C. A. Reber, G. P. Newton, N. W. Spencer, N. C. Kayser, D. Alcayde, P. Bauer, L. Cogger, and J. P. McClure, A Global Thermospheric Model Based on Mass Spectrometer and Incoherent Scatter Data, MSIS 1. N_2 Density and Temperature, J. Geophys. Res., 82, 2139, 1977.

Jacchia, L. G., Thermospheric Temperature, Density, and Composition: New Models, Spec. Rep. 375, Smithsonian Atrophys. Observ., Cambridge Massachusetts, 1977.

Torr, D. G., and M. R. Torr, Review of Rate Coefficients of Ionic Reactions Determined from Measurements Made by the Atmosphere Explorer Satellites, Rev. Geophys. Space Phys., 16, 327, 1978.

EMPIRICAL F-REGION MODEL DEVELOPMENT BASED ON S3-1 SATELLITE MEASUREMENTS

C. R. Philbrick
Air Force Geophysics Laboratory
Hanscom AFB, MA 01731

ABSTRACT

The ion mass spectrometer experiment on the S3-1 satellite collected measurements of the density of N^+, O^+, N_2^+, NO^+, and O_2^+ in the altitude region 150 to 500 km for a period of approximately six months. Measurements from approximately 1800 orbits, which cover all latitudes for both summer and winter and four local time periods, have been studied to characterize the ionospheric variations. Mean profiles and median profiles with upper and lower quartile values, which show the data range, have been deduced from the measurements. The variation of the ion species densities with latitude, solar zenith angle, season, geomagnetic activity, and altitude has been compared with ionospheric models, such as the International Reference Ionosphere (IRI). Differences in profile shapes and relative composition from the IRI model are discussed. The response of the ionosphere to geomagnetic storm effects is shown in several cases where molecular ion densities can become dominant over atomic ion species to altitudes above 300 km.

INTRODUCTION

Several studies have been carried out toward the development of a useful model for predicting the mean ionospheric properties based upon various collections of data. The Bent,[1] CCIR[2] and IRI[3] models rely heavily on the extensive data which have been collected by ground-based ionosondes and satellite top-side sounders. Some in situ electron and ion density profiles from rockets and satellites have also been considered in formulating these models. The parameter which is best described by the models is the electron density near the peak of the F_2-region, which is directly determined from the plasma critical frequency. However, there is significant uncertainty in the shape of the density profiles, particularly in the lower F-region. The only effort to date to empirically model ion composition has been that of the IRI model which is based entirely upon the summary of 41 rocket probes which was prepared by Danilov and Semenov.[4]

Using the data base from the S3-1 satellite mass spectrometer, the initial stages of an empirical model for the altitude region between 150 and 500 km have been completed. This paper is an interim progress report on this development. Examples of the data, a general description of the modeling approach and initial results from the investigation are presented.

EXPERIMENT MEASUREMENTS

The mass spectrometer experiment on the S3-1 satellite[5,6,7] provided a data base of about 1800 orbits of ion density measurements. The density of the five primary ions, N^+, O^+, N_2^+, NO^+ and O_2^+, was measured each second, but this study only considers those measurements made each 10 seconds when the instrument axis was most closely aligned with the direction of motion of the spinning satellite. A comparison study of the ion density measured by the S3-1 as it passed through the F_2 peak with ground-based ionosonde measurements of the f_oF_2 provided an opportunity to test and calibrate the instrument performance. A total of 73 cases were located where the satellite was passing through the F_2 peak within a 5°x5° latitude and longitude box centered on the ionosonde station, and an f_oF_2 measurement was available within 15 minutes of the satellite overpass. No trend which would indicate a long term sensitivity change was observed and the standard deviation of the comparison was 26%. In another study,[7] a direct comparison was made for five nearly coincident measurements with the AEROS-B impedance probe. The ratio between the measurements from the two satellites was $1.00 \pm .06$. Based on these studies and the laboratory calibrations, the reported ion densities should be accurate to better than $\pm$ 15% and the relative ion composition accurate to better than $\pm$ 10% for densities greater than 5×10^2 cm^{-3}.

Figure 1 shows a ..ple of part of an orbit near perigee for daytime conditions. This orbit is typical for a geomagnetically quiet period even though some wave structure is observed in the southern midlatitude region. Note that the N^+ typically follows the O^+ response but at a density of about 1% of the O^+ on the day side. Also, the molecular ions have a generally similar behavior. Figures 2 to 4 show examples of some of the features which are commonly observed in the data set. These examples are included to indicate the quality of the data, show the variability of the ionosphere, and to provide clear examples of three of the features which occur in the F-region.

Figure 2 shows the results for orbit 1596 which occurred 15 orbits (less than one day) after the results of Figure 1, but these results were obtained near the peak of a geomagnetic storm corresponding to a K_p=7. The left-hand side of the figure corresponds to latitudes within $\pm$ 30° of the geomagnetic equator and the character of the measured profiles is similar to that of Figure 1 in the F_1- and F_2-regions. However, the right-hand side of the figure, which corresponds to geomagnetic latitudes between 30° and 70°, shows the strong effect of the storm. The F_2 peak vanished, primarily because of the loss of O^+ due to the change in the $O^+ + N_2 \rightarrow NO^+ + N$ and $O^+ + O_2 \rightarrow O_2^+ + O$ reaction rates (followed by dissociative ion-electron recombination) in the presence of strong electric fields[8] and the large enhancement in the densities of N_2 and O_2 in the thermosphere. The The N_2^+ ion density is increased due to the enhanced N_2 density and ionization rate from precipating electrons. Between 60° and 70° geomagnetic latitude large spatial irregularities are observed with the O^+ density changing by more than a factor of 10 between measurements which are spaced horizontally by about 70 km. During this geomagnetic storm and others that have been examined, the cross-over from molecular to atomic ion predominance has been found as high as 400 km, instead of the normal case which is near 180 km.

Figure 3 shows an example of nighttime measurements which are characterized by the steep density gradient on the bottom side of the F_2-region which generally becomes steeper at the lower geomagnetic latitudes. This orbit also shows the characteristic of the fountain effect, or Appleton anomaly effect [9], over the magnetic equator. The molecular ion profiles exhibit the transport upward from the E-region into the F_1-region directly over the magnetic equator.

Figure 4 gives an example of the ion density as the satellite passes through a strong auroral disturbance near the satellite's perigee. Over a range of a few degrees of latitude, the ion density is increased by more than a factor of 20 due to ionization by energetic particle precipitation.

MODELING APPROACH

The measurements for each ion species and for the total density have been divided into data bins with the following set of ranges.

Altitude (km): 150-160, 160-170,..., 280-290, 290-310, 310-330,...490-510.

Solar Zenith Angle: < 60 , 60°-70°, 70°-90°, 90°-110°, > 110°

Magnetic Index: $0 \leq K_p < 2$, $2 \leq K_p < 4$, $K_p > 4$.

Latitude (geomagnetic): -90 to -70, -70 to -60, -60 to -45, -45 to -15, -15 to 0, 0 to +15, +15 to +45, +45 to +60, +60 to +70, +70 to +90.

Season: Nov-Feb, Mar-Apr, May-Aug, Sep-Oct.

Studies were also made by subdividing the data further to examine the dependence on sunspot number, AE index and DST index.

From examination of the data it was apparent that in order to develop a model of mean conditions, it would be necessary to eliminate some of the measurements which represented the more unusual ionospheric conditions. These measurements which were outside of two standard deviations from the log mean of each bin were eliminated. Most of the analysis has included the calculation of the mean, median, and upper and lower quartile values for each data bin. The data bins that have been used generally contain between 50 and 800 measurements. Figure 5 shows an example of the mean curves for the total ion density in the latitude ranges from 0° to 15° and from 15° to 45° under nighttime conditions. Each point represents the mean of the logarithmic values for that data bin and the bars show the range of the upper and lower quartile values. In Figure 6 the altitude profiles between 150 and 270 km are shown for the five latitude divisions in the winter hemisphere at night. The curves show a strong variation with latitude. At higher altitudes, near the F_2 peak, the highest density is near the equator and this profile has the strongest altitude gradient. The profile, corresponding to the auroral oval, 60° to 70°, is nearly independent of altitude. The major fea-

tures of this latitude distribution are due to the constrainments on the ion diffusion due to the earth's magnetic field.

Figures 7 and 8 show the mean profiles for the ion species and the total ion density at midlatitudes under the conditions of winter night and summer day. Also shown in these figures is the appropriate IRI model mean profile for total ion density at 30° latitude. The mean of each independent data bin, represented by points on these curves, form relatively smooth profiles because of the large number of measurements within each bin. However, the shape of the IRI model profile is considerably different from the S3-1 profile in the lower F-region at night. The agreement between the IRI and S3-1 profiles is much better for the daytime comparison. The curves shown in Figures 7 and 8 represent the midlatitude mean densities. Examination of these profiles indicates the difficulty in obtaining suitable analytic expressions for representing the global mean features in the ion species densities. From our efforts to date, it appears that the total ion density profile for the S3-1 mean conditions can probably be represented analytically. Possibly these shape functions can be used with the much larger data base of other models, which contain a full range of solar and geophysical variations, by normalizing to the F_2 peak density.

CONCLUSIONS

The area where the mass spectrometer measurements of S3-1 and other satellites can make a unique contribution to the modeling efforts is in describing the relative ion composition. Figures 9(a-d) show the relative ion composition for summer day and winter night conditions at midlatitude and high latitude for the five principle ion species. Examination of these figures lead to several interesting conclusions regarding the ion chemistry processes and the neutral species distributions, but these conclusions will not be discussed in this paper. The present reason for examining these cases of relative ion composition is to show that the species densities could be obtained from interpolation within a reasonable size table if an adequate total density model profile is available. This approach of using relative ion composition is the method currently used in obtaining species profiles from the IRI model. Some comparisons of the S3-1 mean ion composition and the IRI composition model were made in a previous paper.[6]

ACKNOWLEDGEMENTS

The major effort of organizing and analyzing the data has been done by K.H. Bhavnani. The author gratefully acknowledges the efforts of R.E. McInerney, D. Delorey, E.J. Ziema, M.E. Gardner and discussions with K.S.W. Champion. The 1980 version of the IRI model (IRIF07) was kindly supplied by K. Rawer. Data used in connection with this article were obtained from WDC-A solar-terrestrial physics (ionosphere).

REFERENCES

1. Llewellyn, S.K. and R. B. Bent, Documentation and description of the Bent ionospheric model, AFCRL-TR-73-0657, 1973.

2. C.C.I.R. Report 340, Atlas of ionospheric characteristics, Union Internationale des Telecommunications, Geneve, 1967.

3. Rawer, K., S. Ramakrishnan and D. Bilitza, International Reference Ionosphere, 1978, U.R.S.I. Brussels, 1978 and private communication, 1980.

4. Danilov, A.D. and V.K. Semenov, Relative ion composition model at midlatitudes, J. Atmos. Terr. Phys., 40, 1093, 1978.

5. Philbrick, C.R., Recent satellite measurements of upper atmospheric composition, Space Research XVI, 289, Akademie-Verlag, Berlin, 1976.

6. Philbrick, C.R., P. Lammerzahl, E. Neske and A. Dumbs, Comparison between plasma densities measured with the AEROS-B and S3-1 satellites and the IRI model, COSPAR Workshop, Budapest, Hungary, submitted for publication, 1980.

7. Philbrick, C.R., P. Lammerzahl, E. Neske and A. Dumbs, Comparison of plasma densities in the daytime polar F-region measured by the AEROS-B and S3-1 satellites, submitted for publication JATP, 1980.

8. Schunk, R.W., W.J. Raitt and P.M. Banks, Effects of electric fields on the daytime high-latitude E and F regions, J. Geophys. Res. 80, 3121, 1975.

9. Appleton, E.V., Two anomalies in the ionosphere, Nature 157, 691, 1946.

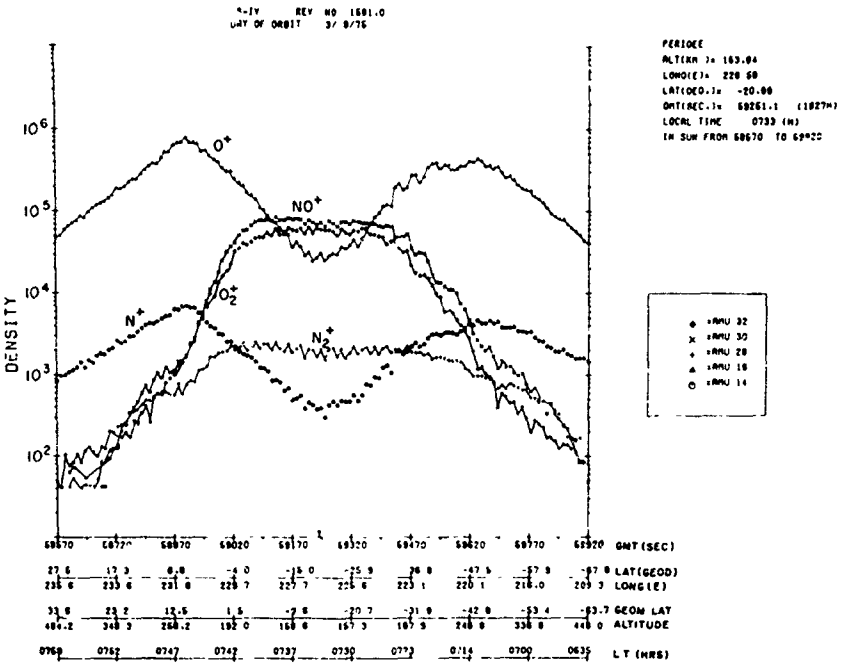


Figure 1. The species density (ions/cm^3) profiles for N^+(14 amu), O^+(16 amu), N_2^+(28 amu), NO^+(30 amu) and O^+(32 amu) measured by the S3-1 on orbit 1581 are shown. The scales show the GMT total seconds, latitude, longitude, geomagnetic latitude, altitude and local time.

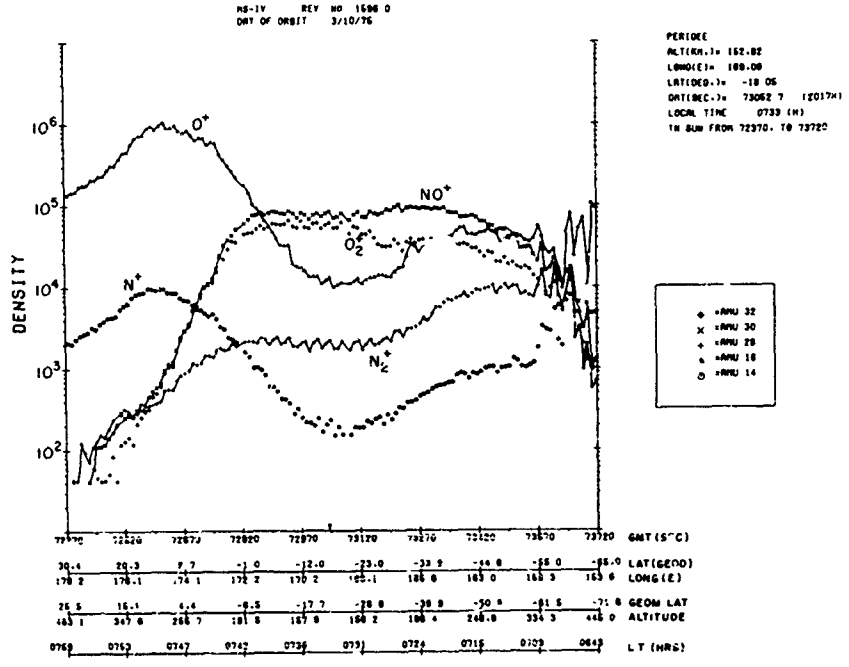


Figure 2. Species profiles similar to those of Figure 1 but for conditions corresponding to a geomagnetic storm.

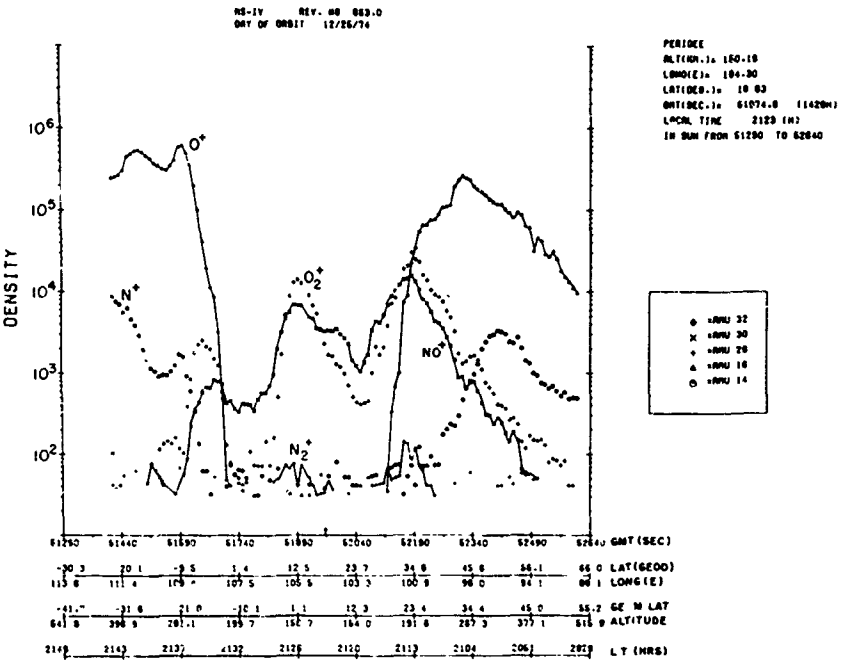


Figure 3. Species profile measurements similar to those of Figure 1 but for nighttime conditions with perigee centered near the magnetic equator.

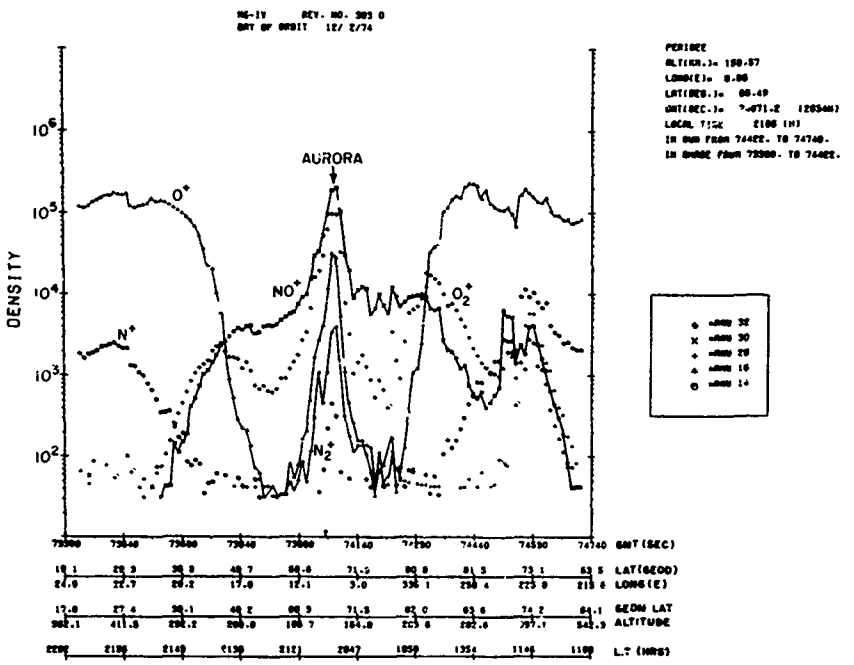


Figure 4. Species profiles similar to those of Figure 1 but for a period of enhanced auroral zone activity.

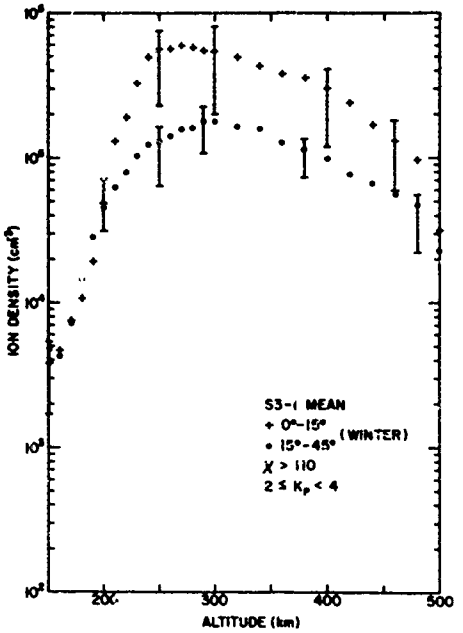


Figure 5. The mean ion density profiles for two latitude regions are shown together with bars representing the range of the upper and lower quartiles for the data bin.

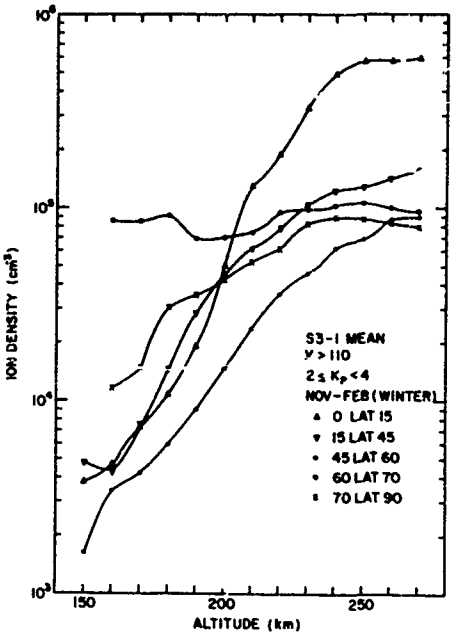


Figure 6. The mean ion density profiles are shown for five latitude range in the winter hemisphere at night.

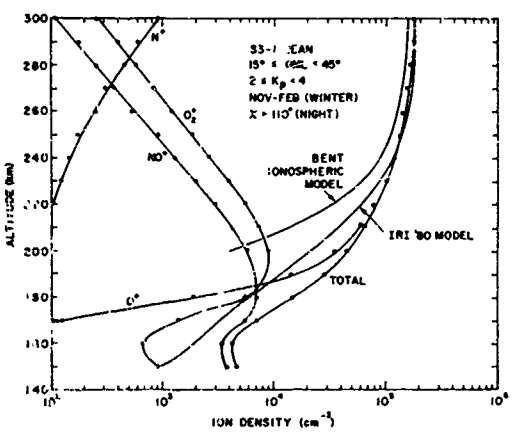


Figure 7. The mean ion species and total density profiles are shown for midlatitude winter nighttime conditions together with the appropriate profile for total ion density from the IRI 80 model.

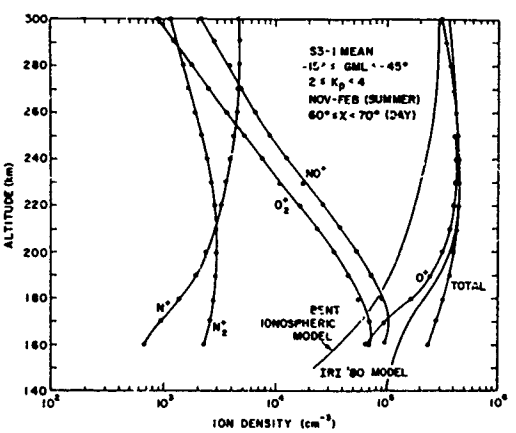


Figure 8. Ion species and total density profiles similar to those of Figure 7 are shown for summer daytime conditions.

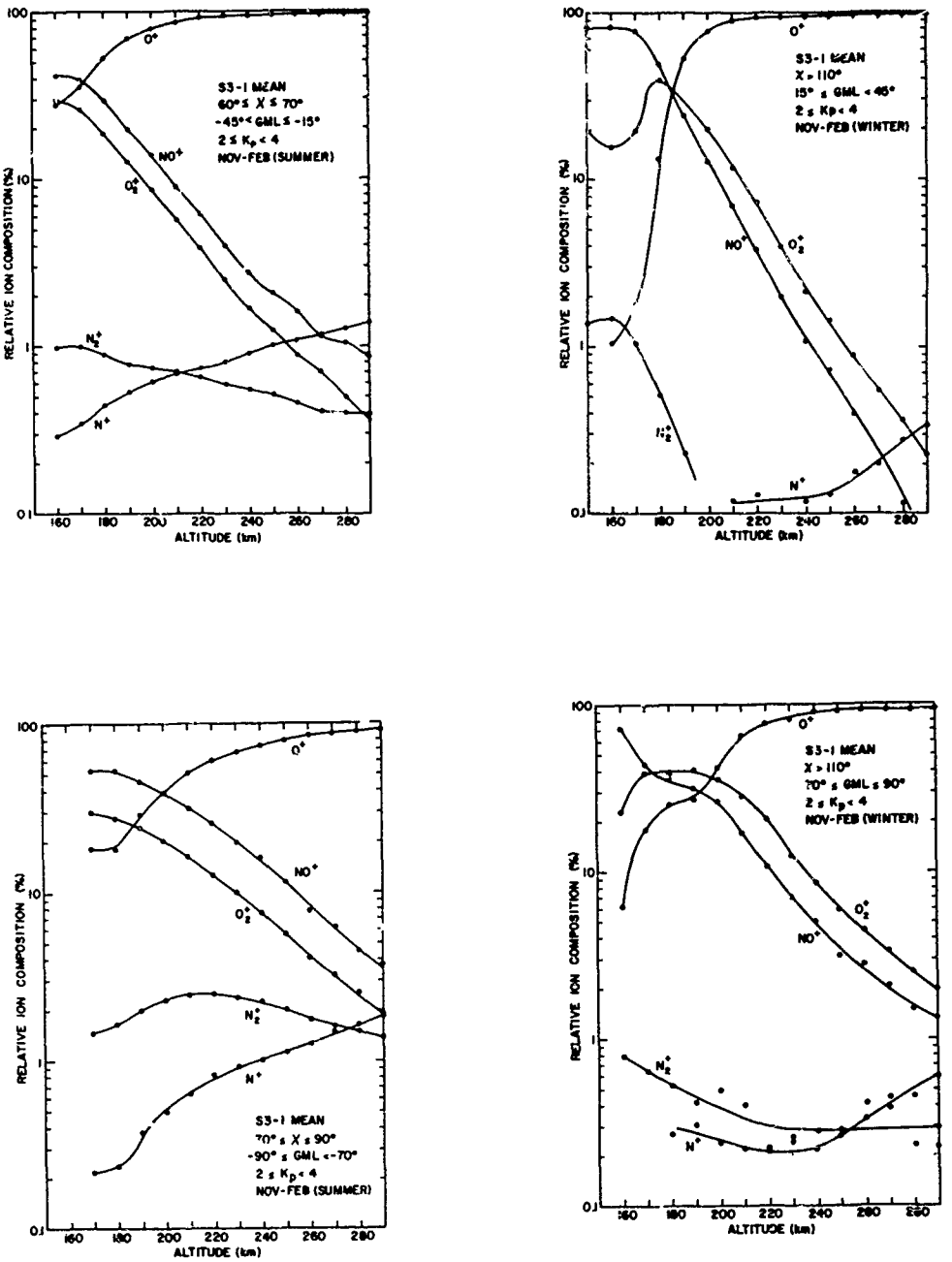


Figure 9. Relative ion composition percentages are shown for the F_1 and lower F_2 regions for the five species measured by the S3-1 satellite for four cases, (a) mid-latitude, day, summer; (b) midlatitude, night, winter; (c)high-latitude,day, summer; (d) high-latitude,night, winter.

SIMULATED PROPAGATION EFFECTS ON TRANSIONOSPHERIC RADIO WAVES

K. C. YEH AND C. H. LIU
DEPARTMENT OF ELECTRICAL ENGINEERING
UNIVERSITY OF ILLINOIS AT URBANA-CHAMPAIGN
URBANA, ILLINOIS 61801

ABSTRACT

An ionospheric model perturbed by nuclear explosions is constructed by superposing three bubbles. Each bubble has a horizontal structure similar to that measured in situ in a natural ionosphere but displaced relative to each other by a small horizontal distance to simulate 50 km to 100 km vertical correlation distance. The propagation effects of this model ionosphere on the wave are obtained by solving the parabolic equation numerically for the complex amplitude. The results show that large amplitude and phase fluctuations are obtained at gigafrequencies and some scintillations are obtained even at a frequency as high as 15 GHz. Properties of amplitude and phase scintillations are studied and shown in the paper.

1. INTRODUCTION

Current evidence from various experiments indicates that the electron density irregularities associated with the equatorial ionospheric bubbles have highly spiky structures (Kelley et al., 1976; McClure et al., 1977; Basu and Kelley, 1979; Morse et al., 1977). Inside the bubbles, the ionization structures are highly complex with steep gradients. Correlative measurements show the occurrence of intense scintillations whenever radio signals traverses the bubble (Basu et al., 1977; Yeh et al., 1979). Using the in situ data as a basis, a model bubble constructed by Wernik et al. (1980) has been shown to be a strong cause of scintillations. Their simulation studies clearly show outbursts of large amplitude excursions whenever the ray from the transmitter to the receiver intersects the spiky ionization structure. This is especially apparent at gigahertz frequencies. These earlier simulations were carried out under ionospheric conditions that can occur naturally. We wish to extend these studies to ionospheric conditions that may exist in man-made events such as atmospheric nuclear explosions.

In the following we first describe the nature of the model ionosphere. Effects on a radio wave propagating through such a model are simulated by solving numerically the parabolic equation. The fluctuations in amplitude and phase are then studied at a number of frequencies. Finally the paper is concluded in the last section.

2. MODEL IONOSPHERE

The ionospheric model used in this paper is essentially a modification of the one used by Wernik et al. (1980). The horizontal structure is unchanged; it is still based on the in situ data (McClure et al., 1977). However, important modifications have been made in two aspects: the vertical extent of the bubble and the vertical correlation distance of the spiky structures. To accomplish both of these aspects we have effectively superposed three bubbles. The top most bubble has a vertical extension of 100 km. The middle bubble has a vertical extension of 50 km and its horizontal structures are displaced horizontally by a distance of 1.26 km. The lowest bubble also has a vertical extension of 50 km and its horizontal structures are displaced horizontally by 2.52 km. Since most of the spiky structures have dimensions less than 1 km, the relative displacement of bubbles by 1.26 km makes these structures in one bubble to be uncorrelated with structures in other

bubbles. These three bubbles are assumed to exist in a background ionosphere of vertical extent 450 km. The maximum electron density of this background ionosphere is 2.8×10^{11} electrons/m^3 which corresponds to a plasma frequency of 4.75 MHz. The integrated electron density or electron content through these bubbles has the shape shown in Fig. 1. The sizable depletion of electron content at a horizontal distance equal to 0 km as compared to 20 km is obvious. The spiky structures are still there. The electron content near the edge of the bubble is approximately 8.9×10^{16} electrons/m^2, a moderate value. The ionospheric slab has therefore a value of 320 km for the background ionosphere.

In our computer simulations we let a wave of a specific radio frequency be incident on the top of the model ionosphere. The wave propagates in the model ionosphere just described for 450 km and exits from the ionosphere at the lower boundary. The wave propagates further in free space for 250 km or 550 km. The total propagation distance from the top of the ionosphere to the ground-based receiver is therefore either 700 km or 1000 km, but in each case scattering from ionization irregularities can take place only within the first 450 km. However, since the difference in behavior is not very large between these two cases (except the scintillation is stronger on the longer path), we will present results only for 700 km propagation distance. In the simulation the forward scatter assumption is made. The complex amplitude under this assumption satisfies a parabolic equation. The numerical scheme used in solving this parabolic equation has been given in an earlier paper (Wernik et al., 1980) and will not be repeated again.

3. HIGH FREQUENCY BEHAVIOR

In our simulation studies computations were made for six radio frequencies, viz 800 MHz, 1.5 GHz, 2.5 GHz, 4 GHz, 7.5 GHz and 15 GHz. On physical grounds the scintillation is expected to be weakest on the highest frequency, i.e. 15 GHz. We discuss results computed for this highest frequency first.

The amplitude and phase received on the ground as a function of horizontal distance are shown in Figs. 2(a) and 2(b) respectively. The amplitude fluctuations are small and become even smaller for horizontal distances larger than 15 km. This distance at 15 km can be viewed as the boundary of the bubble. Within the bubble boundary the S_4 scintillation index is 0.038. The phase of the 15 GHz signal varies with horizontal distance in a manner shown in Fig. 2(b). The optical phase for which diffraction effects are absent is

$$\phi_{op}(x) = \phi_0+(\omega/c) \int n(x,z)dz \quad \text{rad} \tag{1}$$

where ϕ_0 is some reference phase, ω is the angular frequency, c is the free space velocity of light and n is the refractive index assumed to be a function of horizontal distance x and vertical distance z. In the high frequency approximation, the refractive index can be expanded, yielding from (1)

$$\phi_{op}(x) \cong - \frac{84.4\times10^{-8}}{f} [N_T(x)-N_T(0)] \quad \text{rad} \tag{2}$$

where f is the frequency and $N_T(x)$ is the total electron content in SI units. In writing (2) the phase reference is selected at x=0 as done in Fig. 2(b). For the chosen ionospheric model the difference $N_T(25)-N_T(0)$ is 0.67×10^{17} electrons/m^2 as obtained from Fig. 1. The optical phase at x=25 km can be computed easily for 15 GHz by using (2) to be 1.2π in agreement with the results obtained in Fig. 2(b). Therefore, at such a high frequency, diffraction effects are unimportant. The phase perturbations are produced almost entirely by changes in the optical path. As a matter of fact the standard deviation of phase departure from the optical path is computed to be only 0.0086π radians. By closely comparing the amplitude data of Fig. 2(a) against the phase data of Fig. 2(b), one can note that large amplitude fluctuations occur exactly at horizontal distances of large phase change or electron content change. This is the effect noted earlier by Wernik et al. (1980) as being caused by diffraction from sharp irregularity gradients. Such effects are most pronounced at the highest frequencies.

4. AMPLITUDE AND PHASE FLUCTUATIONS

The amplitude fluctuations for all six radio frequencies are shown in Figs. 3(a) through 3(f). The decrease in fluctuations with the increased radio frequency is obvious. The frequency dependence of the scintillation index S_4 has the behavior shown in Fig. 4. In computing the scintillation index the portion of the data from 15 km to 25 km is excluded, since the bubble edge obviously renders the amplitude fluctuations nonstationary. This is particularly obvious in Fig. 3 for 2.5 GHz and higher frequencies. In Fig. 4 at the high frequency limit the S_4 curve approaches asymptotically to a dependence $f^{-1.5}$. As the frequency is lowered this dependence is weakened. Such a behavior is in general agreement with other computational results (Yeh et al., 1975) and experimental results (Umeki et al., 1977), but it should be pointed out that the computations in this paper are made by assuming irregularities to be deterministic while other computations use a stochastic approach.

The phase fluctuations for four radio frequencies are shown in Fig. 5(a) through 5(d). In computing the phase one always encounters the problem of 2π radian phase ambiguity. To resolve this ambiguity problem, we have established several criteria: the change of phase is less than 2π radians in two neighboring points (a distance of 42 meters), and the departure of phase from the optical law (1) at two neighboring frequencies is less than 2π radians. However, as the frequencies are lowered to below 2.5 GHz, the computed phase jumps rapidly between the two neighboring points and two neighboring frequencies that the restoration of phase to the correct value cannot be done with confidence. This is why the phase data for two lower frequencies are not presented. The phase fluctuations shown in Fig. 5(a) are dominated by the optical term which has f^{-1} dependence. The departure from this optical behavior is caused by diffraction. This departure has a standard deviation for various radio frequencies plotted in Fig. 6. A straight line with f^{-2} dependence is also drawn for comparison.

In applications to certain communication problems, one would like to know how frequently large departures in phase from the optical path can occur. This departure can be called the phase error. For this purpose histograms are prepared to show the occurrence frequency with which a particular phase error interval would occur. They are depicted in Fig. 7. For unknown reasons the histograms are slightly skewed.

CONCLUSION

We have simulated the propagation effects on a computer after the wave has propagated a distance of 700 km from the top of the ionosphere. The model ionosphere exists only in the first 450 km with a maximum electron density equal to 2.8×10^{11} electrons/m^3 and an electron content value of 8.9×10^{16} electrons/m^2 (corresponding to a slab thickness of 320 km). Superposed on this background ionosphere is a region of depleted electron densities with a maximum depletion equal to 6.7×10^{16} electrons/m^2. Inside the depleted region there exist sharp gradients that have vertical correlation distances equal to 50 km to 100 km. After propagating through such a model ionosphere, we have found the amplitude scintillations to be appreciable at 4 GHz (S_4=0.26) and clearly discernible even at a frequency as high as 15 GHz (S_4=0.0025), at and above which tropospheric effects are expected to play an important role. The dominating component of the phase fluctuation has its origin in the variation of the optical path which has the inverse frequency dependence. The phase departure from the optical path has a rough f^{-2} dependence. At 4 GHz, the standard deviation of phase departure is 0.09π radians or 16 degrees.

ACKNOWLEDGEMENT

The authors benefited greatly from discussions with Leon A. Wittwer of Defense Nuclear Agency. The work reported in this paper was supported by the Air Force Geophysics Laboratory under contract F19628-78-C-0195.

REFERENCES

Basu, S., and M. C. Kelley (1979), A review of recent observations of equatorial scintillations and their relationship to current theories of F-region irregularity generation, Radio Sci., 14, 471-485.

Basu, S., J. Aarons, J. P. McClure, C. Lahoz, A. Bushby, and R. F. Woodman (1977), Preliminary comparisons of VHF radar maps of F-region irregularities with scintillations in the equatorial region, J. Atmos. Terr. Phys., 39, 1251-1262.

Kelley, M. C., G. Haerendel, H. Kappler, A. Valenguela, B. B. Balsley, D. A. Carter, W. L. Ecklund, C. W. Carlson, B. Havsler and R. Torbert, Evidence of Rayleigh-Taylor type instability and upwelling of depleted density regions during equatorial spread F, Geophys. Res. Lett., 3, 448, 1976.

McClure, J. P., W. B. Hanson, and J. H. Hoffman (1977), Plasma bubbles and irregularities in the equatorial ionosphere, J. Geophys. Res., 82, 2650-2656.

Morse, F. A., B. C. Edgar, H. C. Koons, C. J. Rice, W. J. Heikkila, J. H. Hoffman, B. A. Tinsley, J. D. Winningham, A. B. Christensen, R. F. Woodman, J. Pomalaza, and N. R. Teixeira (1977), Equion: An equatorial ionospheric irregularity experiment, J. Geophys. Res., 82, 578-592.

Umeki, R., C. H. Liu, and K. C. Yeh (1977b), Multifrequency studies of ionospheric scintillations, Radio Sci., 12, 311-317.

Wernik, A. W., C. H. Liu and K. C. Yeh, Model computations of radio wave scintillation caused by equatorial ionospheric bubbles, Radio Sci., 15, 559-572, 1980.

Yeh, K. C., C. H. Liu, and M. Y. Youakim (1975), A theoretical study of the ionospheric scintillation behavior caused by multiple scattering, Radio Sci., 10, 97-106.

Yeh, K. C., H. Soicher and C. H. Liu, Observations of equatorial ionospheric bubbles by the radio propagation method, J. Geophys. Res., 84, 6589-6594, 1979.

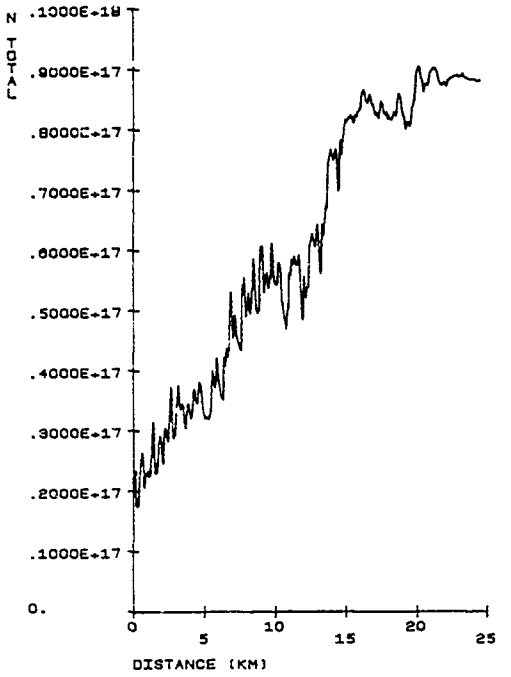
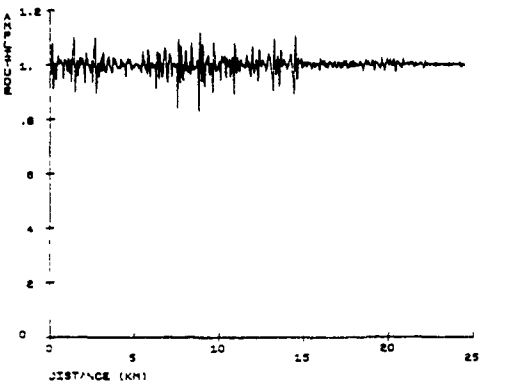
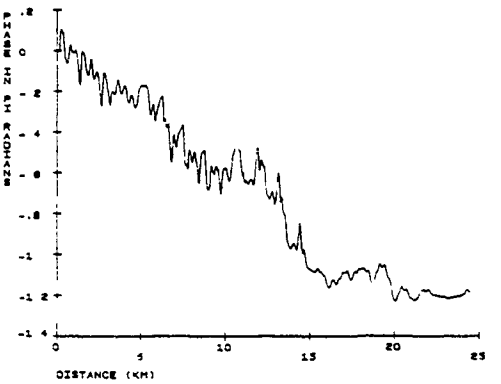


Fig. 1 Vertical total electron content in electrons/m^2 as a function of horizontal distance for a model ionosphere. The model is constructed by superposing three bubbles vertically with a slight horizontal shift to reduce the vertical correlation length.

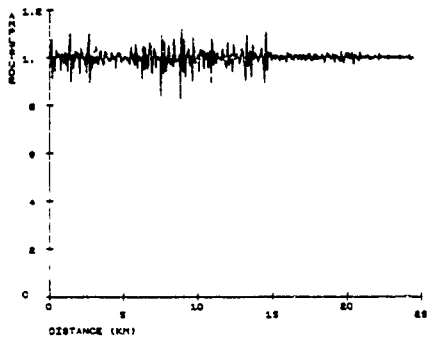


(a)

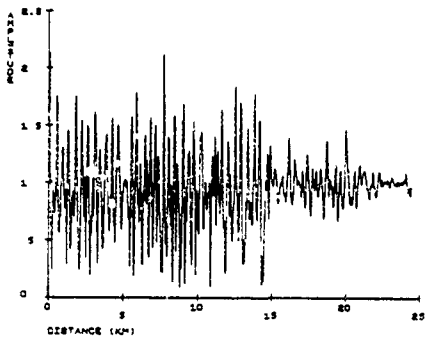


(b)

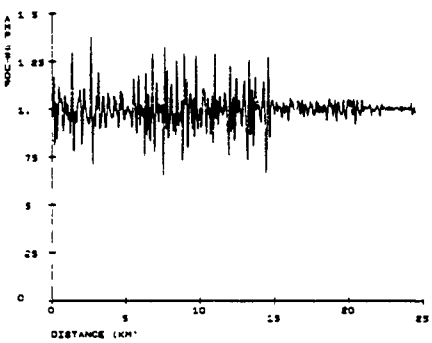
Fig. 2 The variation of the 15 GHz signal as a function of horizontal distance after propagating through the model ionosphere and received at a distance of 700 km from the top of the ionosphere. (a) The amplitude, (b) The phase.



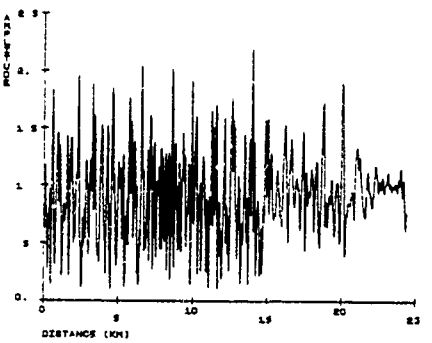
(a)



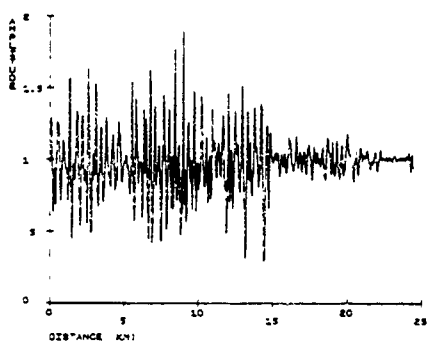
(d)



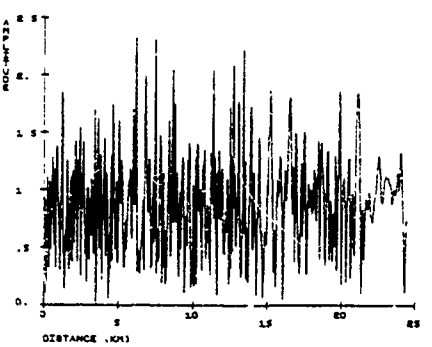
(b)



(e)



(c)



(f)

Fig. 3 Amplitude fluctuations as a function of horizontal distance for six radio frequencies. (a) 15 GHz, (b) 7.5 GHz, (c) 4 GHz, (d) 2.5 GHz, (e) 1.5 GHz, and (f) 800 MHz.

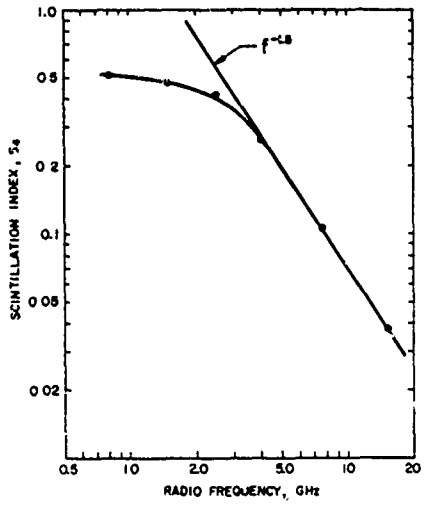
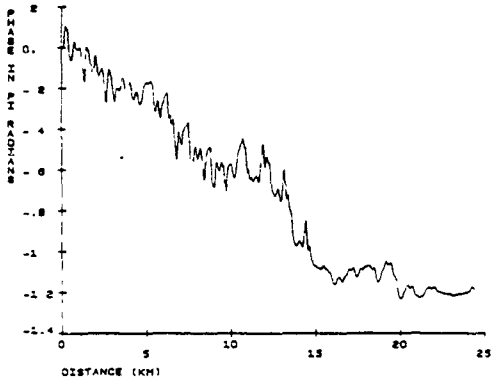
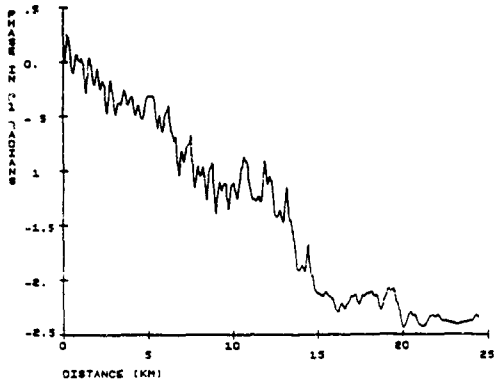


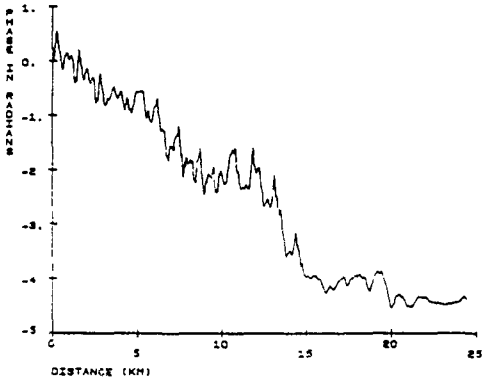
Fig. 4 Scintillation index S_4 as a function of radio frequency. The high frequency asymptotic line has the dependence $f^{-1.5}$.



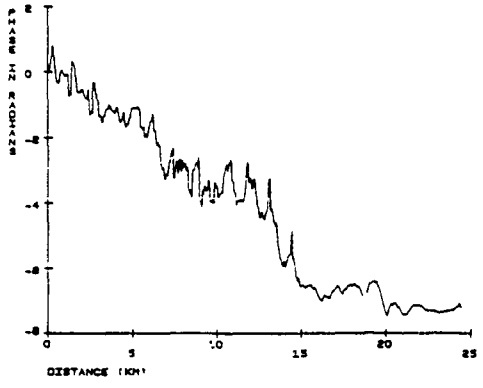
(a)



(b)



(c)



(d)

Fig. 5 Phase fluctuations in π radians as a function of horizontal distance for four radio frequencies. (a) 15 GHz, (b) 7.5 GHz, (c) 4 GHz, and (d) 2.5 GHz.

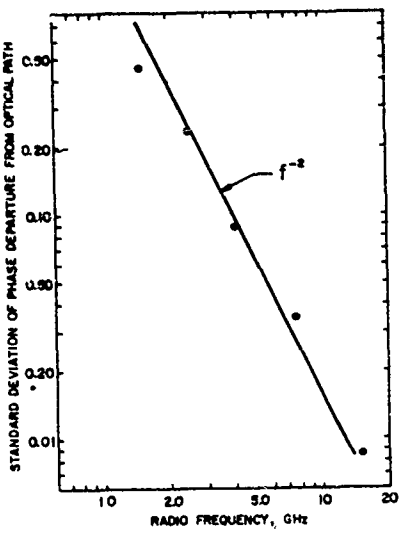
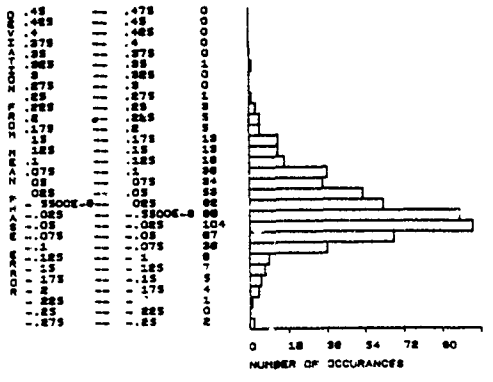
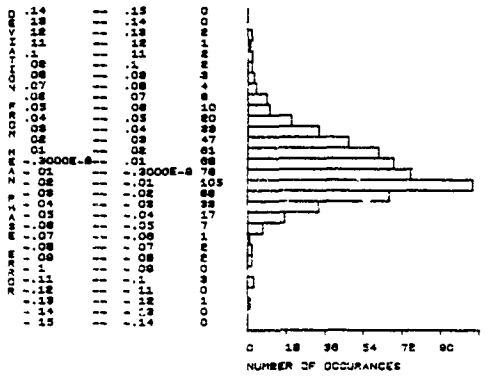


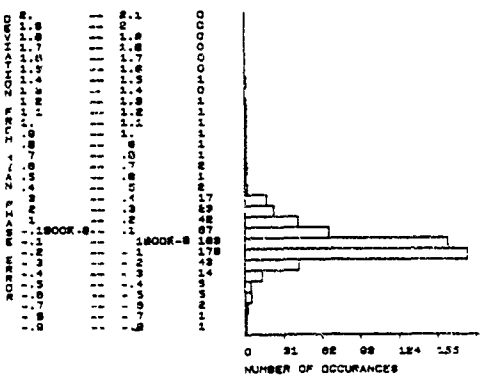
Fig. 6 The standard deviation of phase departure from the optical path, in π radians, as a function of radio frequency.



(a)



(b)



(c)

Fig. 7 Histograms showing the occurrence frequency of phase error in π radians. (a) 7.5 GHz, (b) 4.0 GHz, and (c) 2.5 GHz.

High-Latitude Ionospheric Model: First Step Towards a Predictive Capability

R.W. Schunk, W.J. Raitt and J.J. Sojka
Center for Atmospheric and Space Sciences
Utah State University, Logan, Utah 84322

ABSTRACT

We combined a simple plasma convection model with an ionospheric-atmospheric composition model in order to simulate high-latitude ionospheric behavior. The convection model includes the offset between the geographic and geomagnetic poles, the tendency of plasma to corotate about the geographic pole, and a dawn/dusk magnetospheric electric field mapped to a circular region in the ionosphere about a center offset by a few degrees in the antisunward direction from the magnetic pole. The ionospheric-atmospheric composition model takes account of plasma convection, plasma diffusion, photochemical processes, thermospheric winds, and ion production due to both auroral precipitation and resonantly scattered solar radiation. A typical numerical simulation produces time-dependent, 3-dimensional, ion density distributions (NO^+, O_2^+, N_2^+, O^+, N^+ and He^+) for the high-latitude ionosphere above 42°N magnetic latitude and at altitudes between 160-800 km. One of the early results to emerge from the use of this numerical model was that high-latitude features, such as the 'main trough', the 'ionization hole', the 'tongue of ionization', the 'aurorally produced ionization peaks', and the 'universal time effects', are a natural consequence of the competition between the various chemical and transport processes known to be operating in the high-latitude ionosphere. In this investigation the numerical model was used to study the variations of the electron density with altitude, latitude, longitude, and universal time for ionospheric conditions corresponding to winter solstice and for convection electric field patterns that are representative of both low and high geomagnetic activity conditions. In addition, we studied the morphology of h_mF_2, N_mF_2, the F-region total electron density content, and the topside plasma density scale height.

1. INTRODUCTION

Over the last several years we developed a theoretical model of the convecting high-latitude ionosphere in order to determine the extent to which various chemical and transport processes affect the ion composition and electron density in both the sunlit and dark hemispheres (Banks et al, 1974; Schunk and Banks, 1975; Schunk et al, 1975, 1976; Schunk and Raitt, 1980). From these studies, as well as those by Knudsen (1974), Knudsen et al (1977), Spiro et al (1978), Watkins (1978) and Brinton et al (1978), it is apparent that the high-latitude ionosphere can, by varying degrees, be influenced by solar EUV radiation, energetic particle precipitation, diffusion, thermospheric winds, electrodynamic drifts, polar wind escape, energy-dependent chemical reactions and magnetic storm induced neutral composition changes.

In a recent study (Schunk and Raitt, 1980), we improved our high-latitude ionospheric-atmospheric model so that we could study the solar cycle, seasonal, and geomagnetic activity variations of the daytime high-latitude F-layer. The improvements included updating the various chemical reaction rates; the adoption of the latest solar EUV fluxes measured by the Atmosphere Explorer satellites; the adoption of the most recent MSIS neutral atmosphere model (N_2, O_2, O, He); and the adoption of the Engebretson et al (1977) atomic nitrogen model. Also, our theoretical model was improved by including N^+ and He^+ in addition to the ions NO^+, O_2^+, N_2^+ and O^+.

In parallel with the improvement of our high-latitude ionospheric-atmospheric model, we developed a simple model of high-latitude plasma convection (Sojka et al, 1979a, b; 1980a, b). Our convection model includes the offset between the geographic and geomagnetic poles; the tendency of plasma to corotate about the geographic pole; and a dawn/dusk magnetospheric electric field mapped to a circular region in the ionosphere about a center which is offset in the antisunward direction from the magnetic pole. The convection model allows for asymmetries in the mapped magnetospheric electric field.

We have combined our convection and ionospheric-atmospheric models in order to study the high-latitude F-region. Our model produces time-dependent, 3-dimensional, ion density distributions for the ions NO^+, O_2^+, N_2^+, O^+, N^+ and He^+. We cover the high-latitude ionosphere above 42°N magnetic latitude and at altitudes between about 160-800 km for one complete day. The input parameters used by our model are shown schematically in Figure 1. The neutral composition and thermospheric wind are inputs to our model as are the neutral, ion, and electron temperatures; these parameters vary over the polar cap. The EUV solar radiation spectrum and ion production due to both auroral particle precipitation and resonantly scattered solar radiation are also inputs. Finally, our model requires the cross-tail magnetospheric potential, which is mapped down along dipole geomagnetic field lines. These parameters are used in the calculation of plasma convection, plasma diffusion, and photochemical processes, which in turn yield the ion density distributions.

As the ionospheric plasma convects through the high-latitude neutral atmosphere it is subjected to different chemical and transport processes at different times. Also, at a given time the plasma can be influenced by different physical processes at different altitudes. Figure 2 summarizes how the different processes affect the shape and magnitude of the altitude distribution of electron density. The relative importance of a given process depends on many geophysical phenomena, such as solar cycle, season, geomagnetic activity, longitude, and universal time.

In this study we compared plasma density features in the winter high-latitude F-region for low and high geomagnetic activity conditions. For low geomagnetic activity we adopted a cross-tail magnetospheric electric potential of 20 kV and set A_p=20, while for high geomagnetic activity, the corresponding parameters were 90 kV and 35, respectively. In the ionospheric F-region, a 20 kV cross-tail potential produces plasma convection speeds up to 0.5 km/s, whereas a 90 kV cross-tail potential can produce convection speeds as high as 2 km/s.

2. PLASMA CONVECTION MODEL

Our convection model for low geomagnetic activity included the offset between the geographic and geomagnetic poles, the tendency of plasma to corotate about the geographic pole, and a dawn/dusk magnetospheric electric field mapped to a circular region in the ionosphere about a center offset by 5° in the antisunward direction from the magnetic pole. The radius of the circle corresponded to 17° of latitude and the electric potentials were aligned parallel to the noon/midnight meridian within the circular region. Equatorward of the circle the potential diminished radially and varied inversely as the fourth power of sine magnetic co-latitude (Volland, 1975). Our convection model for high geomagnetic activity was similar, but the circle radius was 18.5° and the center offset was 7.5°. In addition, an enhanced dusk cell convection pattern was adopted. Such an asymmetry in the convection electric field is commonly found in satellite electric field observations (Heppner, 1977) and in incoherent backscatter radar observations (J.C. Foster, private communication). For both low and high geomagnetic activity, we assumed that the cross-tail magnetospheric potential was constant for the duration of our calculations.

Figure 3a shows 10 representative convection trajectories in the magnetic quasi-inertial frame for our low geomagnetic activity case. Magnetic local time is shown by tick marks at 10° intervals around the outer circle, which represents a magnetic latitude of 50°. In addition, the key MLT values are shown. The inner circles represent magnetic latitudes at 10° intervals from 60° to 80°. Field tubes of plasma following different trajectories have different circulation times, owing to the different trajectory lengths and the variation of speed around a given trajectory. Table 1 shows circulation times for the 10 representative trajectories.

The widely different convection times together with the time dependence of the terminator position in the magnetic frame are the source of the UT dependence that our 3-dimensional composition model will display. Figure 3b shows how the terminator, auroral oval, and plasma convection trajectory are coupled. A trajectory that passes near the stagnation region was selected as a test trajectory. The terminator at one particular UT is shown; in one day the terminator moves a total of 22° as the magnetic pole rotates about the geographic pole. As is shown in Figure 3b, most of the polar cap is in darkness in winter. The auroral oval, however, is a source of ionization in the dark hemisphere. For our calculations we selected the quiet time oval of Feldstein and Starkov [1967]. For the UT shown the test trajectory lies in the dark hemisphere; however, it does enter the evening sector oval. This test trajectory was used to study the sensitivity of the F-region ionization to some of our adopted parameters and the results are discussed by Sojka et al (1981a).

In Figure 4a, we present 10 representative plasma drift trajectories in the magnetic quasi-inertial frame for the asymmetric magnetospheric electric field pattern adopted for our high geomagnetic activity case. Magnetic local time is shown as tick marks at hourly intervals on a circle representing 40° magnetic latitude. Magnetic latitude increases linearly to the magnetic pole, as shown by the vertical scale. The circulation times associated with these trajectories are shown in Table 1. The comparison of these circulation times with those obtained for the low geomagnetic

activity condition indicates that in both cases there is a wide range of circulation times, varying from a small fraction of a day to about 1-4 days. Also, in general the circulation times are much shorter for the high geomagnetic activity case.

Also shown in Figure 4a is our adopted auroral oval, which corresponds to the K_p=5 auroral oval of Comfort (1972). When field tubes of plasma enter this region they are subjected to an ion production source due to energetic electron precipitation. The ion production rate profile that we adopted was taken from Knudsen et al (1977) and corresponds to moderately active auroral electron precipitation.

For high geomagnetic activity, the large convection velocities result in an enhancement in the relative importance of certain transport processes, such as ion-neutral frictional heating and induced vertical plasma drifts. The distribution of horizontal drift speeds, which has a direct bearing on frictional heating, is shown in Figure 4b, where contours of the horizontal plasma convection speed are plotted in the magnetic quasi-inertial frame. Each contour is labeled with its appropriate speed in m/s; the region with speeds below 100 m/s is indicated by the shading. A region of high speed, reaching almost 2 km/s, is located in the dusk sector and it corresponds to the enhanced magnetospheric electric field on the dusk side of the polar ionosphere. Over the polar cap the horizontal speed lies in the 200 to 600 m/s range. In contrast, an extended low-speed region is present in the afternoon and evening sectors. The location of this extended low-speed region has a direct bearing on the location of the main or mid-latitude plasma density trough.

Figure 4c shows contours of the vertical component of the plasma convection velocity in the magnetic quasi-inertial frame. Each contour is labeled with the appropriate velocity in m/s; the dashed contours represent downward velocities, while the solid contours represent upward velocities. Upward electrodynamic drifts occur on the dayside where the plasma is convecting toward the magnetic pole, while downward electrodynamic drifts occur in the nightside where the plasma is convecting away from the magnetic pole. For the case considered, the vertical plasma drift ranges from +50 m/s to -70 m/s. Vertical drifts near the extremes of this range have a pronounced effect on both the F-region peak electron density, N_mF_2, and the altitude of the peak, h_mF_2 (cf. Schunk et al, 1976).

In addition to the vertical motion associated with the $\underline{E} \times \underline{B}$ plasma drift, a thermospheric wind also induces a vertical plasma drift (cf. Rishbeth and Garriot, 1969; Watkins and Richards, 1979). The importance of this additional vertical drift to the ion composition and to the maintenance of the nighttime high-latitude ionosphere was discussed by Sojka et al (1981b). In that study we adopted a modified form of the thermospheric wind pattern given by Murphy et al (1976). The Murphy et al (1976) neutral wind is meridional everywhere, flowing towards the pole in the sunlit hemisphere and away from the pole in the dark hemisphere. The meridional flow reaches a maximum equatorward speed of 200 m/s at 0100 local time (LT) and a maximum poleward speed of 60 m/s at about 1300 LT. Near dawn and dusk the wind speed is extremely small. The small modification we made to the Murphy et al (1976) wind model was that we set the low-speed poleward wind to zero. This simplification was motivated by the fact that thermospheric heating in the auroral oval acts to reduce or reverse this poleward wind. However, the equatorward wind in the nightside is enhanced by auroral oval heating, and therefore, this equatorward wind must be included in order to obtain reliable predictions of ionospheric behavior on the nightside.

Figure 4d shows contours of the wind induced upward ionization drift in the magnetic quasi-inertial frame. The contours, which are labeled in m/s, display symmetry about an axis that extends from 1300 to 0100 MLT. Also, the wind induced upward ionization drift increases continuously with decreasing latitude over our latitude range, owing to its dependence on the magnetic field dip angle. Since the thermospheric wind is meridional in the geographic frame, it is UT dependent in the magnetic quasi-inertial frame of Figure 4d. However, the UT variation is relatively small because the induced upward drift depends not only on the thermospheric wind, but also on the magnetic field dip angle, which is UT independent in the magnetic frame.

In our model, the wind induced upward ionization drift is combined with the vertical component of the electrodynamic drift. A comparison of Figures 4c and 4d indicates that the wind induced and electrodynamic drifts oppose each other, producing a net downward drift in the polar cap and a net upward drift at mid-latitudes.

3. HIGH-LATITUDE IONOSPHERIC MORPHOLOGY

The plasma convection patterns for low and high geomagnetic activity that were shown earlier have been used in conjunction with our ionospheric-atmospheric composition model in order to study the variation of the electron density with altitude, latitude, longitude, and universal time. In the following paragraphs, the resultant variation of the electron density with these parameters will be presented and a comparison will be made between the high and low geomagnetic activity cases.

In Figure 5 the O^+ density at 300 km for the low geomagnetic activity case is shown as a gray-scaled contour plot in a magnetic local time (MLT), magnetic latitude polar diagram. The MLT is shown as tick marks at 1 hour intervals, and magnetic latitude varies linearly from the 55^o circle shown to the magnetic pole at the center of each plot. In this figure the gray scale range was chosen to emphasize low density regions. The contour plots in Figure 5 clearly show a marked UT variation of the O^+ density distribution over the region poleward of 55^o latitude. These contour plots show the gross features of a mid-latitude (or main) ionization trough; a region of enhanced ionization in the vicinity of the auroral oval; and a high-latitude ionization hole around local dawn. However, the detailed characteristics of these features differ for the four UT times shown. For example, the depth and extent of the mid-latitude ionization trough differ considerably.

The O^+ density contours shown in Figure 5 can be compared with the Atmosphere Explorer (AE) satellite measurements of Brinton et al (1978), who showed the variation of the minimum and maximum O^+ density at 300 km above the high-latitude region. Although these data correspond to roughly the same geophysical conditions that were adopted for our calculations, it is not possible to do a detailed comparison because the experimental data presented are incomplete and because no account was taken of the UT dependence of the high-latitude ionosphere when the data were plotted. Also, from the theoretical point of view, our model is uncertain with regard to the depth of the main trough because the depth is sensitive to nocturnal maintenance processes, which are not quantitatively well known. Nevertheless, the comparison of our model predictions with the gross features of the AE data produces some useful results. First, both the data and the model predictions display ionospheric features such as the main trough, the polar hole, and the enhanced ionization in the auroral oval. However, our O^+ densities in the oval are generally larger than the measured values, indicating that our auroral production rates are probably too large. The location of the polar hole corresponds well with that observed; its size is strongly UT dependent and at about 0500 UT it extends into the noon polar cap sector. These latter two features are not present in the AE data owing to both the lack of UT data selection and a data gap over the polar cap. With regard to the depth of the polar hole, our model calculations produced O^+ densities as low as $2 \times 10^2 cm^{-3}$, which are in good agreement with the observed densities, but our calculations also indicate a significant UT variation. As a final item we note that the maximum and minimum O^+ density plots display main troughs that are different in extent and form. Our model predicts a UT variation of the extent and form of the main trough, but it is not clear whether or not the variation presented by Brinton et al (1978) is related to a UT effect, again owing to the omission of UT as a data selection criterion.

Figure 6 shows contours of the electron density at 300 km for the high geomagnetic activity case and for four universal times. The four UT's selected correspond to the geographic pole being near the dawn (1 UT), noon (7 UT), dusk (13 UT), and midnight (19 UT) magnetic meridians. In Figure 6, each plot is a polar diagram in MLT and magnetic latitude. The electron density variation is presented as gray scaled contour intervals, with the darkest shades corresponding to the lowest densities. Over the polar region, the electron density varies from about 10^4 to 10^6 cm^{-3}. A particularly evident feature seen in Figure 6 is the marked UT variation of the electron density. However, certain ionospheric features are also evident, such as the main trough, the auroral oval, a polar hole, and a tongue of ionization.

Two high density regions are present at all four UT's, one is related to solar EUV production in the sunlit hemisphere and the other is the nighttime auroral oval - polar cap region. Both of these regions show UT dependent changes. The dayside high density region moves toward the polar cap and then recedes as the terminator moves in the magnetic quasi-inertial frame. The auroral oval - polar cap region has a UT independent equatorward edge which is associated with the equatorward edge of the auroral oval. However, inside this region the electron density varies markedly with UT. For example, at midnight and 80^o latitude, the electron density varies by more than an order of magnitude.

The two high density regions discussed above are, at times (1, 13, and 19 UT), linked by a tongue of high density in the dawn-noon sector. This tongue of ionization results from the poleward convection of dayside, high density field tubes (see Figure 4a). The preference for a pre-noon tongue over a post-noon tongue is a consequence of the higher convection speeds in this sector (see Figure 4b). With higher convection speeds the field tubes of plasma can convect further during a 'typical 'decay' time. In contrast, the slower convection speeds that occur for low geomagnetic activity do not lead to the formation of a pronounced tongue of ionization (see Figure 5).

As discussed above, the polar hole is a region where the plasma density at 300 km falls below 10^3 cm^{-3} during quiet geomagnetic activity conditions. This polar hole region is located in the nightside polar cap adjacent to the poleward edge of the auroral oval in the midnight-dawn MLT sector. In a recent model study (Sojka et al, 1981a), we have shown that the Brinton et al (1978) polar hole results from slow antisunward convection across the dark polar cap in combination with ordinary ionic recombination. In that model study, the plasma at 300 km had sufficient time to decay to a value as low as $2 \times 10^2 cm^{-3}$, which is the level maintained by resonantly scattered solar radiation.

It is evident from Figure 6 that the Brinton et al (1978) polar hole does not form during more active geomagnetic conditions, owing to the much shorter transit times across the polar cap. The electron density in the location of the Brinton et al (1978) polar hole remains above ~ x 10^4 cm^{-3} at all UT's. However, there is another region in the polar cap at 7 UT that contains relatively low electron densities. In contrast to the low geomagnetic activity polar hole, which is due to chemical loss processes acting over a long time period, the high geomagnetic activity polar hole shown in Figure 6 results from the large downward transport velocities that exist in this region (see Figures 4c and 4d).

Another ionospheric feature that is known to exist in the winter polar ionosphere is the 'mid-latitude' or 'main' electron density trough. This trough is a region of low electron density situated just equatorward of the nocturnal auroral oval. It is generally apparent at all universal times (see Figure 6), but its depth and local time extent are strongly UT dependent. At 1 UT, it is deepest in the morning sector; at 13 UT, it is deeper in the evening sector than in the morning sector; while in the middle of this UT range the trough tends to be more symmetrical about midnight. The trough in the afternoon sector at 7 UT extends to 15 MLT at a magnetic latitude of about 68^o, while at 1 UT it extends to only 19 MLT at a latitude of 57^o. This difference of 4 hours of MLT and 11^o of latitude highlights the marked UT variation of the trough location.

A comparison of Figures 5 and 6 indicates that for both low and high geomagnetic activity the main electron density trough covers a wide range of MLT. In both cases, the location of the region of lowest electron density exhibits a significant variation with UT. The major difference between the main troughs for the high and low geomagnetic activity cases is that in general the main trough tends to be much deeper for low geomagnetic activity.

4. REDUCED IONOSPHERIC PARAMETERS

A number of ionospheric parameters have been frequently used to describe the F-region. In this section we will discuss four such parameters, namely: h_mF_2, N_mF_2, the F-region column content, and the topside plasma density scale height. In general, these parameters show the greatest variability for the high geomagnetic activity case, and therefore, we will restrict our discussion to this case.

Figure 7 shows contours of the altitude of the F-region peak density in the magnetic quasi-inertial frame for 10 UT. The contours are labeled in km. Although only one UT is shown, the UT variation of h_mF_2 is small, being on the order of 10 km at a given location. In Figure 7, the altitude of the peak density varies from 250 to 380 km, with there being two distinct regions where h_mF_2 is large. One region, near the noon meridian at 80^o, is associated with a large upward electrodynamic drift, while the other region, the nightside mid-latitude ionosphere, is associated with the upward plasma drift induced by the thermospheric wind. The minimum values of h_mF_2, on the other hand, are found near the midnight meridian at 70^o. This location corresponds to the region where the maximum downward electrodynamic drifts occur. In this regard, it should be noted that the h_mF_2 contours display features that are similar to those found in the contours of the vertical velocities associated with the electrodynamic and wind induced drifts (see Figures 4c and 4d). The close similarity between these sets of contours emphasizes the importance of vertical transport in determining the altitude of the F-region peak electron density. The lack of a significant UT variation in h_mF_2 is a further indication of this relationship, since in the magnetic quasi-inertial frame the vertical velocities are almost UT independent.

Figure 8 shows contours of the logarithm of N_mF_2 (upper panel) and the logarithm of the F-region column density (lower panel) at 10 UT. The column density was computed as the density integral from 160 to 800 km. The most evident feature to emerge from a comparison of the two panels is that the two contour patterns are very similar. Although both N_mF_2 and the column content exhibit a significant UT variation, this similarity is present at all universal times. The obvious inference from the similarity of N_mF_2 and the column content is that the bulk of the column content is contained in a region near the F-region peak. With regard to the variation of N_mF_2 with UT, it follows the general pattern of change shown in the four UT plots given in Figure 6 for an altitude of 300 km.

Figure 9 shows contours of the topside plasma density scale height in the magnetic quasi-inertial frame for 10 UT. The scale heights were calculated in the altitude range from 600 to 800 km. Over the polar region the scale height varies significantly, from 190 to 330 km. A comparison of Figure 9 with Figure 4b indicates that the large scale heights are associated with the region of high convection speeds. Strong convection leads to an enhanced ion-neutral frictional heating rate, which in turn leads to elevated ion temperatures and increased plasma density scale heights. As far as the UT variation of the scale height is concerned, the contour patterns obtained at other UT's are very similar to that shown in Figure 9.

5. SUMMARY

We combined a simple plasma convection model with an ionospheric-atmospheric composition model in order to study the plasma density features associated with both weak and strong convection in the winter high-latitude F-region. The details of our time-dependent ionospheric-atmospheric composition model, including ion chemistry and transport equations, are given by Schunk and Raitt (1980), while the details concerning the combined plasma convection and ionospheric- atmospheric models are given by Sojka et al (1981a). In a model calculation a field tube of plasma is followed as it moves along a convection trajectory through a moving neutral atmosphere. Altitude profiles of the ion densities are obtained by solving the appropriate continuity, momentum, and energy equations including numerous high-latitude processes. The result of following many field tubes of plasma is a time-dependent, 3-dimensional ion density distribution for each of the ions NO^+, O_2^+, N_2^+, O^+, N^+, and He^+. The high-latitude ionosphere above 42^oN magnetic latitude and at altitudes between 160 and 800 km was covered for one complete day.

From our study dealing with low geomagnetic activity we found the following:

1. High-latitude ionospheric features, such as the main trough, the ionization hole, the tongue of ionization, the aurorally produced ionization peaks, and the universal time effects, are a natural consequence of the competition between the various chemical and dynamical processes known to be operating in the high-latitude ionosphere.
2. The high-latitude F-layer ionization exhibits a significant UT variation owing to the displacement between the geomagnetic and geographic poles. The peak electron density at a given location and local time can vary by an order of magnitude due to this UT effect.
3. The depth and longitudinal extent of the main trough exhibit a significant UT dependence.
4. The spatial extent, depth, and location of the polar ionization hole are UT dependent.
5. Ion production due to resonantly scattered radiation can maintain the electron density at 300 km in the polar ionization hole at about 2×10^2 cm^{-3}, in agreement with the measurements of Brinton et al (1978).

From our study dealing with high geomagnetic activity we found the following:

1. For a cross-tail magnetospheric electric potential of 90 kV, the horizontal component of the plasma convection velocity in the ionosphere approaches 2 km/s and the vertical component ranges from +50 m/s to -70 m/s. Vertical drifts near the extremes of this range have a pronounced effect on both N_mF_2 and h_mF_2.
2. As a field tube of plasma traverses a given trajectory it is subjected to different chemical and transport processes at different times. Also, at a given time the plasma in the field tube can be influenced by different physical processes at different altitudes.
3. As was found in our previous study dealing with weak convection (Sojka et al, 1981a), for strong convection the electron density exhibits a significant variation with altitude, latitude, longitude, and universal time.
4. For strong convection, certain ionospheric features, such as the main trough, the aurorally produced ionization peaks, the polar hole, and the tongue of ionization, are evident but they are modified in comparison with those found for weak convection.
5. For strong convection, the tongue of ionization that extends across the polar cap is much more pronounced than for weak convection. This tongue of ionization is highly UT dependent.
6. The polar hole found by Brinton et al (1978) during quiet geomagnetic activity conditions, which results from slow antisunward convection across the dark polar cap in combination with ordinary ionic recombination, does not form for more active geomagnetic conditions, owing to the much shorter transit times across the polar cap.
7. For strong convection, a new polar hole appears in the polar cap at certain universal times. This new polar hole is associated with large, downward, electrodynamic plasma drifts.
8. For strong convection, the main or mid-latitude electron density trough is generally not as deep as that found for quiet geomagnetic activity conditions. However, in both cases the depth and extent of the trough are strongly UT dependent. At certain UT's the trough is deeper in the morning sector than in the evening sector, while at other UT's the reverse is true. At still other times, the trough tends to be more symmetrical about midnight.
9. The altitude of the peak electron density exhibits a significant variation with magnetic latitude and MLT, owing to the vertical plasma drifts induced by the electrodynamic force and the thermospheric wind. The altitude of the peak varies from 250 to 380 km in the high-latitude ionosphere at a given UT.
10. The variations of N_mF_2 and the F-region column density with magnetic latitude and MLT are very similar. Although both of these quantities exhibit a significant UT variation, this similarity is present at all universal times.
11. At a given UT, the topside plasma density scale height varies significantly over the polar region, from 190 to 330 km. The large scale heights are associated with the regions of strong convection.

ACKNOWLEDGEMENTS This research was supported by Air Force contract USAF/ESD F19628-79-C-0025, NASA grant NAGW-77 and NSF grant ATM80-15497 to Utah State University.

REFERENCES

Banks, P.M., R.W. Schunk and W.J. Raitt, NO^+ and O^+ in the high latitude F-region, Geophys. Res. Lett., 1, 239-242,1974.

Brinton, H.C., J.M. Grebowsky, and L.H. Brace, The high-latitude winter F region at 300 km: Thermal plasma observations from AE-C, J. Geophys. Res., 83, 4767-4776, 1978.

Comfort, R.N., Auroral Oval Kinematics Program, NASA Report, CR-61373, 1972.

Engebretson, M.J., K. Mauersberger, D.C. Kayser, W.E. Potter, and A.O. Nier, Empirical model of atomic nitrogen in the upper thermosphere, J. Geophys. Res., 82, 461-471, 1977.

Feldstein, Y.I., and G.V. Starkov, Dynamics of auroral belt and polar geomagnetic disturbances, Planet. Space Sci., 15, 209-229,1967.

Heppner, J.P., Empirical models of high-latitude electric fields, J. Geophys. Res., 82, 1115-1125, 1977.

Knudsen, W.C., Magnetospheric convection and the high-latitude F2 ionosphere, J. Geophys. Res., 79, 1046-1055, 1974.

Knudsen, W.C., P.M. Banks, J.D. Winningham, and D.M. Klumpar, Numerical model of the convecting F_2 ionosphere at high latitudes, J. Geophys. Res., 82, 4784, 1977.

Murphy, J.A., G.J. Bailey and R.J. Moffett, Calculated daily variations of O^+ and H^+ at mid-latitudes, J. Atmos. Terr. Phys., 38, 351-364, 1976.

Raitt, W.J., R.W. Schunk and J.J. Sojka, Modelling the high latitude ionosphere, Proceedings of the AGARD/NATO Symposium on "The Physical Basis of the Ionosphere in the Solar-Terrestrial System", Naples, Italy, 1980.

Rishbeth, H. and O.K. Garriott, Introduction to Ionospheric Physics, Academic Press, 1969.

Schunk, R.W. and P.M. Banks, Auroral N_2 vibrational excitation and the electron density trough, Geophys. Res. Lett., 2, 239-242, 1975.

Schunk, R.W. and W.J. Raitt, Atomic nitrogen and oxygen ions in the daytime high-latitude F-region, J. Geophys. Res., 85, 1255-1272, 1980.

Schunk, R.W., W.J. Raitt and P.M. Banks, Effect of electric fields on the daytime high-latitude E- and F-regions, J. Geophys. Res., 80, 3121-3130, 1975.

Schunk, R.W., P.M. Banks and W.J. Raitt, Effects of electric fields and other processes upon the nighttime high latitude F-layer, J. Geophys. Res., 81, 3271-3282, 1976.

Sojka, J.J., W.J. Raitt and R.W. Schunk, Effect of displaced geomagnetic and geographic poles on high-latitude plasma convection and ionospheric depletions, J. Geophys. Res., 84, 5943-5951, 1979a.

Sojka, J.J., W.J. Raitt and R.W. Schunk, High latitude plasma convection: Predictions for EISCAT and Sondre Stromfjord, Geophys. Res. Lett., 6, 877-880, 1979b.

Sojka, J.J., J.C. Foster, W.J. Raitt, R.W. Schunk and J.R. Doupnik, High-latitude convection: Comparison of a simple model with incoherent scatter observations, J. Geophys. Res., 85, 703-709, 1980a.

Sojka, J.J., W.J. Raitt and R.W. Schunk, A comparison of model predictions for plasma convection in the northern and southern polar regions, J. Geophys. Res., 85, 1762-1768, 1980b.

Sojka, J.J., W.J. Raitt and R.W. Schunk, A theoretical study of the high-latitude winter F-region at solar minimum for low magnetic activity, J. Geophys. Res., in press, 1981a.

Sojka, J.J., W.J. Raitt and R.W. Schunk, Theoretical predictions for ion composition in the high-latitude winter F-region for solar minimum and low magnetic activity, J. Geophys. Res., in press, 1981b.

Sojka, J.J., W.J. Raitt and R.W. Schunk, Plasma density features associated with strong convection in the winter high-latitude F-region, J. Geophys. Res., submitted, 1981c.

Spiro, R.W., R.A. Heelis, and W.B. Hanson, Ion convection and the formation of the mid-latitude F-region ionization trough, J. Geophys. Res., 83, 4255-4264, 1978.

Volland, H., Models of global electric fields within the magnetosphere, Annl. Geophys., 31. 159, 1975.

Watkins, B.J., A numerical computer investigation of the polar F-region ionosphere, Planet. Space Sci., 26. 559-569, 1978.

Watkins, B.J. and P.G. Richards, A theoretical investigation of the role of neutral winds and particle precipitation in the formation of the auroral F-region, J. Atmos. Terr. Phys., 41, 179-187, 1979.

Table I. Circulation Periods for Plasma Field Tubes.

Trajectory	Low Activity Model* times (day)	High Activity Model+ times (day)
1	1.00	1.06
2	1.02	1.35
3	1.13	1.32
4	1.09	0.79
5	0.97	0.58
6	0.78	0.25
7	0.64	0.15
8	0.58	0.27
9	0.72	0.11
10	0.95	0.06

*Trajectories labeled as they appear in Figure 3a.

+Trajectories labeled as they appear in Figure 4a.

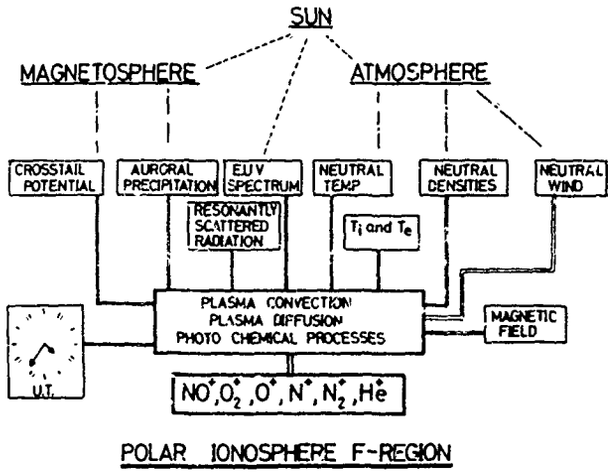


Figure 1. Diagram showing the various physical processes affecting the F-region polar ionosphere. From Raitt et al (1980).

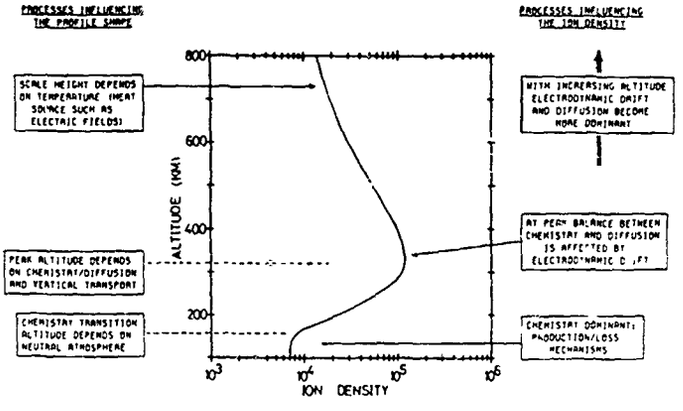


Figure 2. Schematic illustration of a representative altitude distribution of plasma density, showing the major processes that affect its magnitude and shape.

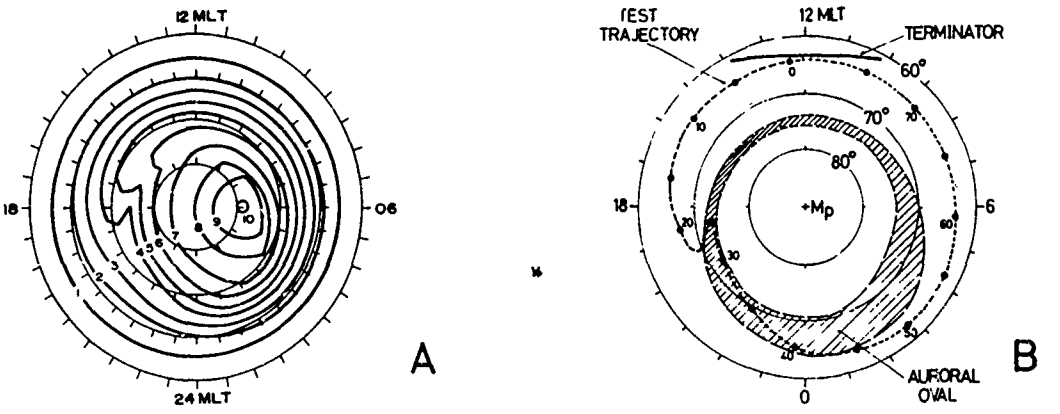


Figure 3. Plasma drift trajectories (panel A) and the location of a northern hemisphere auroral oval (panel B) for low geomagnetic activity. Both panels correspond to the magnetic quasi-inertial frame, with magnetic latitude represented by circles at 10° intervals and magnetic local time by tick marks along the outer circle. From Sojka et al (1981a).

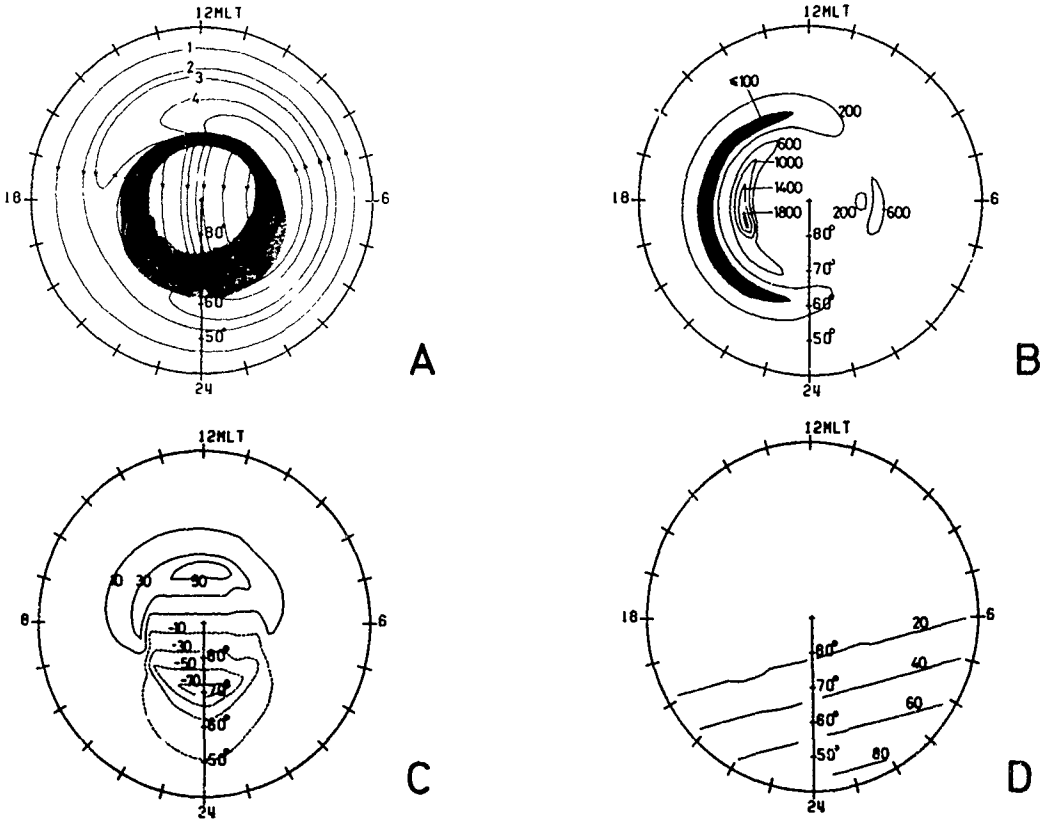


Figure 4. Plasma drift trajectories and the auroral oval (panel A), contours of horizontal plasma convection speeds (panel B), contours of the vertical component of the $\underline{E} \times \underline{B}$ drift velocity (panel C), and contours of the wind induced vertical plasma drift (panel D) for high geomagnetic activity. All four panels correspond to the magnetic quasi-inertial frame. In panels B, C, and D, the contours are labeled in m/s. In panels C and D, upward drifts are positive. From Sojka et al (1981c)

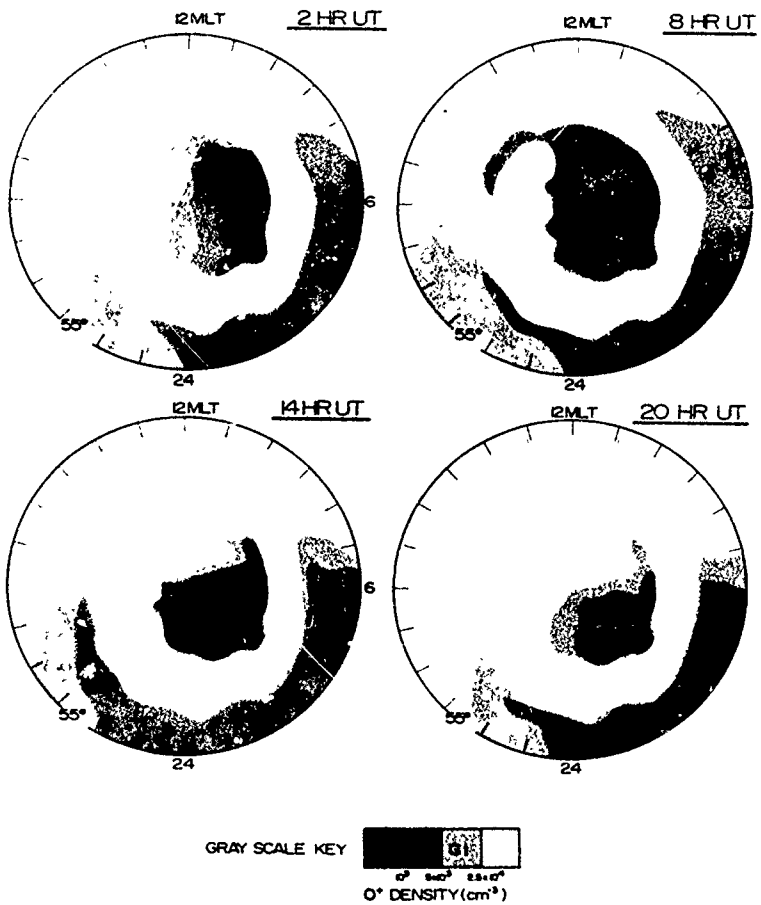


Figure 5. O^+ density contours at 300 km in the magnetic quasi-inertial frame for low geomagnetic activity. From Sojka et al (1981b).

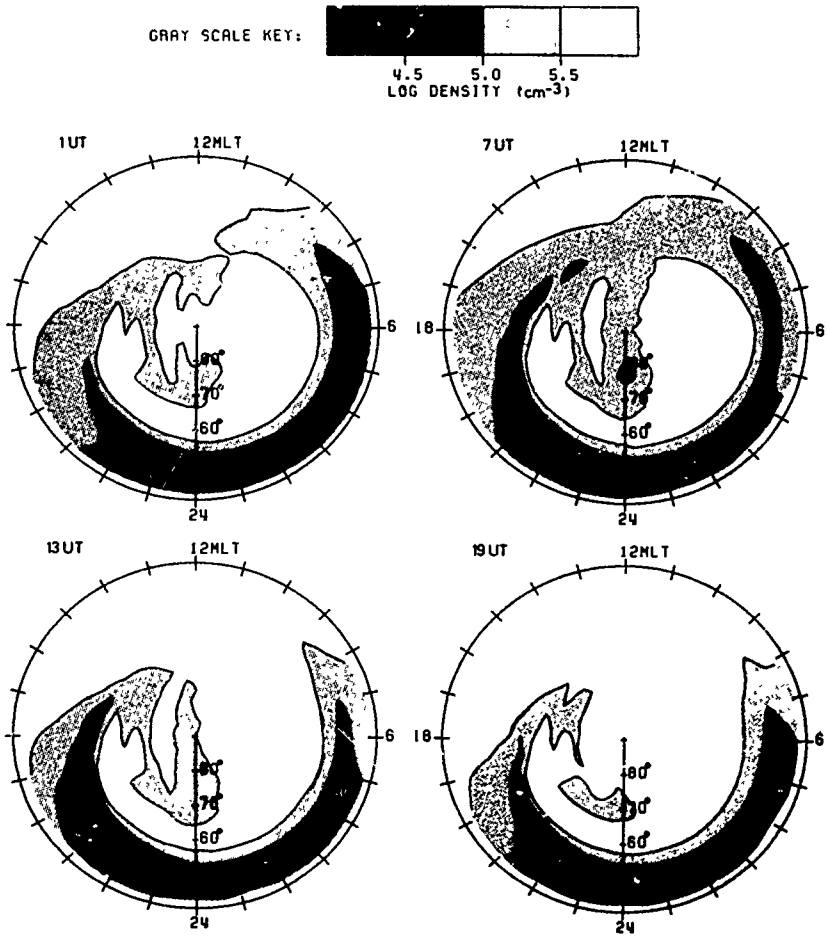


Figure 6. O^+ density contours at 300 km in the magnetic quasi-inertial frame for high geomagnetic activity. From Sojka et al (1981c).

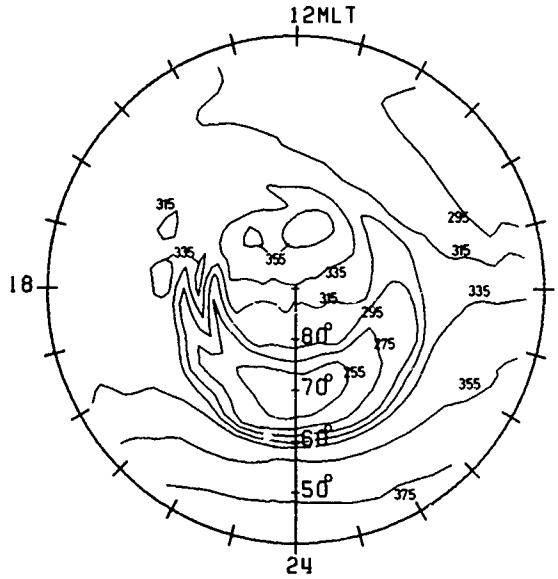
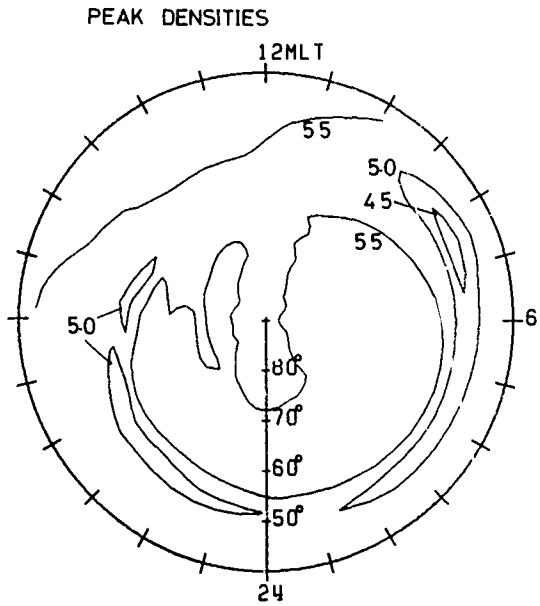


Figure 7. Contours of h_mF_2 (km) in the magnetic quasi-inertial frame for high geomagnetic activity at 10 UT. From Sojka et al (1981c).



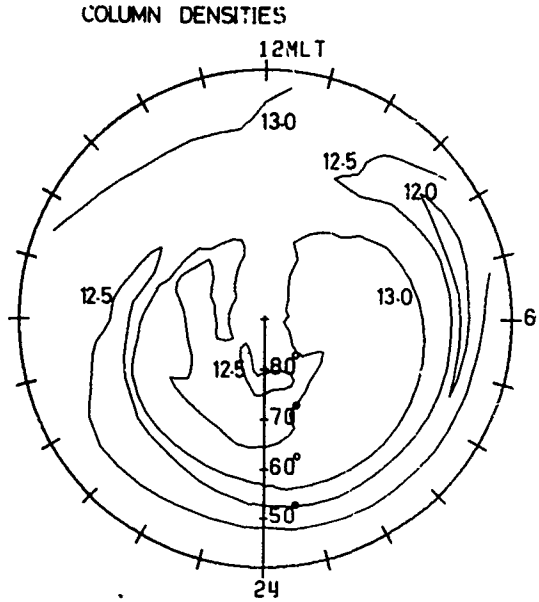


Figure 8. Contours of N_mF_2 (top panel) and column density (bottom panel) in the magnetic quasi-inertial frame for high geomagnetic activity at 10 UT. The contours are labeled in $\log_{10}$ of the appropriate parameter, expressed in cgs units. From Sojka et al (1981c).

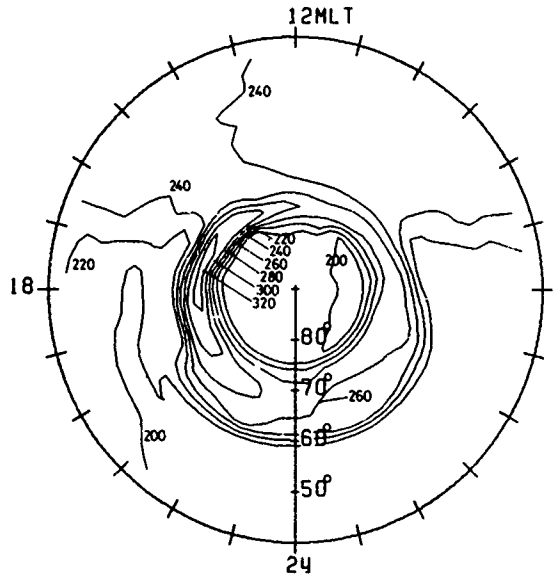


Figure 9. Contours of the plasma density scale height (km) in the magnetic quasi-inertial frame for high geomagnetic activity at 10 UT. From Sojka et al (1981c).

THE APPLICABILITY OF USING THEORETICAL MODELS TO IMPROVE IONOSPHERIC MAPS

D. N. Anderson
NOAA/ERL/SEL
Boulder, Colorado 80303

C. M. Rush, M. PoKempner, and F. G. Stewart
NTIA/ITS
Boulder, Colorado 80303

ABSTRACT

In this investigation we theoretically calculate electron density distributions in the mid-latitude ionospheric F region by solving the time-dependent ion continuity equation numerically, including the effects of production, loss and transport of ionization. The object of the study is to obtain values of the F region critical frequency which can be used to improve the prediction of ionospheric parameters over regions of the earth which are inaccessible to ground based measurements. This is accomplished by including into the theoretical calculations realistic physical processes along with a realistic geomagnetic field model. We adjust input parameters so that agreement is achieved between calculated and observed foF2 values as a function of local time at two different stations and then assume that these same input parameters (i.e., neutral wind field, neutral atmospheric model, etc.) are valid at the intermediate region of interest where the major difference is the geomagnetic field line configuration. Coefficients that yield global representations of foF2 are then determined using the theoretically derived foF2 values and the predicted critical frequencies are compared with observed values to estimate the degree of improvement.

INTRODUCTION

Global maps of ionospheric parameters such as shown in figure 1 form the basis for a number of empirical and statistical models of the ionospheric electron density. The accuracy of the models is tied directly to the accuracy of the maps of the ionospheric parameters. Maps of the critical frequency of the E region, foE (Leftin, 1976), the critical frequency of the F1 region, foF1 (Rosich and Jones, 1973), and the critical frequency of the F2 region, foF2 (Jones et al., 1969; Jones and Obitts, 1970) have been employed in one form or another to determine ionospheric structure (Nisbet, 1971), HF propagation conditions (Barghausen et al., 1969; CCIR, 1978), and transionospheric propagation factors (Bent et al., 1978).

The global maps of ionospheric parameters are generated from numerical coefficients obtained by performing a spherical harmonic analysis on observed monthly median values of foE, foF1, and foF2. The basis for the analysis has been described in detail by Jones and Gallet (1962) and needs not be discussed here. Suffice it to say that the resultant accuracy of the maps is dependent in a significant manner upon the geographical distribution of the data that were used in the generation of the coefficients. The data, foE, foF1, and foF2, that were used in developing the global maps were obtained from between 100 and 150 vertical incidence ionosonde stations that operated throughout the world. These stations provided observations that permit a reasonably accurate map of the ionospheric parameters at those locations for which data were available. At locations for which data were non-existent, such as for ocean areas, the accuracy of the maps is questionable.

It has long been appreciated that the uncertainties in the maps of foF2 are the largest source of potential error in any ionospheric model that utilizes global maps of ionospheric parameters. This is due principally to the fact that the F2 region is the most variable of the ionospheric regions (see Rush and Gibbs, 1973, for example), displaying large changes in both temporal and spatial scales. The variations in the F2 region are the manifestation of complex interactions between neutral and ionized constituents at ionospheric heights, the dependence of F2 region phenomena upon the geomagnetic field, and the influence of the magnetosphere on the ionosphere. The variability of the F2 region renders it difficult to extrapolate observations of foF2 at one location to another location with a degree of accuracy that is commensurate with extrapolation procedures employed for the E and F1 regions.

In order to improve the accuracy of the global maps of foF2 in regions where observations are lacking, it was decided to investigate the possibility of generating values of foF2 from a theoretical model and then re-generating the numerical coefficients that yield the ionospheric maps using both theoretical and actual observations. The theoretical values of foF2 are determined from the time-dependent continuity equation for ions and electrons in the ionosphere. In the following section, the theoretical model used to determine foF2 at locations removed from ground-based ionosondes is described. Also discussed is the procedure adopted in order to assure that the calculated theoretical values of foF2 are consistent with the observed values of foF2 that are available. Emphasis in this study has thus far been directed toward the mid-latitude F2 region. In the third section of this paper, initial results of the study are presented for the southern hemisphere F2 region during September and December 1978. These results are then compared to results obtained from independent data sources (satellites) that describe the mid-latitude F2 region.

THEORETICAL MODEL

To find the ion (O^+) and electron density as a function of altitude, latitude, and local time, the time-dependent ion continuity equation including the effects of production, loss of charge exchange, and transport by diffusion, neutral wind, and $\bar{E} \times \bar{B}$ drift is numerically solved.

The ion continuity equation is given by

$$\partial N_i/\partial t + \nabla \cdot (N\bar{V}_i) = P_i - L_i \tag{1}$$

where N_i is the ion density; P_i, the ion production rate; L_i, the loss rate; and $\bar{V}_i$, the transport velocity. In the ionosphere, plasma is transported along the geomagnetic field lines by diffusion and neutral winds and perpendicular to field lines primarily by $\bar{E} \times \bar{B}$ drift (Kendall and Pickering, 1967). In solving (1), the independent coordinates are transformed to a coordinate system parallel and perpendicular to the magnetic field line (see Anderson, 1973). Equation (1) can then be written

$$\partial N_i/\partial t + \bar{V}_{i\perp} \cdot \nabla N_i = P_i - L_i - \nabla \cdot (N_i \bar{V}_{i||}) - N_i \nabla \cdot \bar{V}_{i\perp} \tag{2}$$

where $\bar{V}_{i\perp}$ is given by $\bar{E} \times \bar{B}/B^2$ and $\bar{V}_{i||}$ includes the effects of plasma diffusion and neutral wind. The left hand side of (2) is the time rate of change of the ion density in a reference frame moving with the $\bar{E} \times \bar{B}$ drift and is such that the plasma simply co-rotates with the earth, $\nabla \cdot \bar{V}_{i\perp} = 0$. The right side of (2) involves terms which are second order in the coordinate parallel to $\bar{B}$. Equation (2) is solved numerically, (see Anderson, 1973 for details) to give N_i ($= N_e$) densities as a function of altitude, latitude, and local time.

The set of coefficients for the ion continuity equation is obtained from models of the neutral composition; neutral temperature; ion and electron temperatures, and production, loss, and diffusion rates as well as the neutral wind model and geomagnetic field model. Briefly, these models are as follows:

1. The MSIS (Hedin et al., 1977) neutral atmosphere model is used to calculate N_2, O_2, and O and the neutral temperature, T_n, as a function of altitude, latitude, and local time. A 10.7 cm solar flux value of 150 is adopted for the September 1978 calculations while 165 is chosen for the December calculations. At all altitudes and local times it is assumed that $T_i = T_e = T_n$.

2. The production rate and loss rate reactions incorporated in the calculation of the O^+ densities are

$O + h\upsilon \rightarrow O^+ + e$	September $P_\infty = 4.0 \times 10^{-7}$ sec $^{-1}$
	December $P_\infty = 4.5 \times 10^{-7}$ sec $^{-1}$
$O^+ + N_2 \xrightarrow{k_1} NO^+ + N$	$k_1 = 7 \times 10^{-13}$ cm^3/sec
$O^+ + O_2 \xrightarrow{k_2} O_2^+ + O$	$k_2 = 2 \times 10^{-11} (T_n/300)^{-1/2}$ cm^3/sec

3. The diffusion coefficient for elastic collisions between ions and neutrals is given by

$$D_{in}^{-1} = \frac{300}{T_n}\left[\frac{[O_2]}{\delta O_2} + \frac{[N_2]}{\delta N_2} + \frac{[O]}{\delta O}\right]$$

where $\delta O_2 = 3.3 \times 10^{18}$ cm^{-1} sec^{-1}, $\delta N_2 = 3.9 \times 10^{18}$ cm^{-1} sec^{-1}, and $\delta O = 3.7 \times 10^{16}$ cm^{-1} sec^{-1}.

4. A realistic geomagnetic field model is used to transform the spherical polar coordinate system r, θ, and φ to one describing directions parallel and perpendicular to B. The "real" field is given by Euler potentials α and β where $\overline{B} = \nabla\alpha \times \nabla\beta$ (Stern, 1965; 1967). A full description of the technique is covered in the paper by Anderson (1973). For this study, eight geomagnetic field lines are chosen, spanning the globe approximately every 45° geographic longitude with the southern end of each field line lying at about 40° south geographic latitude. Table 1 lists the geographic coordinates in the southern hemisphere at the end of each field line and gives the declination angle (positive to the east of north) and dip angle (positive in the northern hemisphere) at each location.

5. The neutral wind models assumed to be appropriate for the September and December 1978 periods were chosen in the following manner. First, the O^+ continuity equation was used to calculate hourly foF2 values at Mundaring, Australia (116.2° E geog. longitude; 32.6° S geog. latitude), where the declination of the field line is 2° W. The meridional component of the neutral wind velocity was adjusted so that agreement between observed and calculated values was obtained. Then a magnetic field line appropriate to Auckland, New Zealand (175° E geog. longitude; 37° S geog. latitude), was chosen where the declination of the field line is 20°E. Incorporating the same meridional component of the neutral wind, the zonal wind component was adjusted so that the calculated hourly foF2 values agreed with the observed values over Auckland. This procedure was carried out for both the September and December periods. These two highly idealized wind systems were then assumed to apply at the eight geographic locations listed in Table 1. Figures 2 and 3 display these wind patterns during September and December, respectively. The wind velocity is assumed to be independent of altitude.

Equation (2) was solved numerically using a Crank-Nicolson (Crank and Nicolson, 1947) implicit finite differencing scheme. The two boundaries are at 125-km altitude at the northern and southern ends of the field line. Boundary conditions for the ion density are photo-chemical equilibrium during the day and 10 ions/cm^3 at night. However, the boundary conditions are not critical, as the appropriate ion densities are reached within a few space steps. The changing size of the magnetic flux tube is taken into account in the calculations.

Table 1. Geographic Coordinates, Dip and Declination at the Southern Hemisphere End of Eight Field Lines

Geographic Longitude (°E)	Geographic Latitude (°S)	Dip Angle	Declination Angle
32.8	40.65	-63.9	-31.3
81.6	40.58	-69.4	-35.4
121.1	35.6	-69.6	- 2.5
170.7	39.7	-65.3	17.96
213.5	40.0	-58.7	18.7
259.2	38.8	-46.7	21.3
305.6	43.4	-40.5	1.3
352.0	38.0	-58.3	-28.3

RESULTS

Results for the September calculations at three different longitudes where the declination of the field line changes but the dip angle is approximately the same are pictured in Figure 4 as curves of foF2 as a function of local time. During daytime hours, the westward declination field line has the greatest peak electron density because the early morning westward blowing zonal wind (80 m/sec) transports ionization upward from below where the production rate is greater to higher altitudes where the loss rate is less. In geographic regions where the declination is nearly zero, the effect of the zonal wind is minimal and the results shown for -2.5° declination are obtained. Conversely, where the declination of the field line is eastward, a westward blowing neutral wind transports ionization downward into a higher loss rate region, decreasing the peak electron density such as shown for 18° declination.

During nighttime hours when the zonal wind blows eastward (1900-2400 LT), the effects of field-line declination on peak electron density are reversed. An eastward blowing wind lowers the F region if the declination is westward, thus decreasing foF2, while the same eastward wind raises the F layer where declination is eastward, causing a slower decay in foF2 with time. This can be seen by comparing the slopes of the curves shown in Figure 4 between the hours 1800 and 2200 local time.

Results for the December calculations are pictured in Figure 5. During the day, the peak electron density is greater for a westward declination field line, while zero and eastward declination field lines have about the same daytime peak electron density of 1×10^6 el/cm^3(9 MHz). At night, the combination of the eastward zonal wind and equatorward meridional wind causes a reversal in the declination effects on foF2 for the same reasons given above.

Using the values of foF2 determined at the eight locations given in Table 1, it is possible to combine these with available observations and values deduced from the existing numerical coefficients to produce maps of foF2. Figure 6 shows a contour map of foF2 for the southern hemisphere for December 1978. The values of foF2 determined by the theoretical calculations are shown encircled on the figure.

In order to gain some estimate of the amount of improvement afforded by incorporating the theoretically deduced foF2 values into a global mapping scheme, it is necessary to compare the theoretically-based values of foF2 with those determined from the existing coefficients alone and with actual observations. It is not possible, in general, to obtain a large number of observations of foF2 in the southern hemisphere. Fortunately, however, the topside sounder observations from the ISS-b satellite, which was launched in February 1978, have provided values of foF2 since then. Matuura (1979) has deduced global maps of foF2 from the ISS-b satellite observations, and these maps of foF2 provide a reasonable data base for assessing the reliability of these models for estimating foF2.

Figure 7 shows three longitudinal distributions of foF2 at three latitudes (30°S, 40°S, and 50°S) for December 1978, 1000 UT: 1) the distribution from the ISS-b satellite observations; 2) the distribution deduced by incorporating theoretical values of foF2 into the mapping scheme; and 3) the distribution of foF2 resulting from the existing numerical coefficients. Because the numerical coefficients have been adopted by International Radio Consultative Committee (CCIR) as the basis for ionospheric parameter specification, the distribution determined from the coefficients is indicated as CCIR. It is readily apparent that differences between observed and mapped values of foF2 exist. However, it is clear that the values of foF2 deduced using the theoretical model tend to represent the observations better than do the values determined from the CCIR coefficients alone.

Figure 8 shows a comparison between the diurnal variation of foF2 deduced from the theoretical calculations (depicted as a dashed line) and the diurnal variation of foF2 deduced from the CCIR coefficients (solid line). The diurnal variation of foF2 in this case was determined for September 1978 at the location closest to Kerguelen (49.4°S and 70.3°E geographic). In addition to the values of foF2 calculated by the theoretical model and by the CCIR coefficients, the median values of foF2 observed at Kerguelen during September 1978 are shown in Figure 8. Also indicated on the figure are values of foF2 deduced from the data collected by the ISS-b satellite. Even though Kerguelen is nearly ten degrees south of the latitude for which the neutral air wind model was deduced, it is readily apparent that the theoretical model yields results that agree much better with the observations than results obtained from the CCIR coefficients.

Figure 9 provides a similar comparison for December 1978. In this example, it is seen that improvement afforded by use of the theoretical model is not as straight-forward as seen in Figure 8. The theoretical model provides a better estimate of the nighttime values of foF2 observed at Kerguelen than do the CCIR coefficients. The CCIR coefficients, on the other hand, provide a better representation for the daytime values of foF2 observed at Kerguelen. The theoretical model, however, appears to provide a better estimate of the daytime values of foF2 observed by the ISS-b satellite.

DISCUSSION

There have been a number of studies directed toward determining the accuracy of the existing maps of foF2. Many of these (King and Slater, 1973, and Wakai and Matuura, 1980, for example) have been directed toward comparing the values determined from the existing maps with average values of foF2 determined from satellite data. Studies have also been undertaken (Rush and Edwards, 1976; Flattery and Ramsay, 1975) to develop methods to specify the global distribution of foF2 on an hourly basis. All these studies rely upon the monthly median maps of foF2 determined from the Jones-Gallet coefficients. The values of foF2 predicted from these coefficients have been found to agree well with observations at those locations for which data were used in the development of the coefficients. However, at regions of the globe that are inaccessible to routine ionospheric sounding, the agreement is less than optimum. A major source of error in these regions can be attributed to the fact that, in developing the coefficients, no account was taken of the effect of the magnetic declination on the foF2 distribution. In order to assure that a stable solution was obtained in generating the CCIR coefficients, it was necessary to perform a "screen analysis" whereby data observed at specific locations were extrapolated using a magnetic dip latitude to regions for which no observations were available. For the most part, the effect of the screen analysis was to extrapolate values of foF2

observed in land regions to ocean areas. In this extrapolation no account was taken of the adjustments in foF2 needed because the magnetic declination is significantly different between regions of observation and regions of extrapolation. The use of the time-dependent equation to calculate values of foF2 provides a means to overcome this limitation. Combining the theoretical values of foF2 with observed monthly median values to regenerate the numerical coefficients can improve the global maps of foF2.

REFERENCES

Anderson, D. N., Plan. Space Sci., 21, 409, 1973.

Barghausen, A. F., J. W. Finney, L. L. Proctor, and L. D. Schultz, ERL 110-ITS-78, ESSA Research Laboratories, 1969.

Bent, R. B., J. R. Lipofsky, S. K. Llewellyn, and P. G. Schmid, AGARD-CP-238-Vol. 1, AGARD, November 1978.

CCIR, Supplement to Report 252-2, Geneva, 1978.

Crank, J. and P. Nicolson, Cambridge Phil. Soc. Proc., 43, 50, 1947.

Flattery, J. W. and A. C. Ramsay, Proc. of Ionospheric Effects Symposium, p. 336, Naval Research Laboratory, 1975.

Hedin, A. E., C. A. Reber, G. P. Newton, N. W. Spencer, H. C. Brinton, H. G. Mayer, and W. E. Potter, J. Geophys. Res., 82, 2148, 1977.

Jones, W. B. and R. M. Gallet, ITU Telecommun. J. 29, 129, 1962.

Jones, W. B. and D. L. Obitts, OT/ITS RR3, NTIS Acces. No. COM75-11143/AS, 1970.

Jones, W. B., R. P. Graham, and M. Leftin, ERL 107-ITS 75, ESSA Research Laboratories, 1969.

Kendall, P. C. and W. M. Pickering, Planet. Space Sc., 15, 825, 1967.

King, J. W. and A. J. Slater, ITU Telecomm. J., 40, 766, 1973.

Leftin, M., OT Rpt 76-88, NTIS Acces. No. PB 255484/AS, 1976.

Matuura, N., Atlas of Ionospheric Critical Frequency (foF2) obtained from Ionospheric Sounding Satellite-b Observation, Part 1, August to December 1978, issued by Radio Research Laboratories of Japan, March 1979.

Nisbet, J. S., Radio Sci., 6, 437, 1971.

Rosich, R. K. and W. B. Jones, OT Rpt 73-22, NTIS Acces. No. COM75-10813/AS, 1973.

Rush, C. M. and W. R. Edwards, Radio Sci., 11, 931, 1976.

Rush, C. M. and J. Gibbs, AFCRL Tec. Rpt 73-0335, Defense Documentation Center, 1973.

Stern, D., J. Geophys. Res., 70, 3629, 1965.

Stern, D., J. Geophys. Res., 72, 3995, 1967.

Wakai, N. and N. Matuura, Acta Astonautica, 7, 999, 1980.

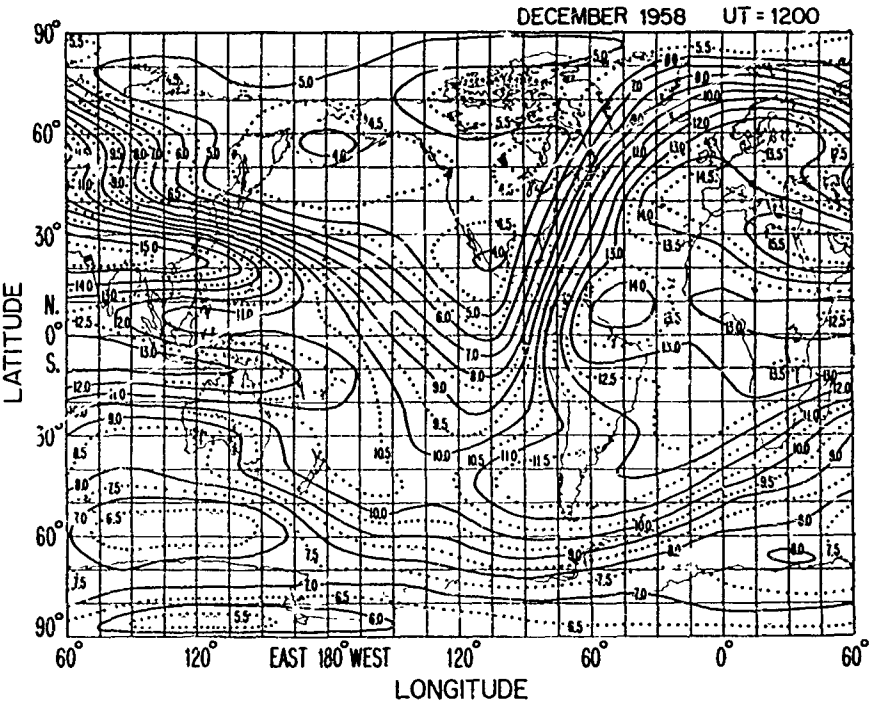


Figure 1. Contours of median values of fxF2 (in MHz) for December 1958, UT = 12 hrs.

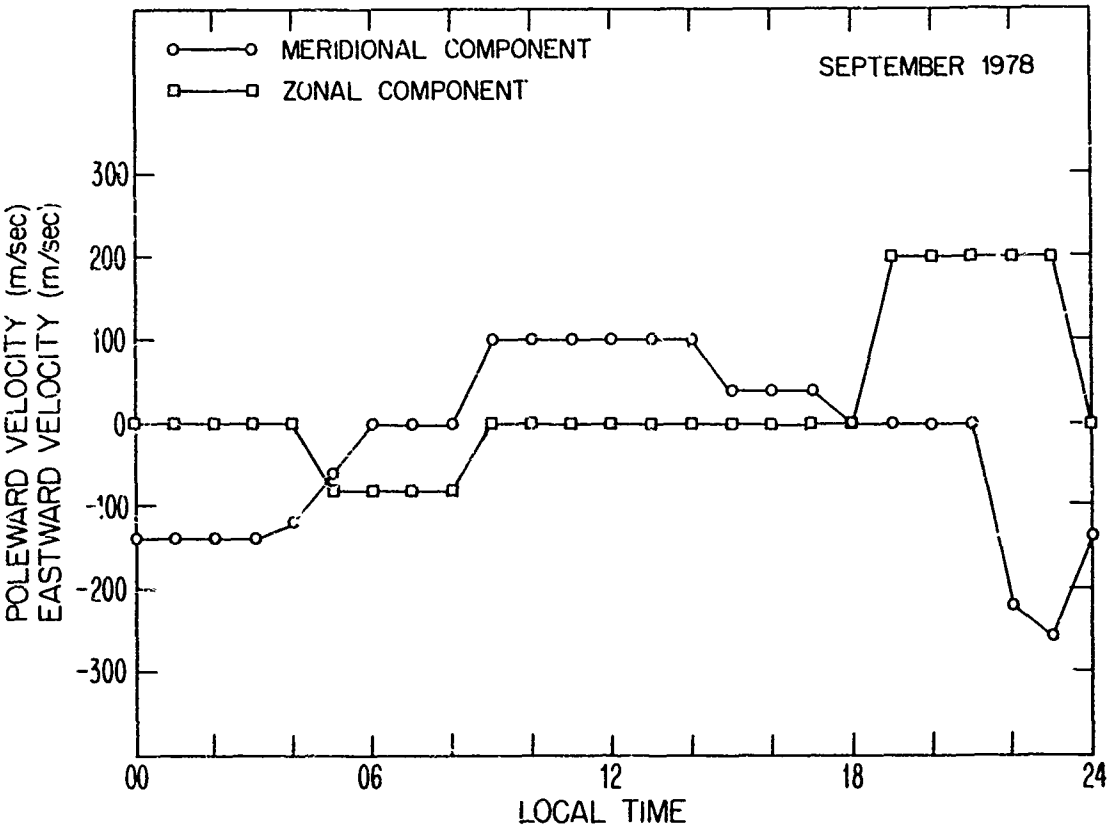


Figure 2. Meridional and zonal wind components for September 1978.

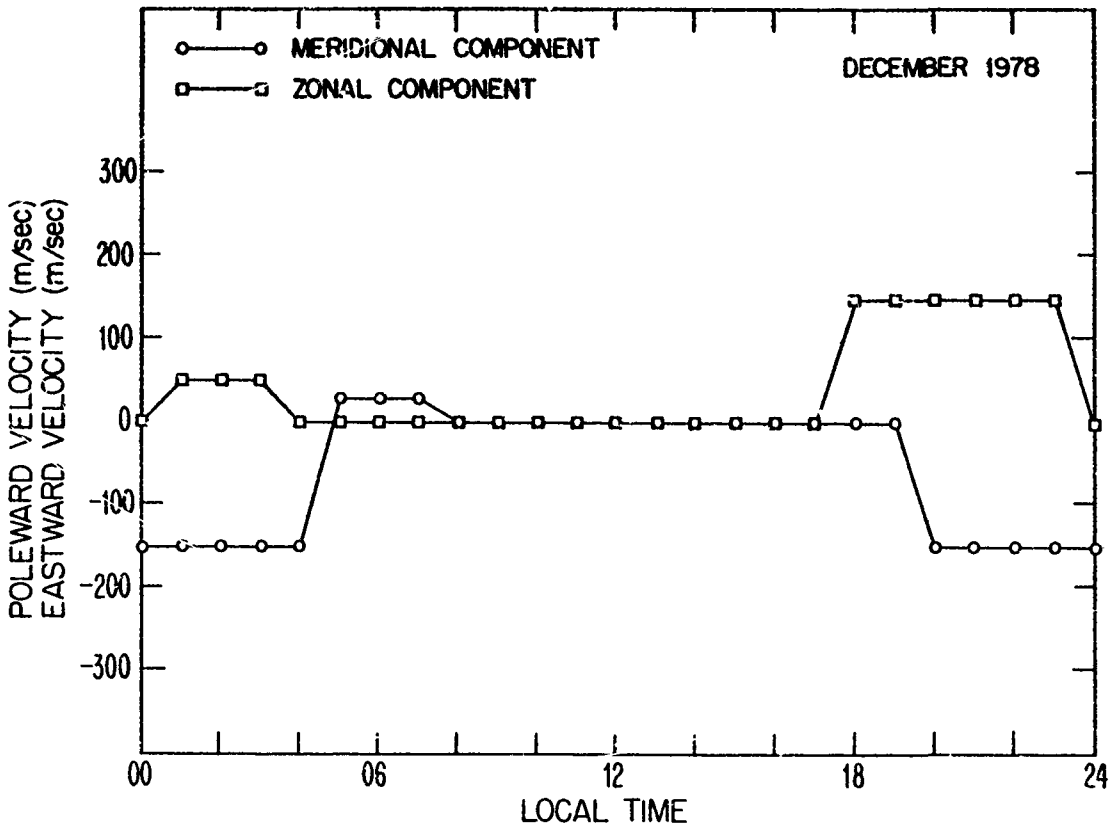


Figure 3. Meridional and zonal wind components for December 1978.

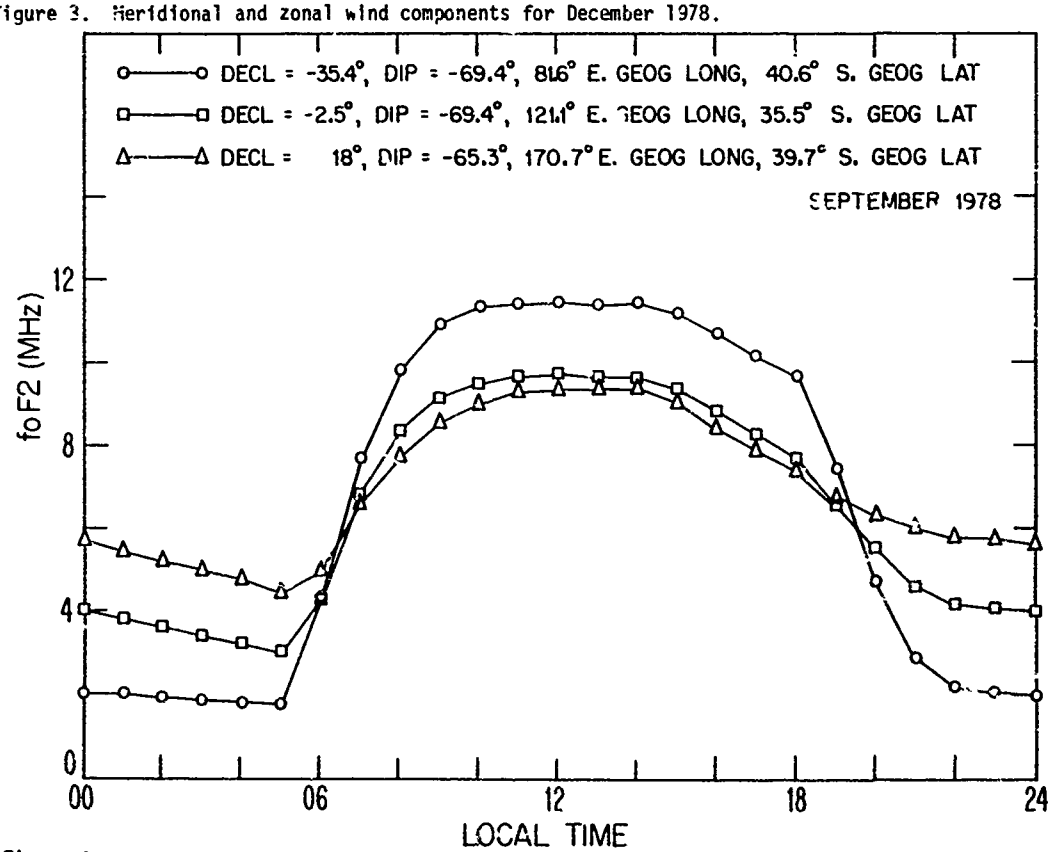


Figure 4. Values of foF2 calculated at locations indicated for September 1978.

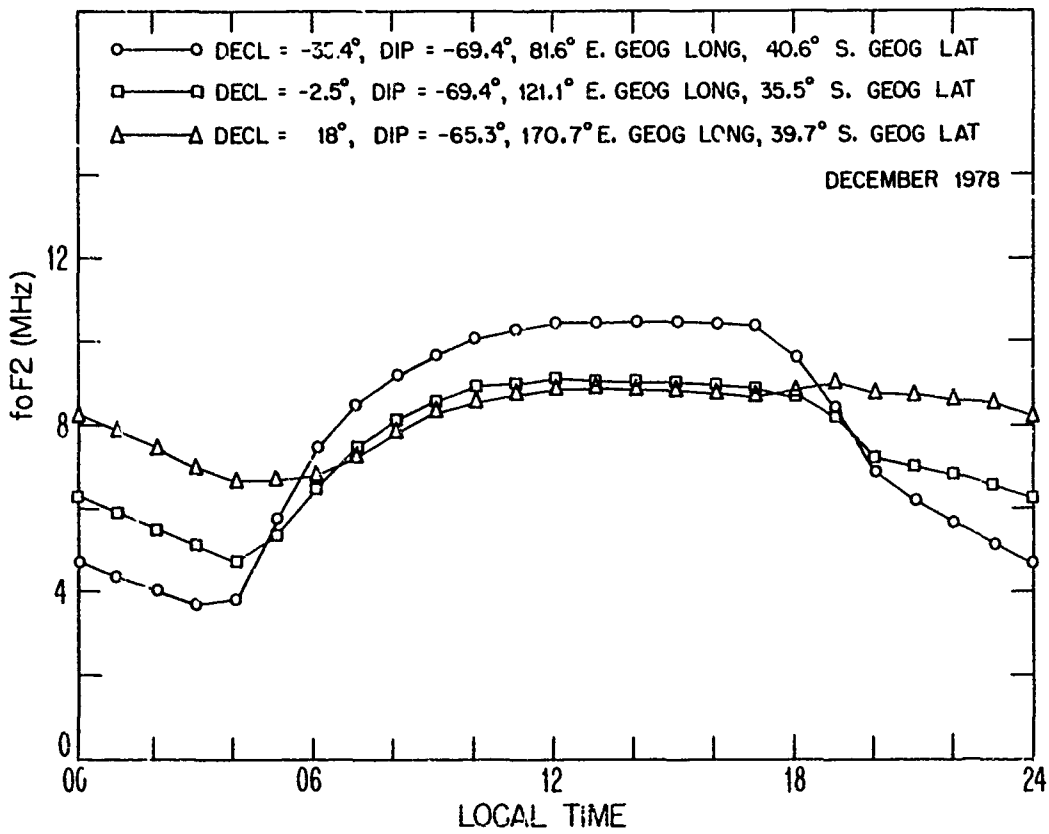


Figure 5. Values of foF2 calculated at locations indicated for December 1978.

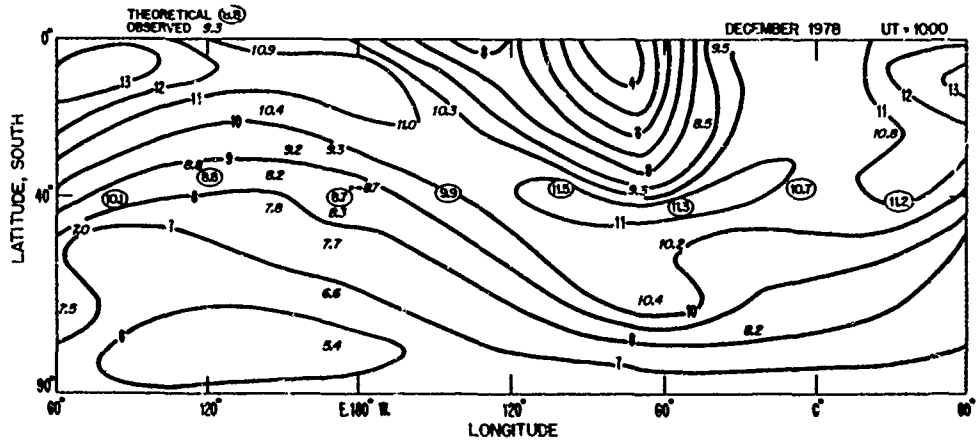


Figure 6. Contours of foF2 (in MHz) obtained by combining theoretical values, observations, and numerical coefficients for December 1978.

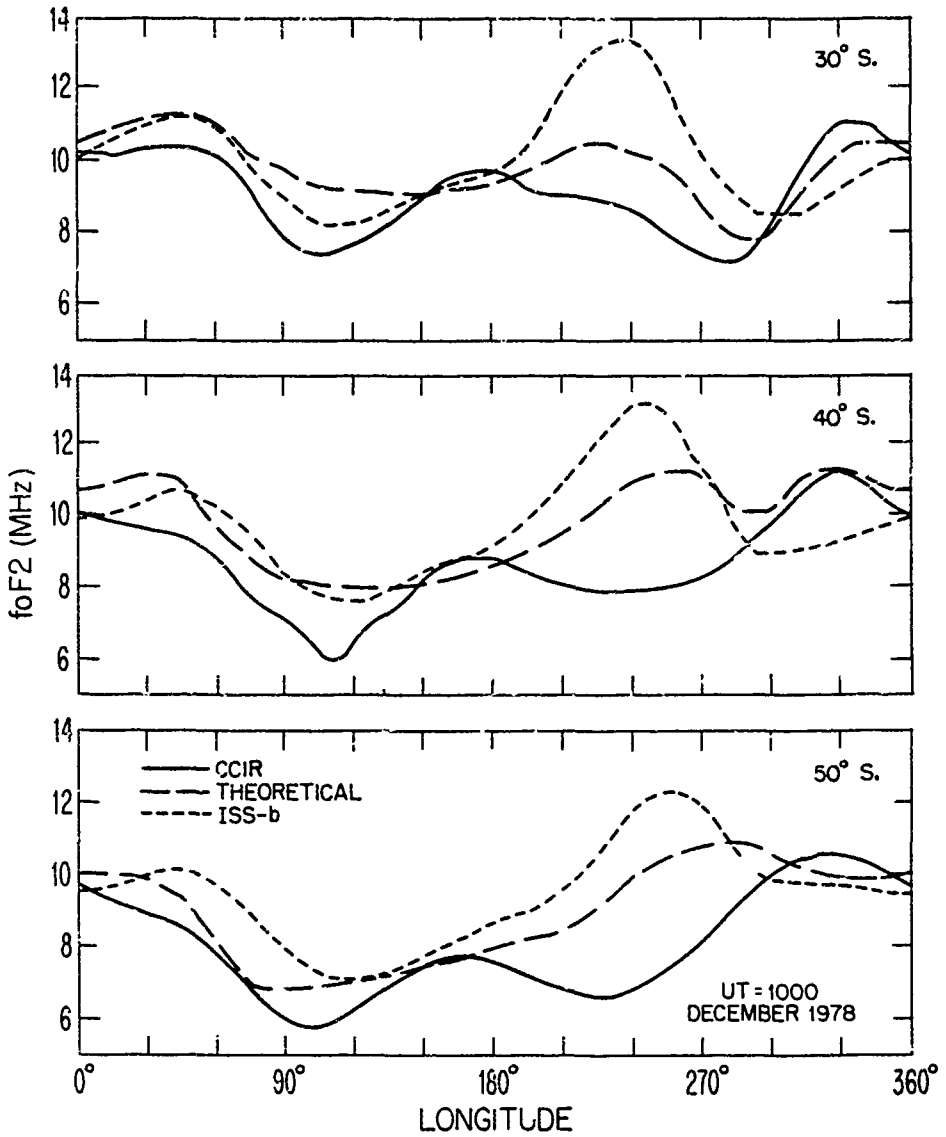


Figure 7. Longitudinal distribution of foF2 for 10 UT, December 1978 at 30°S, 40°S, and 50°S determined from the CCIR coefficients, theoretical calculatio. and ISS-b observations.

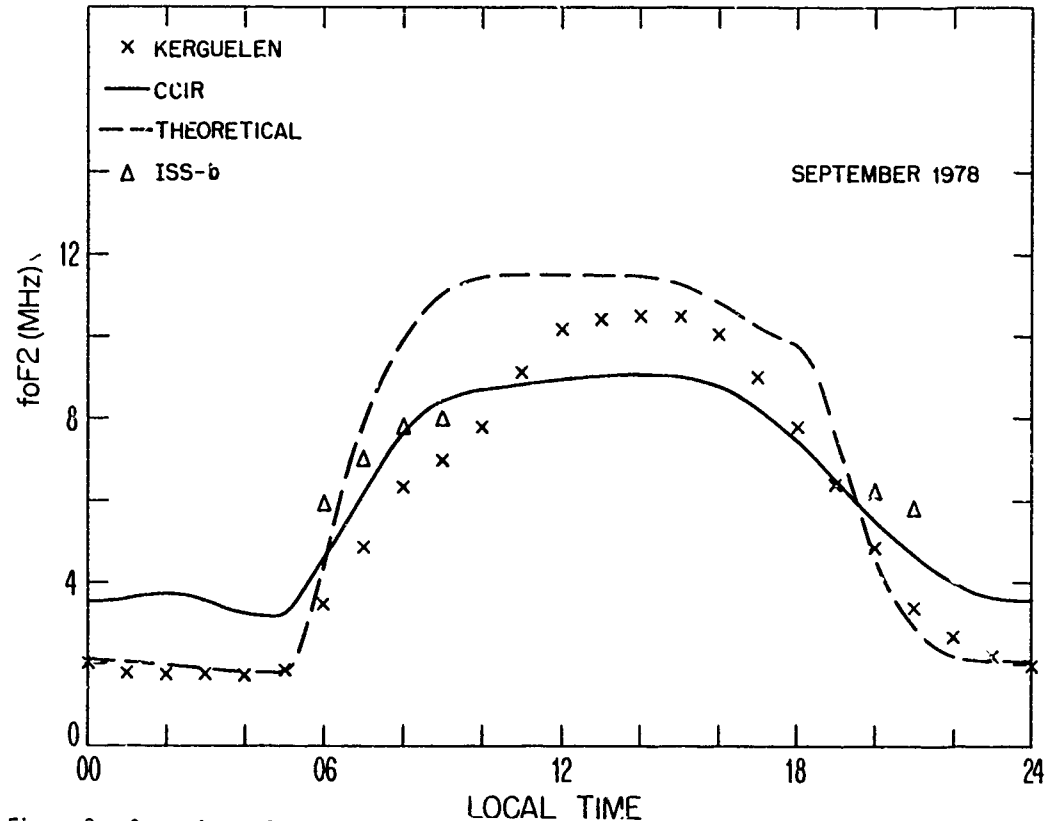


Figure 8. Comparison of foF2 determined from CCIR coefficients theoretical calculations, Kerguelen observations and ISS-b observations for September 1978.

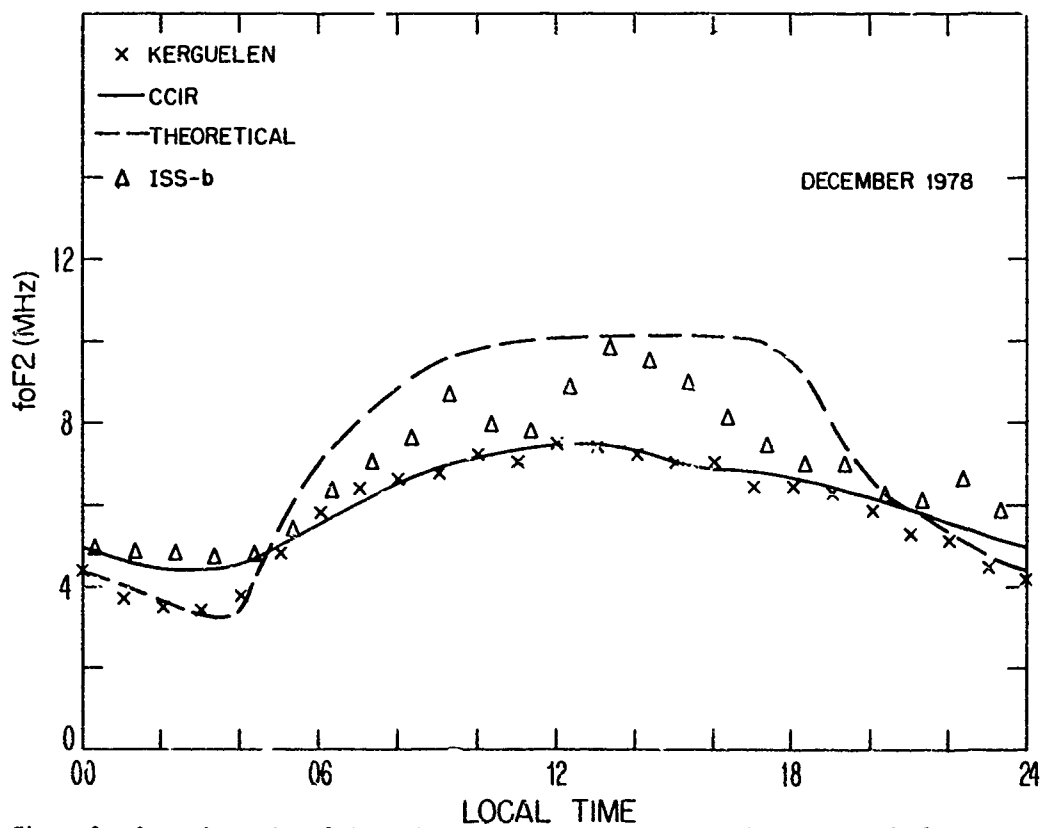


Figure 9. Comparison of foF2 determined from CCIR coefficients, theoretical calculations, Kerguelen observations and ISS-b observations for December 1978.

SATELLITE ORBITS TO RELIEVE IONOSPHERIC SCINTILLATION

Dorothy C. Rogers Paul Christopher
The MITRE Corporation
Bedford, MA 01730

I. BACKGROUND

During the past 20 years, many comprehensive ionospheric scintillation studies have occurred. These have dealt with the nature of the disturbed ionosphere itself and with sophisticated communication techniques for overcoming scintillation. The Air Force Geophysics Laboratory (Aarons, Whitney, Allen 1970) made one of the first morphology estimates of the disturbed ionosphere. It was valuable for advancing satellite communication in the VHF-11 GHz region. Communicators under one of the disturbed regions attempted power supplements, coding, or diversity. R. K. Crane advanced these techniques by describing the standard deviation of the received signal in terms of the spatial power spectrum of the disturbed ionosphere and the relative angle of signal to magnetic field line (Crane, 1974).

This paper takes an unconventional approach by attempting to avoid ionospheric scintillation altogether by offering a choice of satellites at critical ground locations. The possibility of avoiding scintillation by proper choice of satellites has been almost totally ignored in the past. Perhaps this oversight has occurred because geosynchronous satellites have been so successful that they were the only class of satellites considered. As the geosynchronous orbital slots fill, however, we are led to consider other satellites - and indeed to see that they are not only necessary, but useful.

Aaron's early 1968-1970 ionospheric irregularity maps (figure I-1) are used as a first approach to avoid scintillation. These estimates gave clearly defined regions for an auroral ring and for an asymmetric disturbed equatorial region on the nightside. Later estimates of intensity include the sunspot index, but for the purpose of this paper, we omit the intensity fluctuations induced by sunspot and K_p variations and deal only with the locations of the disturbed regions. These valuable Air Force Geophysical Laboratory maps have often been used in a 2-dimensional sense: ground stations lying under a critical region have been assumed susceptible to scintillation. We alter this 2-dimensional morphology to a 3-dimensional structure in an inertial coordinate system (figures I-2 to I-4).

The nightside coordinates (figure I-3) contain all the scintillation regions of interest here. Notice the earth is rotating in this coordinate system, so the high latitude ring wobbles as viewed by an observer on the (-x) axis. The coordinates of the disturbed region require geomagnetic colatitude (but inertial longitude is used because the sun causes extinction). When the disturbed region is at a nominal 300 kilometers altitude, it is seen that a variety of look angles from the ground station may allow the ground station to escape the disturbed region. An analogous situation occurs for a pedestrian peering beneath the bottom edge of an umbrella: the pedestrian may prefer to look up, but he will look horizontally rather than have his vision totally obstructed.

To this 3-dimensional ionospheric map, a dynamic model of satellite orbits is superimposed. The ground-satellite paths are continually calculated by the ground station and the ionospheric piercing points found. All satellite paths can be checked, and any path traversing a disturbed region can be given a low priority.

II. ORBITS TO RELIEVE SCINTILLATION

Efforts to relieve scintillation could concentrate on three equally spaced geosynchronous satellites, but their restriction to the equatorial plane offers little choice to the ground site. This will be seen later when ray paths are examined statistically for the entire Northern Hemisphere. Another class of satellites can also be useful for large area coverage: three inclined elliptic 12 hour (Molniya-type) satellites. These were shown to offer elevation angle statistics comparable to synchronous satellites (Christopher, 1980) for the Northern Hemisphere. Figure II-1 shows the different planes and time phasing required for convenient Molniya coverage. A combination of three

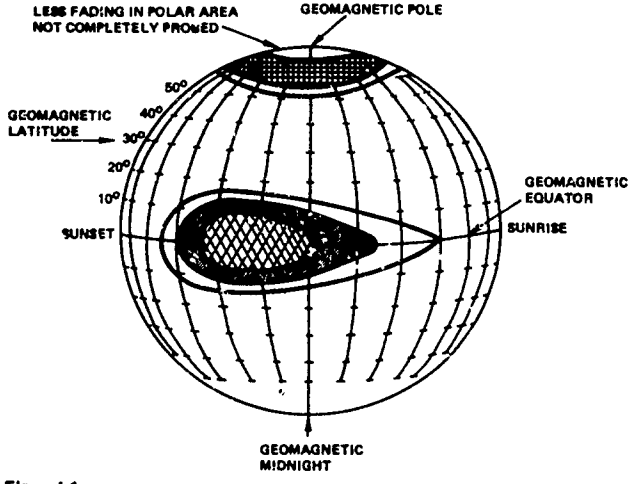


Figure I-1. AARONS' IONOSPHERIC IRREGULARITY STRUCTURE AT NIGHT

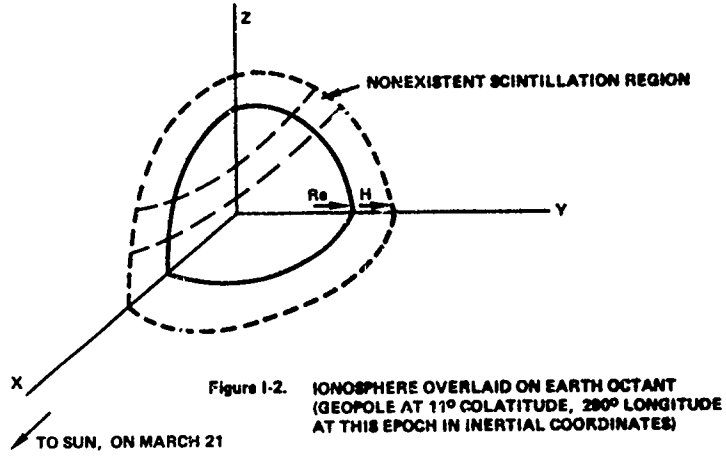


Figure I-2. IONOSPHERE OVERLAID ON EARTH OCTANT (GEOPOLE AT 11° COLATITUDE, 290° LONGITUDE AT THIS EPOCH IN INERTIAL COORDINATES)

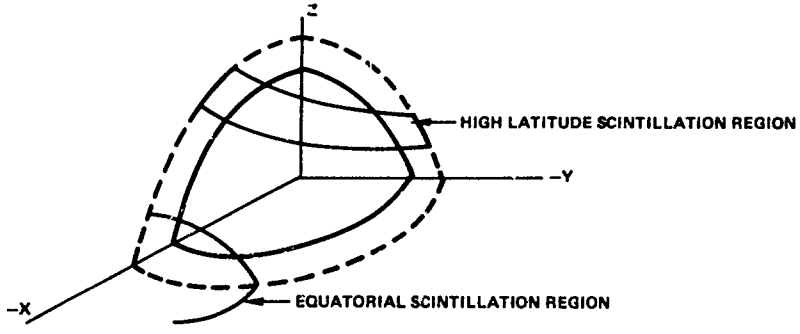


Figure I-3. AARONS' PRIME SCINTILLATION REGIONS IN AN (NIGHTSIDE) INERTIAL COORDINATE SYSTEM

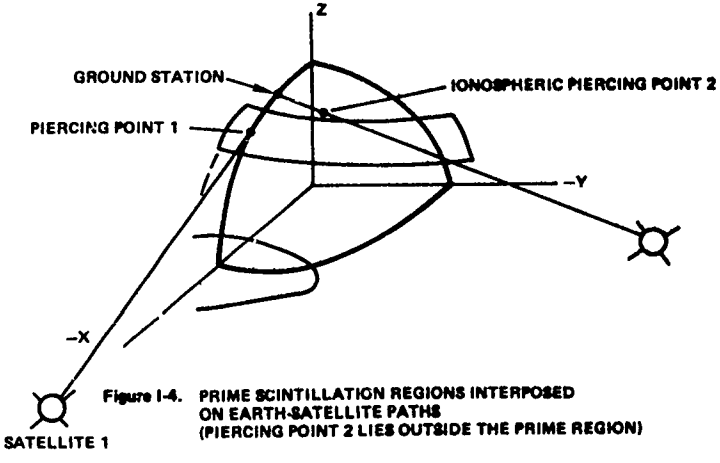


Figure I-4. PRIME SCINTILLATION REGIONS INTERPOSED ON EARTH-SATELLITE PATHS (PIERCING POINT 2 LIES OUTSIDE THE PRIME REGION)

Molniya and 2 synchronous satellites offers outstanding coverage and significant relief from the prime scintillation regions, as will be shown later in the statistical results.

An important class of inclined, 24 hour satellites (Walker, 1970) can also give excellent elevation angle statistics and relief from the prime scintillation regions. Figure II-2 shows a spatial relation of the first two Walker satellites at T=0 hrs. and figure II-3 shows the subsatellite traces over a 24 hour period. Walker pointed out that these five satellites are all that are required for worldwide coverage (unlike earlier predictions which required more satellites).

A ground station with a choice of multiple satellites may wish to maximize expected link performance. He may be hurt by choosing a path through a prime scintillation region, or he may be hurt by choice of low elevation angle (above the local horizon). Low elevation angle can allow the ground receiver to be more susceptible to adjacent terrestrial interference sources.

Specifically, received terrestrial interference power

$$P_{ri} = P_{ti} \; G_{rs} \; G_{ti} \left(\frac{\lambda}{4\pi R_i}\right)^2 \frac{1}{\ell_i}$$

where P_{ri} = received interference power, watts.

P_{ti} = transmitted interference power, watts

G_{rs} = sidelobe gain of receive antenna

G_{ti} = gain of interference source

λ = wavelength, km

R_i = distance between interference source and receiver, km.

ℓ_i = atmospheric loss between interference source and receiver.

The signal level received from the satellite is

$$P_{rs} = P_{ts} \; G_{ro} \; G_{ts} \left(\frac{\lambda}{4\pi R_s}\right)^2 \cdot \frac{1}{\ell_s} \cdot \frac{1}{\ell_p}$$

and

P_{rs} = received signal power, watts

P_{ts} = satellite transmitter power, watts

G_{ro} = mainbeam gain of receive antenna

G_{ts} = mainbeam gain of transmit antenna

R_s = distance between satellite transmitter and receiver, km.

ℓ_s = atmospheric loss on trans-atmospheric path

ℓ_p = equivalent system loss caused by ray traversing a prime scintillation region. An index of link performance can be found for the ground station by forming the ratio of interference-to-signal (I/S).

The distance to the interference source and configuration of that source are beyond the control of the ground receiver, and if he neglects ℓ_i and ℓ_s, the only variables which affect his choice are

$$I/S \propto \frac{G_{rs}}{G_{ro}} \; \ell_p \; R_s^2$$

The ratio of sidelobe envelope-to-mainbeam gain can be found to be proportional to cosecant^3 (off boresight angle) for a uniformly illuminated receiver dish*. The smallest angle which the

*expansions for $J_1(U)/U$ in Abramowitz & Stegun, Handbook of Mathematical Functions, Dover, N.Y., 1964.

interference source can present to the boresight is elevation angle E; the gain ratio will be set at $CSC^3(E)$ for the subsequent discussion. Alsc

$$\ell_p = 10^{\frac{LZ(cscE)}{10}} \quad \text{(ray path traversing a disturbed region)}$$

$$\ell_p = 1 \text{ (ray patn outside a disturbed region)}$$

where

LZ = zenith loss through the disturbed region, dB

E = elevation angle

The ground receiver would then choose a satellite (i) which would minimize the ratio

$$\left(\frac{I}{S}\right)_i = \left[\csc^3(E_i)\right]\left[10^{\frac{LZ}{10}\cdot\csc(E_i)}\right]R_{S_i}^2$$

where the subscripts relate to a particular satellite.

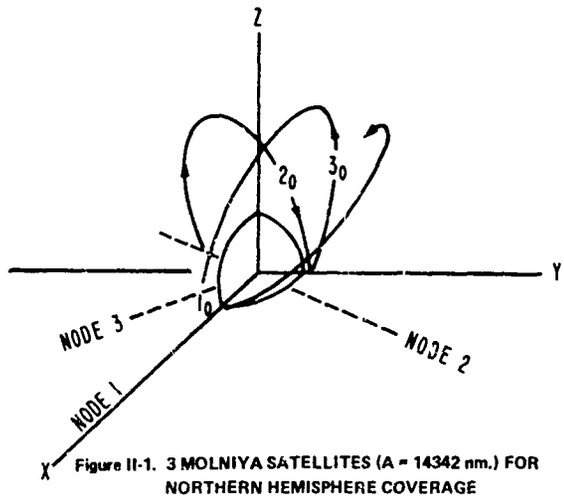


Figure II-1. 3 MOLNIYA SATELLITES (A = 14342 nm.) FOR NORTHERN HEMISPHERE COVERAGE

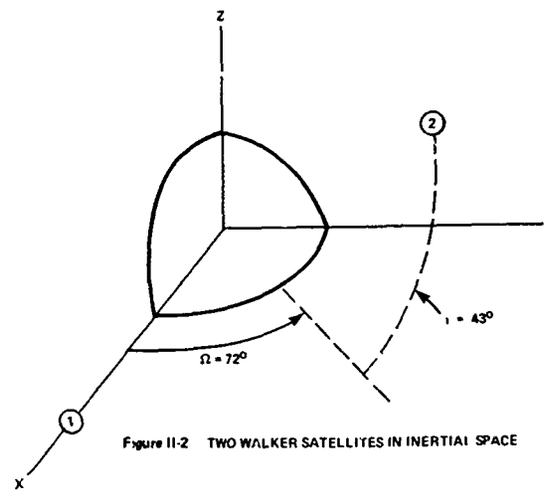


Figure II-2 TWO WALKER SATELLITES IN INERTIAL SPACE

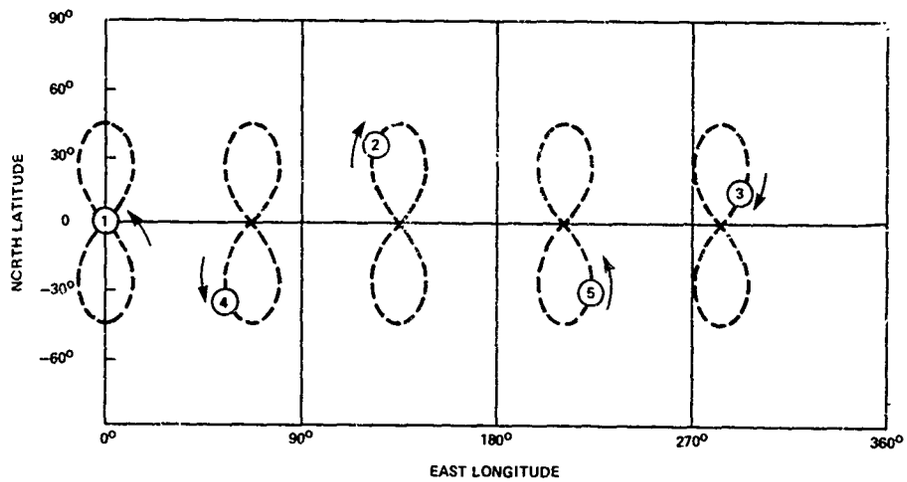


Figure II-3. SUBSATELLITE POINTS OF FIVE SATELLITES IN WALKER ORBITS t = 0

This index of link performance ("objective function") allows the ground receiver to choose, for example, between a satellite at low-elevation angle which allows a ray path outside the prime disturbed region and a satellite at high elevation angle whose ray traverses the prime region. He seeks to minimize the objective function. At LZ = 20 dB (a very disturbed ionosphere), the ground observer would be forced to choose a 10° elevation angle in the clear rather than a 30° elevation angle in the prime region. Satellite constellations which normally would allow excellent elevation angle statistics would then suffer a degradation in elevation angle statistics while improving system performance.

III. DETERMINISTIC RESULTS

A computer program has been developed to calculate ground station-satellite paths and to determine whether their piercing points lie in a disturbed region. The sample computer output shown in tables III-1a, b give link availability for a ground station at 70° North latitude, 0° East longitude communicating with three phased Molniya-type satellites during a 24 hour period. From left to right, the columns are satellite number, time in hours, range in kilometers, elevation angle in degrees, azimuth in degrees, scintillation indicator (SCIN), piercing point inertial longitude, piercing point colatitude relative to the geomagnetic pole assumed to be 11° colatitude, 290° East longitude, subsatellite longitude, and subsatellite colatitude. SCIN = 0 if no scintillation occurs. SCIN = 1 if scintillation occurs in the auroral region; SCIN = 2 if scintillation occurs in the equatorial region. A negative elevation angle indicates that the satellite is not visible from the ground station in which case the scintillation indicator has no meaning. It is seen that ground station communication will be poor or impossible for hours 4, 5, 7 through 13, and 15 through 17 with satellite 2, and it could be seen for hours 7 through 18, 20, and 21 with satellite 3. These hours of poor or nonexistent links do not completely overlap, indicating that the ground station may be able to avoid scintillation during some hours of the night by choosing the appropriate satellite link.

Table III-1a. Path Parameters to Molniya 1 (ground station at 70°N, 0°E)

SAT	HRS	R, KM	ANGLE	AZIMUT	SCIN	FH	THT	SUBL	SUBC
1	1.00	15618.984	29.611	181.099	0.0	14 801	22.777	359.152	63.942
1	2.00	24492.750	59.831	178.13[illegible]	0.0	30 132	20.420	1.097	44 071
1	3.00	31309.266	71.872	179.17[illegible]	0.0	45 034	19.851	0.372	35 090
1	4 00	35937.352	78.133	180.980	0.0	59 974	19.585	359.658	30.086
1	5.00	38617.855	81 363	181.518	0.0	74 970	19.458	359.574	27.415
1	6 00	39495.969	82.378	179 999	0.0	90 000	19.425	0.000	26 565
1	7 00	38617.266	81.362	178 467	1.0	105.030	19.469	0.430	27 416
1	8 00	35936 531	78.133	179.010	1.0	120 027	19 595	0.346	30 087
1	9 00	31307.832	71.870	180.814	1.0	134 966	19.838	359.632	35 092
1	10 00	24490 803	59.827	181.864	1.0	149.868	20.370	358.907	44.074
1	11 00	15616 7 9	29.598	178 895	1.0	165 200	22.855	0.854	63.951
1	12 00	13122 332	-72.463	124.293	0.0	268 105	162 037	90.125	153 435
1	13 00	19930.703	-12.465	0.761	0.0	13 211	20.407	179.157	63.934
1	14 00	27979.926	14.119	359.149	0.0	209 426	13.044	181.101	44 068
1	15 00	34173 609	26.118	359 737	1.0	224 915	15.287	180.376	35 089
1	16 00	38417.411	32.607	0 221	1.0	240 053	16.070	179.662	30 085
1	17 00	40885 793	35.997	0 264	1.0	255.055	16 393	179.578	27 415
1	18 00	41695.375	37 064	360 000	1.0	270 000	16.477	180.000	26.565
1	19 00	40884 543	35 996	359.729	0 0	284.943	16.373	180.434	27 416
1	20.00	38415.094	32 604	359 771	0 0	299.945	16.050	180.350	30 088
1	21 00	34169.699	26 111	0 256	0.0	315 083	15.316	179.635	35.094
1	22.00	27974.375	14.106	0 842	0.0	330 568	13.228	178.911	44 077
1	23 00	19924.168	-12.496	359.224	1.0	163 819	20.008	180.859	63.959
1	24.00	13118.328	-72.398	235.431	0.0	272 327	141.645	270.307	153 435

Table III-1b. Path Parameters to Molniya 2

SAT	HRS	R, KM	ANGLE	AZIMUT	SCIN	FH	THT	SUBL	SUBC
2	1.00	31307.832	71.870	180 814	0.0	14.966	19.838	359 632	35 092
2	2 00	24490 803	59.82[illegible]	181.864	0.0	29.868	20.370	358.907	44.074
2	3 00	15616 719	29.598	178.895	0.0	45 200	22.855	0.854	63 951
2	4.00	13122.332	-72.463	124 293	0.0	148 106	162.037	90.125	153.435
2	5.00	19930 703	-12 465	0 761	1 0	253.211	20.407	179.157	63 934
2	6 00	27979.926	14 119	359 149	0.0	89.426	13.044	181 101	44 068
2	7 00	34173.609	26 118	359 737	1.0	104.915	15.287	180.376	35 089
2	8 00	38417 441	32 607	0.221	1.0	120.053	16.070	179.662	30 085
2	9.00	40885 793	35.997	0.264	1.0	135.055	16.393	179.578	27 415
2	10 00	41695 375	37 064	360.000	1.0	150 000	16.477	180.000	26 565
2	11.00	40884 543	35.996	359.729	1.0	164.944	16.373	180.434	27.416
2	12 00	38415.094	32.604	359.771	1.0	179.945	16.050	180.349	30.088
2	13 00	34169.699	26.111	0.256	1.0	195 083	15.316	179.635	35.094
2	14 00	27974.375	14.106	0 842	0.0	210.568	13.227	178.911	44.077
2	15.00	19924.168	-12.496	359.224	0.0	46 819	20.008	180.859	63.959
2	16 00	13118 328	-72.398	235.431	0 0	152 327	141.645	270.307	153 435
2	17.00	15623.633	29.639	181.085	1.0	254.804	22.774	359.162	63 925
2	18.00	24496.863	59.840	178.116	0.0	270.133	20.419	1.104	44 065
2	19.00	31312 371	71.876	179.162	0.0	285 034	19.851	0.379	35 087
2	20.00	35939.270	78.136	180.957	0.0	299.974	19.585	359.666	30 084
2	21.00	38618.738	81.364	181 490	0.0	314.971	19.458	359.583	27 414
2	22 00	39495.969	82.378	179 999	0.0	330 000	19.425	0.000	26 565
2	23.00	38618.383	81.361	178.440	0.0	345 031	19.469	0.437	27.416
2	24.00	35934.609	78.130	178.989	0.0	0.027	19.595	0.353	30.089

The possibility of avoiding scintillation through the correct choice of satellites is demonstrated in figure III-1 for a ground station at 70° North, 0° East. The index of link performance described in Section II is plotted against time for two synchronous equatorial satellites, S1 and S2, with right ascensions at 0° and 60°, respectively, and three phased Molniya satellites, M1, M2, and M3.

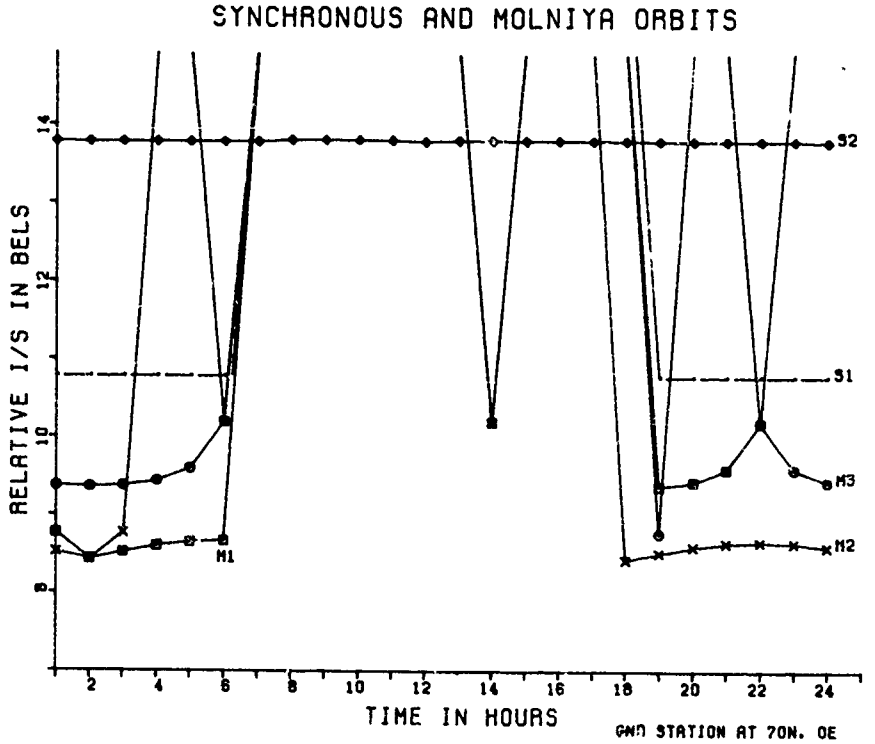


FIGURE III-1 OBJECTIVE FUNCTIONS FOR SYNCHRONOUS AND MOLNIYA ORBITS. NORTHERN SITE.

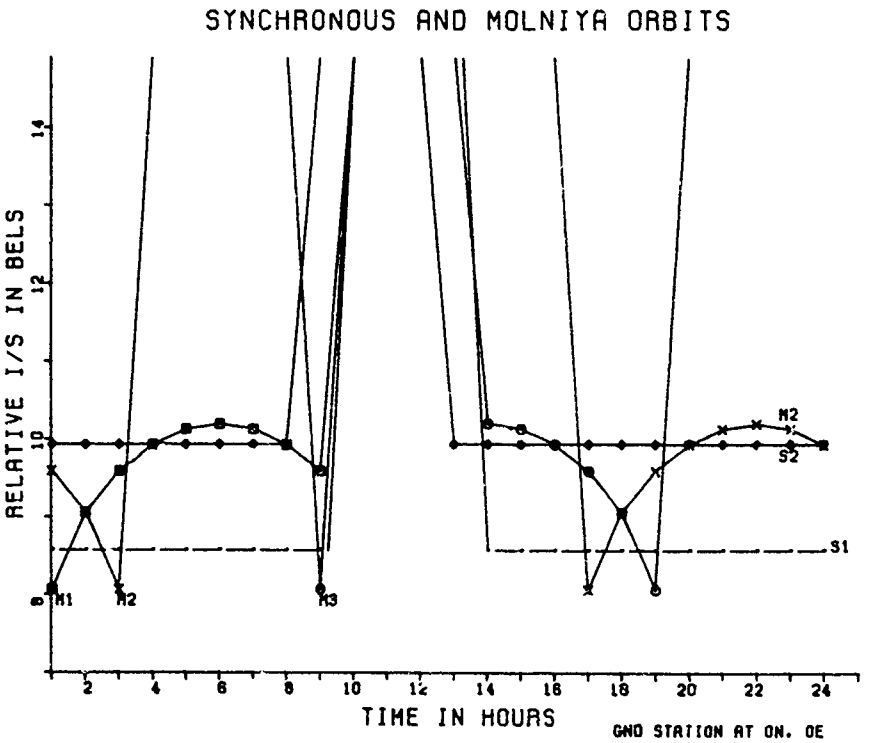


FIGURE III-2 OBJECTIVE FUNCTIONS FOR SYNCHRONOUS AND MOLNIYA ORBITS. EQUATORIAL SITE.

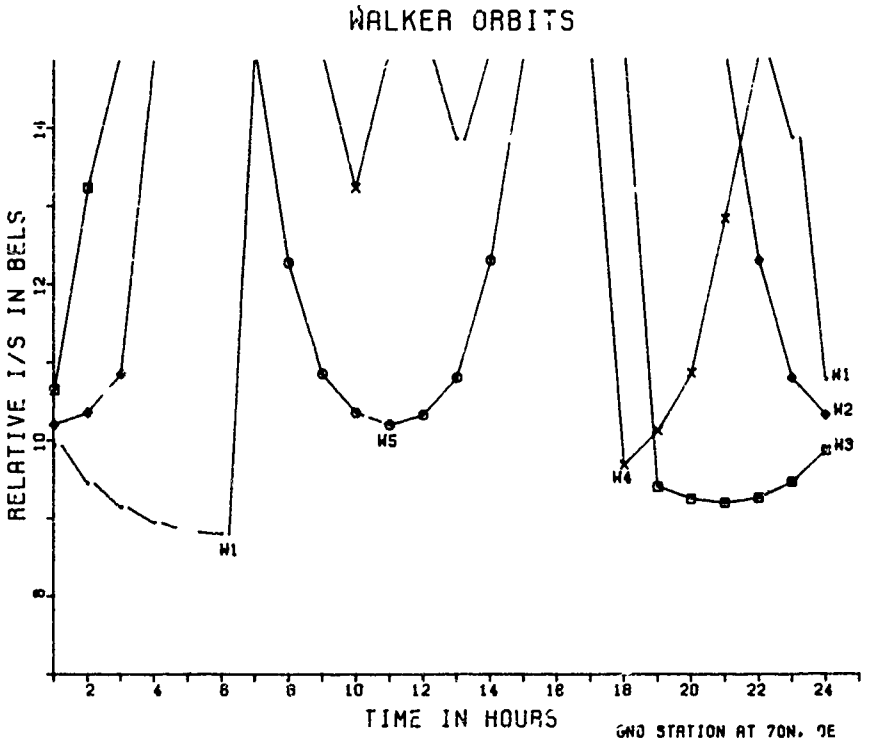


FIGURE III-3 OBJECTIVE FUNCTIONS FOR WALKER ORBITS. NORTHERN SITE.

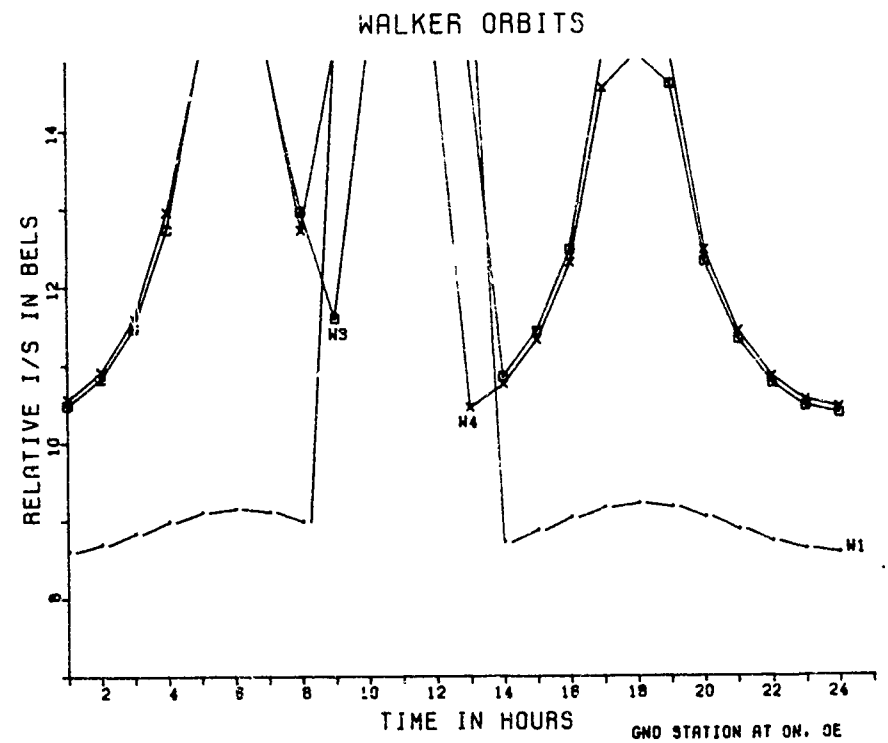


FIGURE III-4 OBJECTIVE FUNCTIONS FOR WALKER ORBITS. EQUATORIAL SITE.

Objective function values in excess of 15 Bels are represented by lines going off the top of the graph, indicating that either the satellite is below the horizon or that scintillation occurs at that time. Although no scintillation occurs with satellite S2 during the entire 24 hour period, better link performance is achieved with satellite M1 during hours 2 through 6 due to its higher elevation angle. Satellite S2 becomes useful during hours 6 through 13 and 15 through 17 when scintillation in the auroral region is occurring for the other satellites. For hours 1, 14, and 18 through 24, satellite M2 provides the best link, again due to a higher elevation angle. Note that objective function values remain fairly constant for M1 during hours 1 through 6 and for M2 during hours 18 through 24 because the increase in elevation angles compensates for the increase in range.

Figure III-2 shows the same satellite constellation with the ground receiver now at 0° East. A comparison of this with the preceding figure reveals that the Molniya satellites are less useful for a ground station on the equator than for a more northerly site. This is because the Molniya satellites spend less time above the horizon and have greater ranges and lower elevation angles with respect to the equatorial site.

Link performances for five Walker satellite orbits and a ground station at 70° North, and 0° East is displayed in figure III-3. Satellite W1 is seen to be rising during hours 1 through 6, providing a good link until scintillation occurs at hour 7. Scintillation ceases at hour 12, but by then W1 has begun to descend below the horizon. Satellite W5 is visible between hours 8 and 15 and provides a link for most of the time that scintillation is occurring for W1. Scintillation ceases for satellite W4 and W3 at hours 17 and 18, respectively, but W3 shows better performance since W4 is beginning to drop rapidly.

Figure III-4 shows that only three of the Walker satellites are ever visible to the ground station at 0° North, 0° East. Satellite W1 is always visible with a high elevation angle and fails to provide communications only during the hours when scintillation occurs.

IV. STATISTICAL RESULTS

The measure of link performance changes continuously in a moving satellite system at a fixed ground site, as shown in Section III. A choice of "best" satellite should occur at every instant. Local statistics of link performance can be found by examining preferred links over a 24 hour period. However, satellite systems are usually intended to give large area coverage, and more useful statistics can be generated by examining every site (Christopher, 1980) in the Northern Hemisphere. This can be done systematically by choosing a site near the geophysical North Pole, choosing a best satellite at a given instant, then continuing calculations until 24 hours' performance statistics are gathered. A new site is then chosen and calculations continue until the grand ensemble of performance statistics for the entire Northern Hemisphere is generated. This can be a useful way to compare the relative performance of different satellite constellations.

Table IV-1 shows some statistical performance results for exhaustive searches in the Northern Hemisphere. Scintillation regions of prime concern are interposed on the satellite-ground paths: the high latitude region lies between 15° - 30° geomagnetic colatitude and 90° - 270° inertial longitude (nightside). The equatorial region is centered on 165° inertial longitude and extends ±10° latitude. The computer is encouraged to choose ray paths outside these regions by setting zenith path loss LZ=20 dB in the prime disturbed region. The disturbed region is at a 300 km altitude. At a given instant, a path through the prime region is chosen only if the elevation angle is high enough (e.g., E>30° for the five satellite systems) to offer better link performance than other possibilities. The first two rows show disappointing performance for three synchronous satellites; the columns show average elevation angle $\bar{E}$, probability P (PRIME) of a chosen ray path lying within a disturbed region, average link performance of chosen links, probability (P<30) that elevation angle was less than 30°, followed by P<20, P<10, P<0, and comments. P (PRIME) was less than 3% versus approximately 7.5% if an effort had not been made to avoid the prime regions. The best elevation angle failed to rise to 0° for 1.5% of the total samples. The second row represents 3 phased Molniya and also shows disappointing performance. Synchronous and Molniya satellites cover each others weaknesses, however, and 3 phased Molniya and 2 antipodal synchronous offer the promising performance of row three. P (PRIME) becomes 2.32% of (7.9% with no selection) and probability of elevation angle P<20 is 1.7%. The performance index 9.5 is the best of the table, better even than the 5 Walker satellites on the final row. The 5 Walker satellites have been seriously studied during the past decade as the best worldwide coverage system, but they are seen to be outperformed by the Molniya/synchronous combination in the Northern Hemisphere. Molniya/synchronous satellites can also be shown to be preferred in the absence of ionospheric disturbances.

The advantages of multiple satellites are more notable against the smaller, more severe regions. The high latitude severe region is approximated by colatitudes 17.5° - 27.5°. Table IV-2 shows the five satellite constellations reduce P (PRIME) to less than 1% (vs. 4% with no selection). It is possible to reduce this to negligible levels with the eight satellite system consisting of three

Table IV-1

LARGE SCINTILLATION REGIONS, LZ=20dB, SEARCHES AT 3 HOUR INTERVALS OVER 24 HOURS

Constellation	$\overline{E^0}$	P(PRIME)	$\overline{LOG_{10}(I/S)}$	P<30°	P<20°	P<10°	P<0°	Comment
3 synchronous	36.28°	0.0289	10.3289	0.4183	0.2092	0.067	0.015	20° x 20° search P(PRIME)=0.075 if no selection
3 Molniya	39.34	0.042	10.016	0.3323	0.1710	0.082	0	P(PRIME)=0.075 if no selection
3 Molniya + 2synchronous	50.07	0.02374	9.500	0.1035	0.01718	0.00082	0	Right ascension for MI=.0355 Rad P(PRIME)=.079 if no selection
5 Walker	47 4	0.02934	9.7595	0.18098	0.03647	0.0081	0	Right ascension for WI=0 Rad. P(PRIME)=.0645 if no selection

TABLE IV-2

SATELLITE PERFORMANCE FOR LARGE SCINTILLATION REGIONS, LZ=20dB, SEARCHES AT 3 HOUR INTERVALS OVER 24 HOURS

Constellation	$\overline{E}$	P(PRIME)	$\overline{LOG_{10}(I/S)}$	P<30	P<20	P<10	P<0	Comment
3 synchronous	36.54°	0.00964	10.1933	0.3759	0.1985	0.0871	0.013	20° x 20° search
3 Molniya	39.508°	0.0177	9 9326	0.329.	0.16702	0.07717	0	
3 Molniya + 2synchronous	50.76°	0.00911	9.4439	0.06217	0 009011	0	0	Right ascension of MI=0 Rad. P(PRIME)=0.0432 if no selection
5 Walker	47.52°	0.00743	9.6834	0.1794	0.0343	0.00251	0	Right ascension of WI=0 Rad. P(PRIME)=0.0385 if no selection
3 modified Molniya + 3 Molniya + 2 Synchronous	55.47°	0	9.2540	0.0811	0.00984	0.00241	0	Right Ascension for MI=1.892 Rad; MI^1 is modified Molniya with argument of perigee changed from -90° to -45°.

modified Molniya, three Molniya, and two synchronous satellites. The modified Molniya have argument of perigee changed from -90° to -45° and the time of perigee has been advanced two hours relative to Molniya satellites.

CONCLUSIONS

Early Air Force Geophysical Laboratory maps of worldwide scintillation morphology have been changed to a three-dimensional structure and interposed on ground-satellite paths. Link performance has been estimated as a function of ionospheric piercing point (and its relation to the prime disturbed region), elevation angle and range. When multiple satellites were available, the satellite which offered best link performance was chosen. Link performance statistics were generated for the entire Northern Hemisphere. Two 5 satellite systems gave promising performance in the Northern Hemisphere: three phased Molniya plus two synchronous satellites reduced the probability of a ray path entering the disturbed regions by over two-thirds and five Walker satellites did almost as well. Overall link performance was better on the former system.

Multiple satellite systems were even more capable of reducing the probability of ray paths

entering the small, intense regions. The probability was typically reduced from 4% to less than 1%.

ACKNOWLEDGEMENTS

A. Cohn of MITRE D-91 pointed out a critical error in an early program for the calculation of ionospheric piercing point. Helen King applied unreasonable patience to the first drafts. Dr. B. E. White's interest in the scintillation regions and their relation to satellite communication provided the impetus for this work.

REFERENCES

Aarons, J. Whitney, H. E., and Allen, R. S., "World Wide Morphology of Scintillations," Air Force Cambridge Research Laboratories, Radio Astronomy Branch, Ionospheric Physics Laboratory, January, 1970.

Ballard, A. H., "Rosette Constellations of Earth Satellites," IEEE Transactions on Aerospace and Electronic Systems, AES-16, No. 5, pp. 656-673.

Christopher, P., "Orbit Selection for Optimum Satellite System Performance," Proceedings of National Telecommunications Conference (NTC '80), Houston, Texas, 1980.

Crane, R. K., "Morphology of Ionospheric Scintillation, "Proceedings of National Telecommunications Conference (NTC '74), 1974.

Walker, J. G., "Circular Orbit Patterns Providing Continuous Whole Earth Coverage," Royal Aircraft Establishment, Technical Report 70211 (UDC 629.195:521.6), November, 1970.

INFLUENCE OF SOLAR SECTOR BOUNDARIES ON IONOSPHERIC VARIABILITY

Michael Mendillo and G. Iuliano
Departments of Astronomy and Mathematics
Boston University
Boston, Massachusetts 02215

ABSTRACT

A latitudinal network of total electron content (TEC) observations has been used to assess the contributions to day-to-day F-region variability associated with a wide variety of solar-terrestrial phenomena. Recent discussions of a proposed "solar-weather coupling" effect related to solar wind sector boundary (SB) crossings prompted a systematic search to document any upper atmospheric components of such a linkage. TEC data from several North American sites close to the 70°W meridian (Narssarssuaq, Greenland (L ≈ 5), Goose Bay, Labrador (L 4), Sagamore Hill, Massachusetts (L = 3), and the Kennedy Space Center, Florida (L = 2)) were examined during the years 1973 to 1975. Average daytime (0900-1700 LT) TEC values were compared with a 27-day mean centered on the day of an SB-crossing. The resultant percentage deviations were examined for sector polarity (+ to - versus - to +) effects and for overall latitudinal patterns during a period of ±5 days from the SB-crossings. Average results for the entire data set show peak excursions well within the ±20-25% deviations associated with day-to-day variability. There is, however, a remarkable consistency in the coherence of the small amplitude latitudinal patterns. At middle to high latitudes (L = 3-5), the TEC variations show an ordered transition from enhancements to depletions during the days spanning an SB-crossing. At lower latitudes (L = 2), the opposite pattern of depletions to enhancements occurs. An identical analysis for geomagnetic activity variations (using the planetary index Ap) reveals the well-known pattern of a transition from relatively quiet to relatively disturbed activity as a sector is crossed. A merging of these two results suggests that the manifestation in the F-region of SB-crossings is nothing more than a small-scale version of the average "geomagnetic/ionospheric storm effect" that is well-documented over the L = 2-5 range.

I. INTRODUCTION

In recent years, the study of solar-terrestrial relations has been extended to include the possibility of a so-called "solar-weather" effect (see McCormac and Seliga, 1979 for a recent survey of the field). A key element often mentioned in such discussions deals with the role of solar magnetic sector structure as an ordering parameter for meteorological phenomena (Wilcox et al., 1974, Wilcox et al., 1977, Reiter, 1977, Wilcox, 1979). The main difficulties encountered in these analyses center on questionable repetition rates during different solar epochs and, perhaps more fundamental, that no acceptable theory exists for a solar wind - magnetosphere - ionosphere - troposphere coupling mechanism.

In quite a different area of study, attempts to understand the day-to-day variability of the ionosphere have increasingly turned away from the relatively well-known sources of ionospheric perturbations (e.g., geomagnetic storms (Mendillo 1971, 1973; Mendillo and Klobuchar, 1979)) to more subtle mechanisms (such as lunar tidal effects (Bernhardt, 1978)). Attempts to model and/or to predict sources of ionospheric variability continue to receive wide attention in the ionospheric community (Donnelly, 1979). In a practical sense, few useful results have emerged from these efforts, and thus questions persist as to whether the ±20-25% variability usually associated with ionospheric (F-region) electron densities represents an inherent "noise" of the upper atmosphere, or merely a statement of our ignorance of the relevant aeronomic processes.

In this study, we review the few attempts made to isolate the influence of solar sectors on upper atmospheric fluctuations, and then summarize the results of a new study that uses a multi-site observing network to search for such an influence. The question of a possible solar magnetic sector structure effect on ionospheric variability is, of course, an intriguing question in its own right; in addition, any evidence for an energy transfer mechanism from the solar wind to the upper atmosphere could provide clues for a possible further linkage to the troposphere.

II. PREVIOUS ANALYSES OF SECTOR CROSSING EFFECTS ON THE F-REGION

The standard method for analyzing solar sector boundary (SB) crossing effects is to use a superimposed epoch analysis for the variations of a parameter from some appropriately chosen control curve. Lyon and Bhatnagar (1979) examined SB effects on the maximum diurnal value of total electron content (TEC) and critical frequency of the F-region (foF2) from North American mid-latitude sites during the years 1968-1972. Zevakina and Lavrova (1980) used hourly foF2 data during selected local time intervals from several stations in both hemispheres and from various time periods during the years 1958 to 1968. It is unfortunate that in searching for such an effect, two independent investigators should choose to use such different data bases. It is difficult, therefore, to compare their results in a quantitative way, but nevertheless some common trends do exist:

1) For comparable mid-latitude ionosonde sites in the northern hemisphere, Wallops Island/Ottawa versus Moscow/Murmansk, there is a tendency for daytime foF2 values to decrease as the Earth passes from a solar sector of outward (+) to inward (-) directed magnetic fields.

2) For the - to + SB crossings, Zevakina and Lavrova report the same effect while Lyon and Bhatnagar find the reverse trend (foF2 increases); Lyon and Bhatnagar reported further that the - to + crossings accompanied by ionospheric enhancements were even more pronounced for TEC data, and particularly so during Winter months.

Zevakina and Lavrova described a variety of other correlations they found between foF2 behavior and solar wind configurations. These may be divided into a few broad categories:

1) When the Earth is <u>within</u> a given sector, ΔfoF2 fluctuations tend to be ordered by the radial polarity of the sector: ΔfoF2 is positive during radially outward (+) sectors and negative during radially inward (-) sectors.

2) foF2 exhibits the greatest variability on the day when the Earth passes from one sector to another.

3) For North/South differences in the IMF, the resultant ΔfoF2 patterns are similar to the +/- IMF sector effects, that is, ΔfoF2 tends to be positive during periods of northward IMF and negative during periods of southward IMF.

4) Seasonal and local time analyses show that the influence of the IMF geometry is most pronounced near 12:00 LT and ordered in severity from Equinox (high) to Winter (medium) to Summer (low).

5) The IMF N/S correlations with ΔfoF2 tended to be stronger than the IMF +/- correlations, and finally,

6) The + to - or N to S changes in the IMF produced F-region variations that were similar, but of lower intensities, to the types of perturbations found during large ionospheric storms.

The high degree of specificity found in the Zevakina and Lavrova study prompted us to re-examine the issue using a latitudinal network of TEC observations that had been subjected previously to an exhaustive study of storm-time and quiet-time variations (Mendillo and Klobuchar, 1979). The continuity and reliability of ionospheric TEC data obtained using the satellite radio beacon (Faraday rotation) technique have been demonstrated in many studies (e.g., Mendillo and Klobuchar, 1975), and thus the search for a possible subtle influence of IMF geometry on the ionosphere is well within the capabilities of TEC-based investigations.

III. THE PRESENT STUDY

Routine measurements of the ionospheric total electron content (TEC) are carried out by the Air Force Geophysics Laboratory using a network of observatories close to the 70°W meridian. Data taken from Narssarssuaq (Greenland), Goose Bay (Labrador), Sagamore Hill (Massachusetts) and the Kennedy Space Center (Florida) refer to ionospheric points located at geomagnetic L-shell values of 5, 4, 3 and 2, respectively. For the period 1973-1975, a total of 114 station-months of TEC data were available for analysis from these stations, and 84 solar sector boundary crossing dates were reported (Svalgaard, 1976). For each of the SB events, a TEC control value was established by computing the mean of hourly TEC values over a selected local time range, using data from ±13 days from the date of the SB-crossing. The differences between the daily values (average over the same LT window) and the 27-day mean were then averaged in superimposed epoch fashion over the period ± 5 days from the SB crossing date.

Figure 1 presents the results of such an analysis upon the Goose Bay (L ≈ 4) data base, where six local time windows are defined and the SB's were not separated by +/- or -/+ changes. During

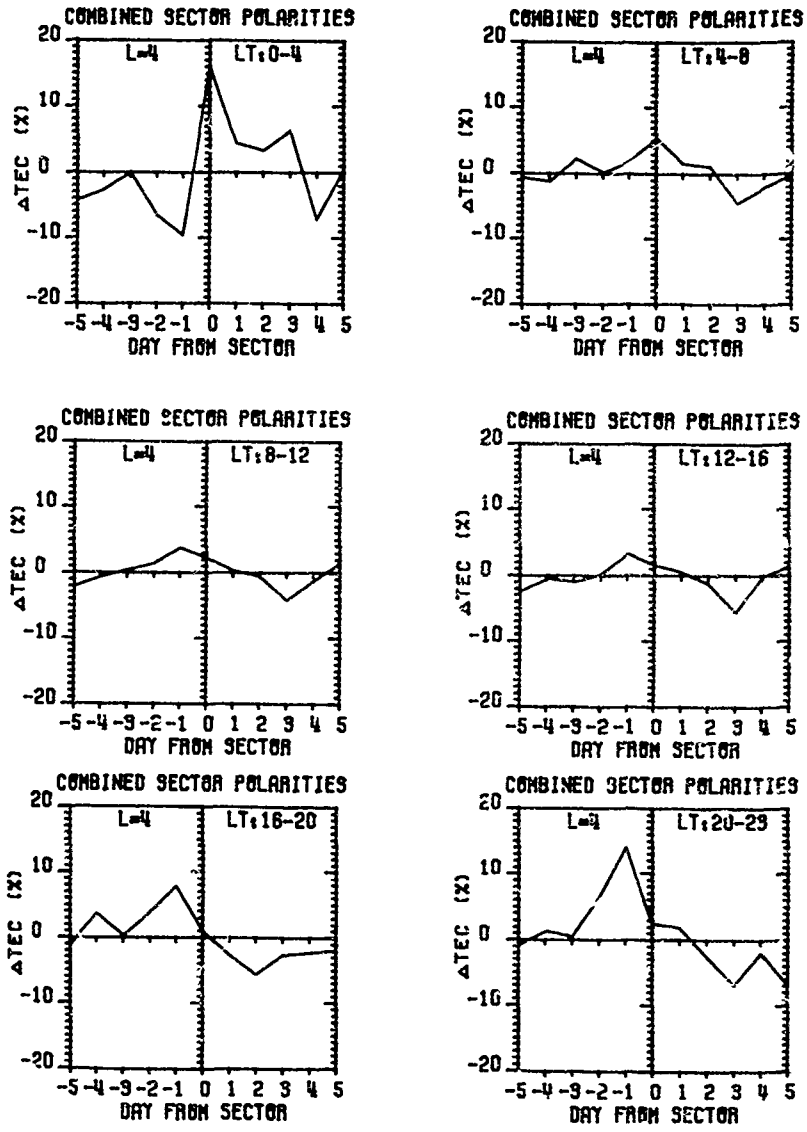


FIGURE 1. Average patterns for ΔTEC (%) during different local time intervals, as ordered by a superimposed epoch analysis for days before and after a solar wind sector boundary crossing. The TEC data were taken from Goose Bay, Labrador, and refer to an L ≈ 4 ionospheric point.

the daytime periods (08-12, 12-16, 16-20 LT) there is a trend for the SB to mark the transition from relatively enhanced to relatively depleted TEC values. While the effect tends to be proportionally greatest during the late afternoon/early evening hours, the overall daytime pattern is a relatively consistent one. At night, the pre-midnight (20-23 LT) and post-midnight (00-04 LT) patterns show a change of phase associated with the SB data.

At the low latitude (L ≈ 2) KSC site (shown in Figure 2), the opposite pattern occurs, i.e., the SB crossing is associated with a transition from low to high TEC values during all local time periods. On the basis of similar analyses for the other stations, it was decided to combine the daytime periods into a single daily value averaged over the 07-19 LT period. The analysis of purely daytime effects seemed an appropriate avenue to follow since mechanisms associated with daytime phenomena are fairly well understood. Moreover, small variations in low nighttime TEC values, while comparable to daytime perturbations on a <u>percentage basis</u>, represent very small effects in the actual total plasma contents.

Average results for the daytime (07-19LT) variations seen over the L = 5-2 range, using all SB types, are given in Figure 3. The ΔTEC (%) curves contained in the lower four panels exhibit a well-defined latitudinal pattern: from middle to high latitudes (L ≈ 3, 4 and 5), the SB's mark a transition from enhanced to depleted TEC values, while at lower latitudes (L ≈ 2), the opposite pattern occurs. The amplitude of the variations decreases over the in-phase domain (L = 5 to 3), but returns to the L ≈ 5 value (~5%) at L ≈ 2. The key to the understanding of this overall morphology is contained in the top panel of Figure 3. A superimposed epoch analysis for the daily planetary index for geomagnetic activity (Ap) reveals that a sector boundary generally signals the transition from periods of relative geomagnetic quiet (so-called Q-days) to days that are more disturbed (D-days). This is a well-known effect that was described by Wilcox et al. (1974). Gross ionospheric behavior during geomagnetic disturbances is also well understood, especially along the L = 2-5 region near the 70°W meridian (Mendillo and Klobuchar, 1975, 1979). These characteristic F-region disturbance features may be summarized as follows:

a) During geomagnetic storms (typified by a SSC followed by one or more days with Ap ≥ 30), ΔTEC is positive over the L = 3-5 range during the first daytime period and negative on the subsequent few days. At L = 2 (and below), the TEC is generally enhanced for several days following an SSC.

b) During periods of geomagnetic quiet, the F-region exhibits a behavior that is opposite from the <u>dominant</u> or <u>main phase behavior</u> seen during disturbed periods. Thus, at L = 3, 4, 5 where the dominant D-day behavior is one of F-region depletions, the quiet time behavior is one or relative enhancements. At lower latitudes, the opposite case of enhancemments on D-days and depletions on Q-days occurs. While seasonal effects may modify the sign and LT dependence of the perturbations, the Q vs. D dichotomy is always preserved with respect to monthly mean conditions.

The general conclusion drawn from these studies was that geomagnetic activity is the single best parameter for ordering the day-to-day variability of the ionosphere. (Mendillo et al., 1979)

Within the context just discussed, the "sector boundary/ionospheric variability effect" described in Figure 3 is simply a restatement of the geomagnetic activity effect. The ΔAp results in Figure 3 show that, on the average, the final day the Earth spends within a sector (Day -1) is generally the one of lowest geomagnetic activity. The ionosphere at L = 5, 4 and 3 is consequently enhanced with respect to average conditions, while at L ≈ 2 the ionosphere is depleted. The transition to disturbances on Day 0 probably includes a positive phase due to an ionospheric storm, and thus the enhancements at L = 5, 4 and 3 are maintained for another day. Eventually, the negative phase (associated with disturbances of the neutral atmosphere) occurs on the subsequent days (with Day +3 showing the maximum effect). At L ≈ 2, where the disturbed behavior is one of prolonged enhancements, the SB day marks a transition from Q-type depletions to D-type enhancements.

The analysis carried out to obtain Figure 3 was repeated after dividing the total data set into cases where the solar sector boundaries marked a transition from outward to inward directed fields (called + to - SB's) and the reverse case (- to + SB's). These results are given in Figure 4. The ΔAp results show that there are no essential differences between the geomagnetic activity transitions that occur as the Earth passes from either + to - or - to + sectors. The ionospheric patterns are also similar. The + to - cases tend to be more pronounced, but not at all of the stations. At some sites, it might appear that the two cases are out of phase on the days leading up to an SB, or that a one day lag exists between the two cases once the SB is crossed. Given the fact that the analysis used daily mean values, it would, however, seem inappropriate to infer 24-hour periodicities in the derived patterns.

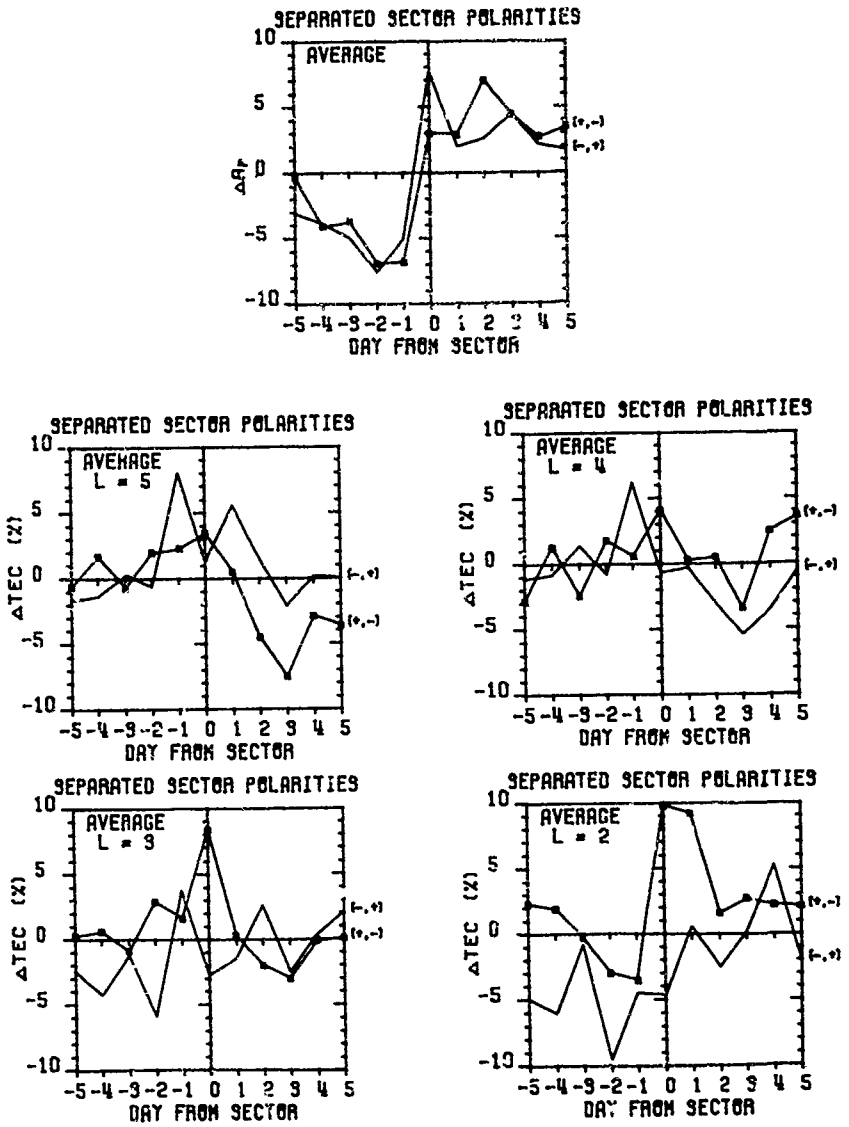


FIGURE 2. Same as Figure 3, with the data set separated by SB polarities.

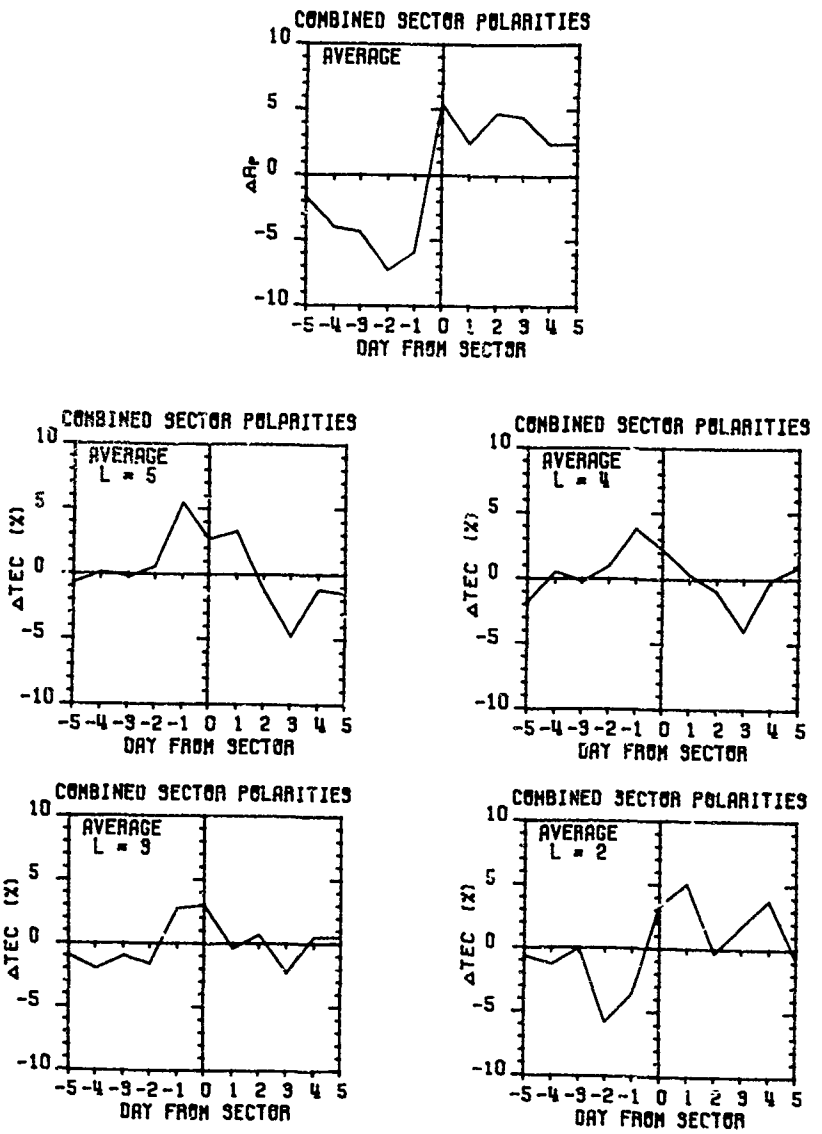


FIGURE 3. Average results for daytime (07-19 LT) TEC/SB effects over the L = 2 to 5 range, together with simultaneous variations in the geomagnetic index Ap.

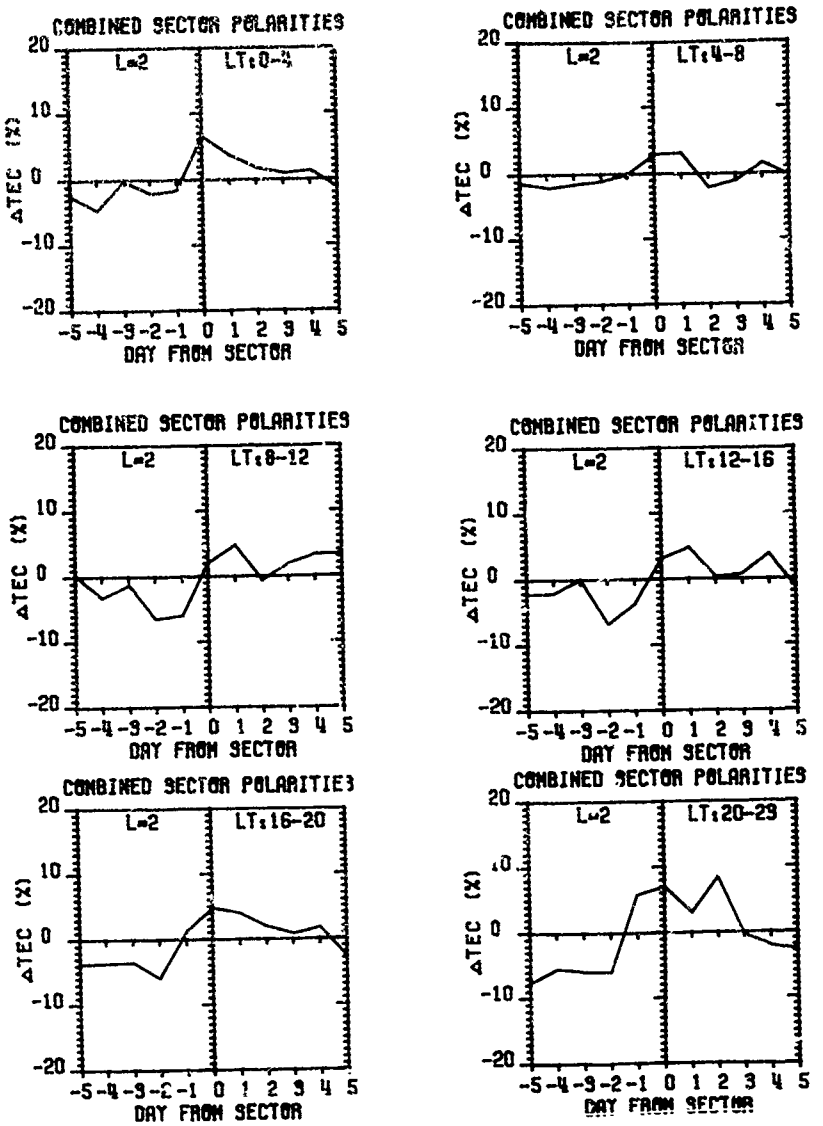


FIGURE 4. Same as Figure 1, using TEC data from the Kennedy Space Center (L=2).

IV. CONCLUSIONS

An analysis of solar sector boundary (SB) related perturbations of the ionosphere showed that as the Earth passes from one sector to another, the ionospheric total electron content (TEC) at middle to high latitudes (L = 3-5) generally undergoes a change from relatively enhanced to relatively depleted TEC values. The opposite trend occurs at low latitudes (L = 2). The variations are small in comparison to normal characterizations of F-region variability (±20-25%) and are associated with the increased geomagnetic activity as the Earth enters into a new sector. The exact nature of the polarity switch at the SB does not play a major role in determining the subsequent ionospheric response. The SB effect on the ionosphere is therefore presented as a scaled-down version of normal ionospheric storm effects.

Acknowledgements

This work was supported, in part, by Contract #77-C-0270 from the Air Force Geophysics Laboratory to Boston University. We are grateful to J. A. Klobuchar for providing the TEC data used in this study.

V. REFERENCES

Bernhardt, P. A., J. Geophys. Res., 83, 5212, 1978.

Donnelly, R. (Editor), Proceedings of the International Solar-Terrestrial Predictions Workshop, Boulder, CO, April, 1979.

Lyon, G. F. and V. P. Bhatnagar, Can. J. Phys., 57, 218, 1979.

McCormac, Billy M. and Thomas A. Seliga (Editors), Solar-Terrestrial Influences on Weather and Climate, D. Reidel Pub. Co., Dordrecht, Holland, 1979.

Mendillo, M., Nature, 234, 1971.

Mendillo, M., Planet. Space Sci., 21, 349, 1978.

Mendillo, M. and J. A. Klobuchar, J. Geophys. Res., 80, 643, 1975.

Mendillo, M. and J. A. Klbuchar, Proceedings of the International Solar-Terrestrial Predictions Workshop (R. Donnelly, Ed.), Boulder, CO, April, 1979.

Mendillo, M. Lynch, F. X. and J. A. Klobuchar, Proceedings of the International Solar-Terrestrial Predictions Workshop (R. Donnelly, Ed.), Boulder, CO, April, 1979.

Reiter, R. J. Atmos. Terr. Phys., 39, 95, 1977.

Svalgaard, L., An Atlas of Interplanetary Sector Structure 1947-1975, Stanford University Institute for Plasma Research Report No. 648, 1976.

Wilcox, J. M., Solar-Terrestrial Influences on Weather and Climate, B. M. McCormac and T. A. Seliga, (Eds.), 149-159, D. Reidel Pub. Co., Dordrecht, Holland, 1979.

Wilcox, J. M., Scherrer, P. H., Svalgaard, L., Roberts, W. O., Olson, R. and R. L. Jenne, J. Atmos. Sci., 31, 581, 1974.

Wilcox, J. M., Svalgaard, L. and P. H. Scherrer, J. Atmos. Sci., 33, 1113, 1976.

Zevakina, R. A. and E. V. Lavrova, Proceedings of the International Solar-Terrestrial Predictions Workshop, R. Donnelly (Ed.), Boulder, CO, April, 1979.

WAVES IN SPACE PLASMAS PROGRAM

R. W. Fredricks and W. W. L. Taylor
Space Sciences Department
TRW Defense and Space Systems Group
One Space Park
Redondo Beach, California 90278

INTRODUCTION

A new era in space research will begin in the early 1980's, made possible by the use of the Shuttle/Spacelab system which will allow instrumentation to be lifted into orbit at a relatively low cost to actively probe the medium remotely and in situ. This instrumentation can be large and heavy, can consume relatively large amounts of power, and can be reflown as many times as required to perform the experimentation. WISP is an excellent example of this type of instrumentation which can be used to perturb the plasma with radio waves to solve some of the outstanding problems in ionospheric, atmospheric, magnetospheric and plasma physics. The initial payload including WISP is tentatively planned for 1986. The transmitters included will cover the frequency range from about 300 Hz to 30 MHz at power levels up to one kilowatt. Receivers will be flown simultaneously on Spacelab and on a recoverable subsatellite (RPDP). An electric dipole antenna of selectable length up to 300 m tip-to-tip is proposed. Ground receivers will also be utilized when feasible. The availability of a payload specialist and the WISP minicomputer to quickly reduce data will provide fast response and human interaction, which is expected to increase the scientific output of the WISP investigation. The WISP equipment will be augmented from flight to flight in an evolutionary way, relying heavily on WISP measurements, theoretical analyses and modeling. WISP is also expected to be used on large, free flying facility-class satellites such as the Solar Terrestrial Observatory and the Science and Applications Space Platform when they are available in the late 1980's.

The WISP investigation is a joint international effort, involving instrumentation to be designed and fabricated by funding from both NASA and the National Research Council of Canada (NRCC). Scientists from both the U.S. and Canada will participate in all phases of the data analysis, interpretation and reporting of results to the scientific community. The NRCC will be responsible to produce the HF transmitter (0.3-30 MHz), HF receivers on the Spacelab and the free-flying subsatellite; NASA will support the production of the VLF transmitter (0.3-300 kHz), receivers, plasma diagnostics, monitoring and control electronics, antenna subsystem, and overall integration of the instrumentation into the Shuttle/Spacelab. NASA will support the pre- and post-launch science and data analysis done by U.S. scientists, while NRCC will support the Canadian participants in WISP in the same areas. The overall effort is directed by the U.S. Principal Investigator (R. W. Fredricks), while the Canadian effort is directed by the Canadian Principal Investigator (H. G. James).

Among the ionospheric and plasma phenomena to be investigated using WISP instrumentation are VLF wave-particle interactions; ELF/VLF propagation; traveling ionospheric disturbances and gravity wave coupling; equatorial plasma bubble phenomena; plasma wave physics such as mode-coupling, dispersion, instabilities (linear and nonlinear); plasma physics of the antenna-plasma interactions. A brief description of the use of WISP instrumentation to gain insights into these important questions will be given subsequently.

VLF WAVE-PARTICLE INTERACTIONS

A principal goal of the proposed WISP investigation is to study the coupling and energy transfer processes between regions in the magnetosphere-ionosphere-atmosphere system (see the reviews by Fredricks, 1975a,b). An important mechanism which transfers energy from the outer magnetosphere to the ionosphere is the interaction between VLF waves, both natural and man-made, and the energetic radiation belt particles. Both incoherent [Kennel and Petschek, 1966; Williams and Lyons, 1974a,b; Taylor and Lyons, 1976] and coherent whistler mode waves [Rosenberg et al., 1971; Helliwell and Katsufrakis, 1974] play important roles in this interaction. While much has been learned about these effects through passive and active investigations, some important questions remain to be answered, e.g., (1) what are the characteristics of the waves that interact with energetic particles? (2) what is the dominant wave growth mechanism in such interactions? (3) what are the particle

distribution functions before and after the interaction? (4) can particle precipitation into the ionosphere be controlled by VLF wave transmissions? WISP should help provide answers to these questions.

Nonlinear Wave Growth and Triggering by Coherent Waves

A discrete, narrowband, variable frequency emission which typically lasts about one second [Helliwell, 1965] is the principal evidence for the existence of a strong nonlinear interaction between coherent VLF waves and energetic particles in the magnetosphere. During these emission-triggering interactions, significant amplification of the triggering wave takes place and large pitch-angle changes can be induced in the resonant electron population. Some quantitative measurements of these wave-growth effects are being made from Siple Station, Antarctica [Helliwell and Katsufrakis, 1974]. Results indicate that a radiated power of a few watts is sufficient to initiate nonlinear wave growth at rates up to about 100 db/sec to amplitudes of about 30 db above the input wave.

Although the active VLF wave injection experiments proposed for the WISP investigation are an outgrowth of the Siple program, these WISP experiments cannot be carried out from the ground. In fact, the WISP instrumentation provides critical capabilities that ground sources cannot supply. First is the ability of the WISP equipment to perturb the plasma into nonlinearities. Second is the ability of the WISP equipment to excite a full range of wave normal directions in the plasma so that all wave modes can be studied. From the ground, the high refractive index of the lower ionosphere limits the initial wave normals in the ionosphere to essentially the vertical direction. Third is the moveable source feature.

Particle precipitation. WISP will be able to study longitudinal and latitudinal variations in wave growth, emission activity, and induced particle precipitation. It has been shown experimentally that naturally-occurring coherent VLF waves can precipitate energetic electrons into the ionosphere, where they produce marked increases in ionization, X-rays and light emissions [Rosenberg et al., 1971; Helliwell et al., 1973; Doolittle et al., 1978]. Furthermore, it has been shown theoretically that energy fluxes as high as 0.1 ergs-cm^{-2}-sec^{-1} (= 10^{-4} watts-cm^{-2}) can be precipitated by moderately intense, coherent VLF waves ($\sim$ 10 milligammas), and that these particles create perturbations in the nighttime ionosphere [Inan et al., 1978].

A remarkable feature of wave-induced particle precipitation is the enormous leverage exerted by the exciting waves on the precipitating particles. For example, under typical conditions the short-term precipitated energy flux induced by waves is estimated to be about 50 db larger than the energy flux in those waves [Inan, 1977]. The corresponding leverage for a Spacelab VLF transmitter should be equally large. Thus, VLF wave energy may act as a valve that regulates the flow of trapped radiation into the ionosphere.

One of the goals of the WISP investigation is to verify that particle precipitation into the ionosphere can be controlled by man-made VLF transmissions. A number of applications may be possible: (1) study of the physics of the auroral atmosphere and lower ionosphere by controlling recombination processes; (2) production of ionospheric conductivity perturbations and subsequent generation of waves with frequencies below about 1 Hz; (3) control of the density in the D-region of the ionospheric waveguide, causing phase and amplitude scintillation of VLF waves in the waveguide; and (4) reduction of energetic electrons from the radiation belts using space-borne transmitters, reducing the radiation hazard for systems in the belts.

To determine whether such control can be established, the WISP equipment will inject VLF waves into the magnetosphere and will measure the resulting electron precipitation at low altitudes directly with a quadrispherical electron detector on the RPDP and indirectly with the low light level TV on Spacelab.

A variety of ground-based measurements will be required to support the WISP investigation. The ground-based VLF instruments exist and are expected to be operated routinely during the 1980's by Stanford University. The two primary stations are the conjugate pair of Roberval, Canada and Siple, Antarctica. These two stations, which are at $L \approx 4$ and have ongoing programs of coordinated geophysical measurements that involve about a dozen U.S. investigators, are expected to be equipped with VLF antennas, receivers, direction finders, photometers, riometers, digital HF sounders, magnetometers and low light level TV systems. Data from these instruments will be available for the WISP investigations.

Augmenting RPDP measurements at low altitudes would be satellite measurements of waves and particle distribution functions on orbits that follow along lines of $\vec{B}$ in the interaction regions (e.g., the DE-A satellite if simultaneous, coordinated measurements can be arranged and made).

VLF/LF WAVE PROPAGATION

In order to understand the physics of the interactions between VLF/LF waves and energetic particles in the magnetosphere, it is necessary first to understand the propagation paths for these waves. It is anticipated that a substantial portion of the VLF electromagnetic energy radiated from the Shuttle will enter the whistler mode, and will propagate either in a ducted or nonducted manner. Ducted signals follow geomagnetic field-aligned density enhancements or depletions and may emerge from the ionosphere and be observed at ground stations [Helliwell, 1965]. Nonducted waves follow more complicated paths; they tend to remain above the lower boundary of the ionosphere, and are not usually observed on the ground [Smith and Angerami, 1968]. The properties of ducted signals are fairly well understood; however, much less is known about nonducted waves, though it has been established that such waves can trigger VLF emissions.

Radio amateurs or students could be involved in some propagation experiments in a meaningful way; for example, to record times of signal acquisition and loss during beacon mode experiments.

Field-Aligned VLF Ducts

One of the outstanding questions of plasma wave propagation in the magnetosphere concerns the origin, structure and dynamics of the field-aligned VLF ducts which guide whistler-mode waves between conjugate points. Using ground transmitters and receivers to locate ducts, and the RPDP for information on wave distribution functions and plasma structure, WISP transmissions may be employed to study the manner of duct excitation, the nature of wave propagation and reflection near the ionospheric ends of ducts, the structure of ducts, and their relationship to other finer structures in the medium.

VLF Wave Propagation Along the Plasmapause

VLF wave propagation just outside the plasmapause at high ($\simeq 0.8$) ratios of wave frequency to equatorial electron gyrofrequency and favorable conditions for wave growth on paths near the plasmapause have been reported by Carpenter [1978]. Theoretical studies show efficient guiding of whistler (VLF) waves along a density gradient such as that at the plasmapause, but do not explain the high frequency propagation seen just outside the plasmapause. Wave injection from WISP on field lines threading the plasmapause, with subsequent reception at ground stations, can explore the nature of this propagation.

TRAVELING IONOSPHERIC DISTURBANCES AND ATMOSPHERIC GRAVITY WAVES

Energy and momentum may be transported over large regions of the atmosphere-ionosphere-magnetosphere system by atmospheric gravity waves. Their ionospheric manifestations are called Traveling Ionospheric Disturbances (TID's) [Hines, 1974]. WISP instrumentation, using remote radio techniques, will measure such ionospheric structures for the purpose of determining their role in transporting energy and momentum, establishing their driving forces and inferring corresponding behavior of neutral particles. WISP instrumentation will provide an almost instantaneous spatial distribution of wave structures. Since ground instrumentation provides temporal distributions at a fixed location, attempts will be made to coordinate measurements by incoherent scatter radars, Doppler sounders, and oblique propagation facilities. In situ data on neutral properties from Spacelab would also be invaluable: e.g., measurements of neutral wind, temperature and composition (as made on the Atmospheric Explorer Satellite [Spencer et al., 1973] which can determine the characteristics of gravity waves at the satellite [Hoegy et al., 1978] and would complement the WISP remote measurement capability of ionized wave structures.

Joule heating by auroral precipitation is one of the most important sources of TID's, and couples the magnetosphere and atmosphere. Winds and pressure perturbations associated with wave fields are produced: e.g., with scale lengths on the order of tens of centimeters (acoustic waves) to thousands of kilometers (gravity waves). The effects are most noticeable during magnetic storms and are observed on a global scale. WISP goals in this area are to: (1) determine the wave characteristics of TID's in the region within a few hundred kilometers of the peak of the F-layer (wavenumbers, direction of travel, frequency, and amplitude as a function of altitude); (2) infer identity and location of sources for the waves; help will be sought from correlative data from ground ionospheric facilities and other geophysical instrumentation (e.g., auroral zone and meteorological measurements); (3) infer the transport of energy and momentum by TID's, and the same for the neutral atmospheric waves driving the ionization; (4) deduce the global distribution of energy and momentum wave fluxes from measurements over various regions; and (5) determine occurrence and structure of smaller-scale disturbances and relate these structures to TID's.

Instrumentation

The WISP MF/HF subsystem will be required for gravity wave and TID observations which will measure a wide range of disturbance structures with scale sizes from meters to thousands of kilometers.

The subsystem is capable of monostatic transmission and reception from the Spacelab and bistatic operation between Spacelab and the RPDP. Operation is in the frequency range from 0.3 to 30 MHz.

The flexible MF/HF subsystem will be capable of coherent detection, and will measure time delay, phase, amplitude and Doppler shift of received signals. Ionization structures and motions of TID's will be measured, both along the orbit and remote from Spacelab. The subsystem will operate in a number of preprogrammed modes, including a survey mode which will search for natural phenomena of interest. This survey mode may be followed by a mode to investigate some feature in greater detail by providing information in a more magnified form. Other modes may change the frequency range of operation, pulse width, pulse repetition frequency and/or the power radiated. Modes may select Doppler shift information or operate at a single frequency to monitor continuously the strength and variation of the received signals. Several types of data presentations are possible: e.g., displays of propagation time vs frequency, or Doppler vs time, Doppler vs range, and range vs time at selected frequencies. Two complementary methods for improving range resolution are available. In the first, pulse width may be reduced. In the second, phase difference of the echoes at two frequencies separated by a small amount ($\sim$ 10 kHz) is determined and provides a vernier to the standard time delay method [Bibl and Reinisch, 1978] with range resolutions from 0.7 to 20 km.

This range-measuring capability will provide data on vertical structure of TID's as well as horizontal variation along the orbit. In contrast with earlier remote sounding systems, WISP will provide rapid frequency sweeps over limited frequency bands, enabling considerably improved horizontal resolution. The system will determine the wave vector of the TID [Dyson, 1978]. Frequency may be determined, as well as all other wave properties, from TID gravity wave theory: these in turn may be corroborated with ground measurements, if available.

The Doppler-measuring capability of the MF/HF subsystem will represent a unique opportunity for ionospheric measurements from space. It provides a means for distinguishing traces from different regions at the same virtual distance (echo delay) by supplying information on the angle of arrival. Since Doppler for a given frequency varies as the cosine of the angle between the wave normal and vehicle velocity, the Doppler is a measure of the angle of arrival. The minimum discernible frequency shift depends on oscillator stability. For WISP the stability will be 10^{-7}, corresponding to an angular resolution of $\sim 4°$ for propagation along the orbit and $\sim 0.1°$ for vertical incidence. Though measurements are rapid, some may require up to $\sim$ 6 seconds. Phase data as a function of time, ideally will be processed by a fast Fourier transform to obtain the Doppler spectrum. Ambiguities in the Doppler spectrum and ranges are avoided, when required, by using two different pulse repetition frequencies sequentially.

Doppler measurements will also be made rapidly as well by utilizing, when required, a small number of pulses (2-5) at a fixed frequency. Average Doppler is determined, and angles of arrival are determined by observing the average Doppler for discrete echoes as frequency is changed, permitting rapid measurement of Doppler over a wide range of frequencies [Dyson, 1978]. TID structure will be obtained by combining angle of arrival and time delay data using true height analysis and ray tracing. Doppler techniques also allow vertical returns to be distinguished from field-aligned returns and discrete echoes to be discerned in the presence of Spread-F.

Ground-based measurements of TID's and atmospheric gravity waves will be made in Eastern Australia by LaTrobe University as part of the WISP investigation to provide two-hemisphere coverage. These measurements will follow TID's over a limited range and will, for example, help to determine wave vectors and periods of TID's and should aid in following atmospheric gravity waves as they propagate from the southern auroral region toward the equator. Simultaneous use of moving (WISP) and stationary (ground) systems helps reduce ambiguities in calculations of density irregularity spectra [Fredricks and Coroniti, 1976].

EQUATORIAL BUBBLES

There has been a surge of interest in the equatorial ionosphere in the last few years resulting from new observational and theoretical results indicating that spectacular, large-scale, dynamic plasma processes are taking place. Of particular importance are recently-discovered large (up to hundreds of kilometers) regions of depleted ionospheric plasma in the vicinity of the magnetic equator which bubble through from the bottomside to the topside of the ionosphere (see review by Basu and Kelley, 1977, and references there). These bubbles may be the cause of the unexpectedly severe scintillations observed on communication links with geosynchronous satellites even at GigaHertz frequencies [Taur, 1973; Basu and Kelley, 1971].

Theoretical and numerical simulation studies of bubbles which follow the nonlinear evolution of the collisional Rayleigh-Taylor instability, reveal that a small perturbation on the bottomside ionosphere will grow in amplitude and rise into the topside [Ossakow et al., 1978]. This process is very dependent upon the steepness of the bottomside electron density gradient and on the altitude of the F-peak. Other instabilities have been suggested to explain the wide range of observed irregularities

with bubble scale size; it may be necessary to invoke several processes to completely model bubbles [Hudson and Kennel, 1975]. Besides determining the relative importance of various instability processes, the following basic questions must be answered: (1) what are the physical dimensions of the bubbles? (2) what is the relationship between equatorial bubbles and field-aligned ducts? (3) is topside Spread-F always associated with bubbles? (4) which stage of the instability process is the dominant contributor to the scintillation of communication signals? (5) what is the triggering mechanism for the plasma instabilities? The WISP investigation is expected to answer questions (1), (2), and (3). WISP results, combined with the results of other experimental and theoretical programs, should lead to a resolution of questions (4) and (5) as well.

Instrumentation

The WISP investigation of equatorial bubbles will make use of the remote measurement capability of the MF/HF subsystem and *in situ* plasma measurements from the subsatellite. The potential of the sounding technique has been illustrated by Dyson and Benson [1978] using ISIS-1 post-sunset perigee data near the magnetic equator. The WISP instrumentation will be able to capitalize uniquely on this potential by utilizing its capability for flexible control, bi-static operation, and angle of arrival information from Doppler measurements. In addition, the Spacelab orbit will traverse the region of greatest interest, i.e., within a few hundred kilometers above and below the altitude of the F peak at the magnetic equator. By filtering based on Doppler shift, ionograms free of Spread-F will be produced in the turbulent regions near equatorial bubbles for the first time.

The WISP instrumentation, with its probing capability utilizing plasma waves, provides a practical method of determining the variation of bubble structure along $\vec{B}$. The nature of these variations must be known in order to understand the evolution of bubbles and to test the applicability of current two-dimensional theoretical models which assume that bubbles are of infinite extent along $\vec{B}$. In addition, the information will provide the data base for constructing three-dimensional models of bubbles. During transits of equatorial bubbles above the F-peak field aligned echo traces can be used to determine the electron density distribution along $\vec{B}$ from the Orbiter to the F-peak, and changes in composition along $\vec{B}$ can be inferred from changes in the calculated scale height [Dyson and Benson, 1978]. These inferred remote composition changes will be compared with the *in situ* measurements made by the ion mass spectrometer on the RPDP. The importance of *in situ* measurements during transits of equatorial bubbles has been demonstrated by McClure et al. [1977]. When the bubble extends into the conjugate hemisphere, echoes are returned from that hemisphere. Thus the total length of the bubble along $\vec{B}$ can be estimated, and asymmetries in structure with respect to the magnetic equator can be detected [Dyson and Benson, 1978].

RELATED PLASMA WAVE PHYSICS

Plasma Instabilities and Nonlinear Effects

Plasma instabilities are believed to be of fundamental importance in atmosphere-ionosphere-magnetosphere energy coupling processes. These instabilities excite a number of plasma waves which can grow in amplitude to the point where nonlinear effects become important. In some cases, naturally-occurring magnetospheric plasma wave phenomena (in the kilohertz frequency range) can be investigated by stimulating similar phenomena in the ionosphere where they occur in the megahertz range.

An important example of a stimulated ionospheric phenomenon that has possible relevance to naturally-occurring magnetospheric phenomena is the diffuse resonance stimulated by the ISIS topside sounders near $3/2\ f_H$, where f_H is the electron cyclotron frequency. In contrast to the stimulated electrostatic waves which are observed to persist in the neighborhood of the satellite when the wave frequency corresponds to ambient plasma resonant conditions, the diffuse resonance corresponds to nonresonant plasma waves. A nonlinear mechanism is required to explain their long time persistence (see the review by Benson, 1977).

Dispersion Relations

A fundamental property of plasma waves is that they exhibit dispersion. That is, their phase velocity varies with frequency. Many different electromagnetic (e.g., whistler) and electrostatic (e.g., electrostatic ion cyclotron) waves have been observed in space plasmas, and each mode has its own dispersion relation.

Dispersion measurements will use a WISP transmitter on Spacelab and a correlation receiver on the RPDP. By also transmitting a reference signal to the RPDP on a data link, the phase of the received signal can be determined. This technique is equivalent to measuring the phase velocity, and when carried out over a range of frequencies yields the dispersion relation directly. It is highly accurate, with 0.1 percent accuracies expected. The flexible control capability of WISP Spacelab systems and the WISP systems on the moveable RPDP will result in significant advances in understanding these problems.

Antenna Properties

An important aspect of the WISP investigation related to local plasma wave physics is the study of the properties of antennas immersed in the magnetoplasma. In this compressible and anisotropic medium the electromagnetic fields of the antenna will have a marked influence on the free charges of the plasma, which in turn drastically modify the radiated fields.

Many theoretical treatments of antennas immersed in plasmas have been made. However, in only a few restricted cases have experimental studies been made in space, and possible laboratory work is limited due to wall effects and scaling difficulties. However, it must be noted that due to the very complex nature of the problem, all past theoretical treatments have been restricted by simplifying assumptions whose justification is not always clear.

In addition to this general scientific significance, the antenna impedance measurements in the first flight will be of significant importance for the operation of the transmitters during the later flights, since the antennas will have to be tuned in most cases in order to couple efficiently the electrical energy into plasma wave energy.

During the flights, the WISP instrumentation will measure the voltage and current waveforms and their relative phase at the antenna terminals, thereby determining the antenna impedance. The parameters that the impedance should depend upon (frequency and level of exciting signal, the strength and direction of the ambient magnetic field, and the plasma composition and density) will also be measured by the WISP instruments.

ACKNOWLEDGMENTS

The WISP investigation as proposed to NASA is the result of the combined efforts of the investigator team, each member of which has made indispensable contributions to both science objectives and experiment definition. It is with great pleasure that the authors acknowledge these individuals, who are:
R. F. Benson, NASA Goddard Space Flight Center
Wynne Calvert, University of Iowa
Peter Dyson, La Trobe University, Australia
R. W. Fredricks, TRW, Inc. (Principal Investigator for NASA)
Owen Garriott, NASA Johnson Space Center
Stanley Gross, Polytechnic Institute of New York
Mario Grossi, Harvard-Smithsonian Astrophysical Observatory
R. A. Helliwell, Stanford University
H. Gordon James, Communications Research Centre, Ottawa (Canadian Principal Investigator)
J. P. Katsufrakis, Stanford University
A. G. McNamara, Herzberg Institute of Astrophysics, NRCC, Canada
D. R. Moordroft, University of Western Ontario, Canada
D. B. Muldrew, Communications Research Centre, Ottawa, Canada
David Reasoner, NASA Marshall Space Flight Center
Stanley Shawhan, University of Iowa
W. W. L. Taylor, TRW, Inc.

The authors gratefully acknowledge the support of the WISP investigation in the U. S. under contract NAS8-33971 between NASA Marshall Space Flight Center and TRW, Inc. The constant support and encouragement of Dr. E. R. Schmerling of NASA Headquarters throughout the years is gratefully acknowledged.

REFERENCES

Basu S. and M. C. Kelley (1977), J. Atmos. Terr Phys., 39, 1229-1242.
Benson, R. F. (1977), Radio Sci., 12, 861-878.
Bibl, K. and B. W. Reinisch (1978), Radio Sci., 13, 519-530.
Carpenter, D. L. (1978), J. Geophys. Res., 83, 45-57.
Doolittle, J. A., W. C. Armstrong, R. A. Helliwell, S. B. Mende, J. P. Katsufrakis and D. L. Carpenter (1978), EOS, 59, 362.
Dyson, P. L. (1978), NASA TM 79657 (Goddard Space Flight Center, September 1978).
Dyson, P. L. and R. F. Benson (1978), Geophys. Res. Lett., 5, 795-798.
Fredricks, R. W. (1975a), Space Sci. Rev., 17, 449-480.
Fredricks, R. W. (1975b), in The Magnetospheres of the Earth and Jupiter, V. Formisano, ed., pp. 113-123, Reidel Publ. Co., Dordrecht, Holland.
Fredricks, R. W. and F. V. Coroniti (1976), J. Geophys. Res., 81, 5591-5595.
Helliwell, R. A. (1965), Whistlers and Related Ionospheric Phenomena, Stanford Univ. Press.

Helliwell, R. A. and J. P. Katsufrakis (1974), J. Geophys. Res., 79, 2511-2518.
Helliwell, R. A., J. P. Katsufrakis and M. L. Trimpe (1973), J. Geophys. Res., 70, 177-183.
Hines, C. O. (1974), Geophys. Monograph 18, Amer. Geophys. Union, Washington, D.C.
Hoegy, W. R., P. L. Dyson, L. E. Wharton and N. W. Spencer (1978), NASA TM 79670, Goddard Space Flight Center, November 1978.
Hudson, M. K. and C. F. Kennel (1975), J. Geophys. Res., 80, 4581.4590.
Inan, U. S. (1977), Radioscience Lab. Tech. Rpt. No. 3414-3, Stanford Univ., August 1977.
Inan. U. S., T. F. Bell and R. A. Helliwell (1978), J. Geophys. Res., 83, 3235-3253.
Kennel, C. F. and H. E. Petschek (1966), J. Geophys. Res., 71, 1-28.
McClure, J. P., W. P. Hansen and J. H. Hoffman (1977), J. Geophys. Res., 82, 2650-2656.
Ossakow, S. L., S. T. Zalesak, B. F. McDonald and P. K. Chaturvedi (1978), Paper 2-7 in Proc. of 1978 Symposium on the Effect of the Ionosphere on Space and Terrestrial Systems, sponsored by U. S. Naval Res. Lab., Washington, D. C., June 24-26, 1978.
Rosenberg, T. J., R. A. Helliwell and J. P. Katsufrakis (1971), J. Geophys. Res., 76, 8445-8452.
Spencer, N. W., H. B. Niemann and G. R. Carrignan (1973), Radio Sci., 8, 287-296.
Smith, R. L. and J. J. Angerami (1968), J. Geophys. Res., 73, 1-20.
Taur, R. (1973), Comsat Tech. Rev., 3, 145-163.
Taylor, W. W. L. and L. R. Lyons (1976), J. Geophys. Res., 81, 6177-6183.
Williams, D. J. and L. R. Lyons (1974a), J. Geophys. Res., 79, 4195-4204.
Williams D. J. and L. R. Lyons (1974b), J. Geophys. Res., 79, 4791-4798.

DARN: A RADAR NETWORK FOR STUDYING THE LARGE-SCALE STRUCTURE AND DYNAMICS OF IONOSPHERIC ELECTRIC FIELDS

R. A. Greenwald
Applied Physics Laboratory
The Johns Hopkins University
Laurel, Maryland 20810

Ionospheric electric fields with magnitudes of several tens of millivolts per meter have a profound effect on ground-based radar systems since they are capable of producing E-region electron density irregularities via several plasma streaming instabilities. These irregularities often have Doppler velocities of the order of the acoustic speed and thus they represent a form of clutter that cannot readily be eliminated from radar systems without also eliminating the backscattered signals from more desired targets. The manner in which the ionospheric electric field exciting these irregularities is coupled to the magnetospheric and atmospheric electric fields is often not well understood. This is particularly true for impulsive or oscillatory electric fields where inductive effects are important. However, it is known that enhanced ionospheric electric fields have a profound effect on the ionosphere through Joule heating, changes in the electron density distribution and as a driving force for high altitude neutral winds. It is also known that enhanced ionospheric electric fields, which are normally confined to latitudes poleward of 60° geomagnetic extend to lower latitudes during periods of strong geomagnetic disturbance. This is paticularly true near the onset of the disturbance.

For all of its apparent importance there presently exists no global or large scale system for monitoring the electric field such as the worldwide magnetometer network provides for magnetic fields and the DMSP satellites provide for auroral luminosity. This deficiency can be somewhat alleviated through the establishment of a Dual Auroral Radar Network (DARN) across North America. The network, comprised of 10-15 medium-powered VHF radars, would determine the two-dimensional dynamics of the ionospheric electric field in the latitude range 60°-67° geomagnetic and the longitude range 70°-160° W. DARN utilizes backscatter from the same ionospheric clutter that disturbs surveillance radars to make these measurements. Since DARN is a real-time system that would be connected by telephone line to a host computer, it will not only provide an instantaneous picture of regions of ionospheric clutter over all of Central Canada, but also, it will be an early warning system for increased ionospheric disturbance. Moreover, the data from this system will represent a set of reference measurements for studies of coupling between magnetospheric-ionospheric-atmospheric electric fields and be an important input for dynamic ionospheric models.

INTRODUCTION

At the present time the auroral imager on the DMSP satellites is the only instrument capable of providing large scale two-dimensional images of processes occurring in the high latitude ionosphere. Every 51 minutes one of these satellites produces a broad-band optical image of one of the Earth's high latitude regions. The longitudinal width of these images is approximately 3000 km in extent.

In Figure 1 we see a composite photo of auroral activity over North America obtained from three consecutive passes of the DMSP satellite. One can see that the aurora is present over most of Canada and extends into some of the north central states during the second pass of this series. Typically the aurora is confined to an oval-shaped region that is confined between 60° and 75° geomagnetic latitude.

When viewed from a spacecraft located at a slightly greater distance the auroral belt might appear as shown in Figure 2. In this example we are viewing the nightside auroral zone. A spacecraft located at this greater altitude would be able to image the aurora for longer periods of

FIGURE 1 Composite DMSP satellite photograph of North America continent illustrating auroral activity over Canada.

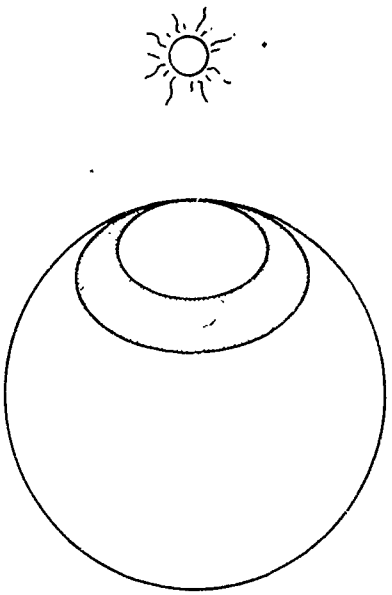


FIGURE 2 Sketch of the appearance of the auroral oval as viewed from a high altitude spacecraft.

time and thereby provide new information on its dynamics. Such high altitude imagers have already been flown on Japanese satellites and they will be flown on United States spacecraft in the near future.

One should keep in mind, however, that the optical aurora is only one of several phenomena occurring on high latitude field lines. A more complete version of Figure 2 might appear as shown in Figure 3. Here, one can see large-scale sheets of field-aligned currents entering and emerging from the auroral regions and connecting to regions of the Earth's magnetosphere that are still not completely known. Moreover, electric fields are observed in the ionosphere and in the distant magnetosphere. These electric fields are self-consistently related to convective plasma motions in the magnetosphere and in the ionosphere. In the lower reaches of the ionosphere - the E-region - these motions manifest themselves as currents since the convective ion drifts are hindered by collisions with the neutral atmosphere. These collisions also cause energy dissipation in the ionospheric via the process of Joule heating.

All of these phenomena, auroral precipitation, field-aligned and ionospheric currents, electric fields, and plasma convection, interact through a self consistent set of equations. At the present time we understand the average behavior of the various elements (e.g., Meng, 1978; Potemra et al., 1980; Rostoker, 1980; Banks and Doupnik, 1975). Unfortunately, the high latitude ionospheric and magnetospheric environments are extremely dynamic and gross departures from the mean can occur on time scales of a few minutes. The changes are caused ultimately by variations in the magnetospheric plasma distribution. In the ionosphere they are manifested by changes in particle precipitation, field-aligned currents, electron density gradients and electric fields.

Sudden variations in high latitude ionospheric conditions can have particularly detrimental effects on radiowave systems. Propagation paths may be altered, absorption may be increased, E and F region clutter may be excited and scintillations may occur on satellite-ground communication links. During geomagnetic disturbances, high latitude effects also expand to lower latitudes and affect a greater portion of the Earth's ionosphere. Responses to high latitude disturbances have even been noted in the equatorial ionosphere (e.g., Gonzoles et al., 1979).

In order to understand the dynamics of these large scale processes, techniques must be developed whereby it is possible to image the large scale patterns of the ionospheric electric fields, field aligned and horizontal currents to an extent currently provided by auroral imagers. Presently, satellite measurements of electric fields, plasma drifts and field-aligned currents are

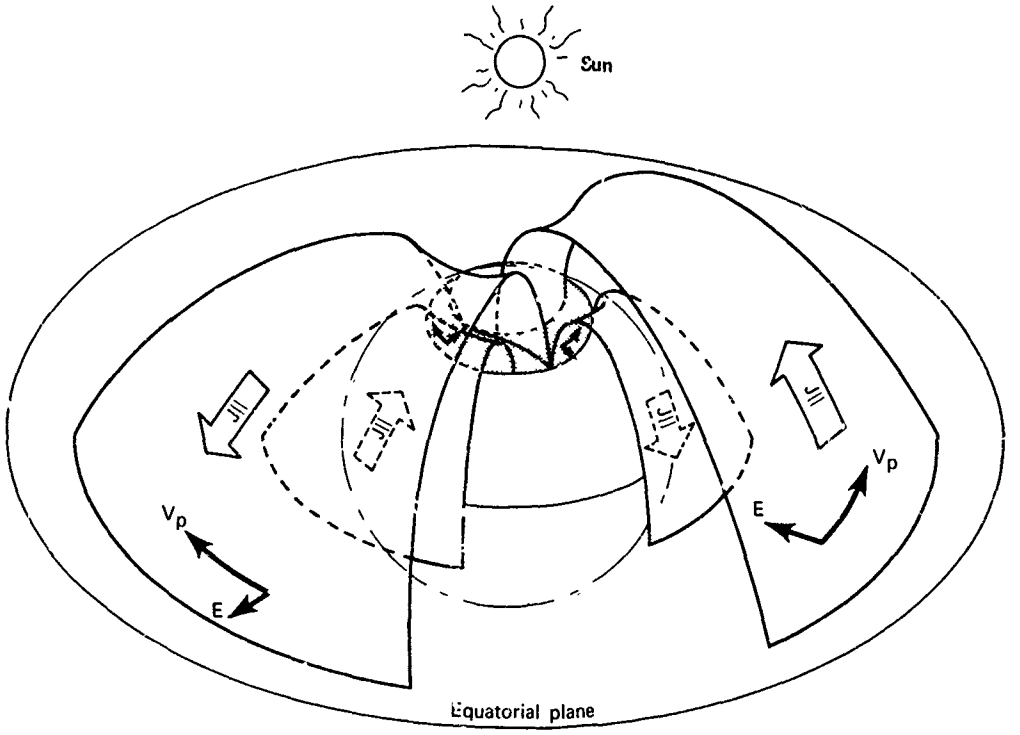


FIGURE 3 Illustration of the interrelationship of the auroral oval, magnetic field-aligned currents flowing between the ionosphere and magnetosphere, and ionospheric and magnetospheric electric fields and plasma flow.

made only along the spacecraft trajectory. Moreover, measurements of a given latitude and local time sector are often separated temporally by the orbital period of the spacecraft. Ground-based incoherent scatter radars are able to determine these quantities over much greater spatial areas (e.g., Banks and Doupnik, 1975). However, the temporal resolution of incoherent scatter measurements is typically of the order of minutes per single vector velocity measurement. The relatively long measurement time is determined primarily by the low cross section of the incoherent scatter process.

Another technique for measuring ionospheric drift velocities and electric fields has been developed recently by the German Max-Planck Institute for Aeronomy (Greenwald et al., 1978). It utilizes VHF radar backscatter from electron density irregularities produced by plasma streaming instabilities in the E-region of the high latitude ionosphere (see Greenwald, 1980 and references therein). The Max-Planck facility, known as STARE, is capable of measuring irregularity drift velocities throughout a 230,000 km^2 region over northern Scandinavia (see Figure 4). The measurements are made with 20 km x 20 km spatial resolution and 20-60 second temporal resolution.

Studies by Ecklund et al. (1977) and Cahill et al. (1978) have shown that the E-region irregularity drift velocity is nearly equal to the ionospheric F-region plasma drift velocity and, hence, that it can be used to determine the ionospheric electric field. STARE has been the first instrument that has really imaged the structure and dynamics of the ionospheric electric field and drift velocity over an appreciable spatial region.

Currently, a second facility known as SABRE is being constructed by British scientists and scientists at the Max-Planck Institute for Aeronomy. The viewing area of this instrument will be located to the southwest of the STARE viewing area.

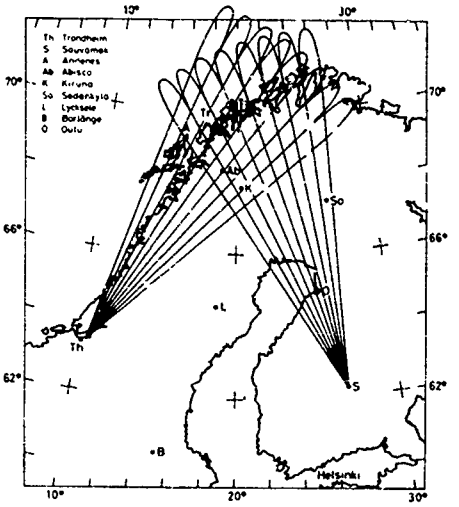


FIGURE 4 Map of northern Scandinavia showing the common viewing area of the STARE radars. The region of overlap covers 230,000 km².

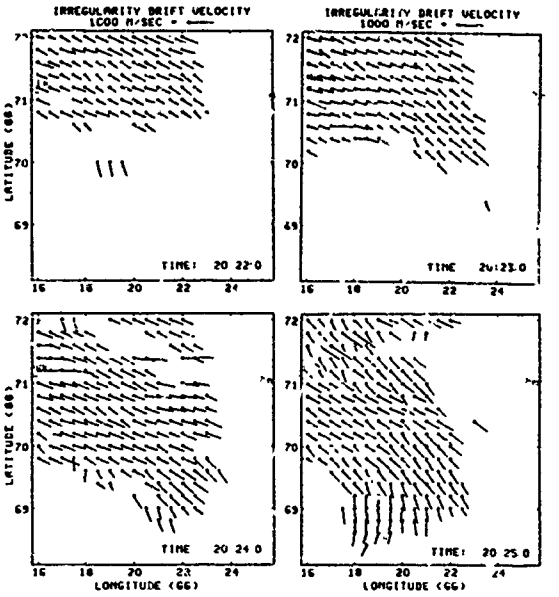


FIGURE 6 Temporal variation in the ionospheric plasma drift in association with the passage of a westward travelling auroral surge.

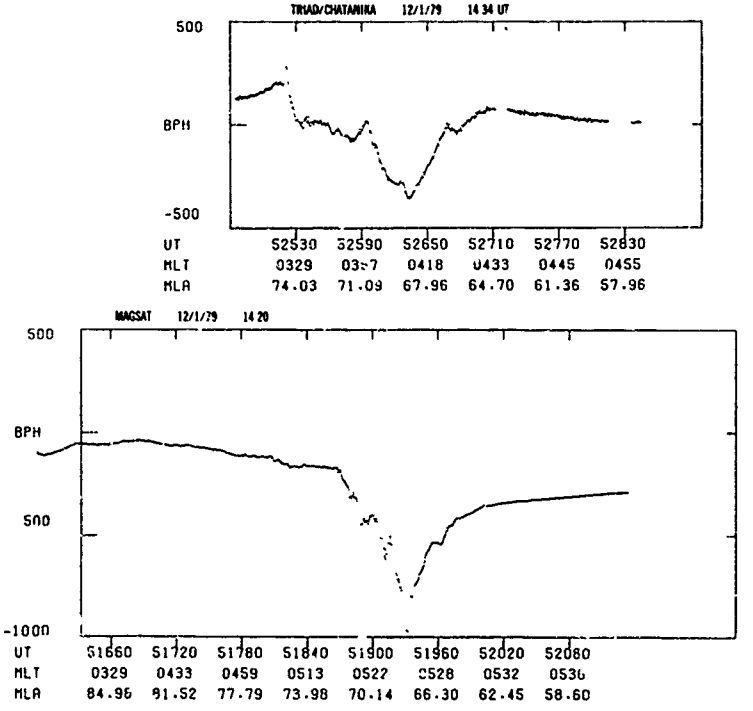


FIGURE 5 Comparison of transverse magnetic fields due to field-aligned currents as observed with the MAGSAT and TRIAD spacecraft.

While STARE and SABRE are beginning to provide imagery of large scale electric field dynamics, their concept may be extended considerably on the North American continent. Due to the extensive east-west land mass, it would be possible to construct an east-west chain of 10-15 simplified STARE-type radars extending from Maine to Alaska. Such a network would provide the following useful information:

(1) Continuous coverage of the location of E-region auroral clutter over most of the North American auroral zone,

(2) Aid in determining the location of F-region clutter and regions in which scintillation would be expected to occur on satellite-ground communication links, and

(3) Large area imagery of the dynamics of auroral zone electric fields and plasma drifts.

This information will not only be useful for monitoring the effects of naturally-occurring phenomena on high latitude radiowave systems, but also it will provide considerable new information on the dynamics of high latitude electric fields and plasma flow. Such information will be particularly useful in developing reliable dynamic models of the high latitude ionosphere.

STRUCTURE AND DYNAMICS OF PARTICLES AND FIELDS IN THE HIGH LATITUDE IONOSPHERE

Before describing the North American radar network that may be developed, let us consider several examples illustrating the dynamic nature and large scale spatial structure associated with particles and fields in the high latitude region.

The optical aurora shown in Figure 1 is representative of regions into which magnetospheric particles are precipitating. It is also to some degree indicative of regions in which the ionospheric conductivity is enhanced relative to their surroundings. As has been noted, the aurora and hence the precipitating particles and density enhancements extend further equatorward during the second pass in this series. This is one example of the dynamic variations in the auroral oval on 100 minute time scales. People who have witnessed auroral breakups, have seen similar movements on minute time scales. One can also see considerable structure in the aurora on several hundred kilometer scale sizes.

A second example of structure and dynamics can be seen in the comparison of magnetic perturbations due to field-aligned currents as shown in Figure 5. Here, the TRIAD and MAGSAT low altitude satellites were in nearly co-planar orbits with TRIAD lagging MAGSAT by approximately 12 minutes. One can see that the equatorward region of upward field-aligned currents (decreasing negative magnetic perturbation with decreasing latitude) is quite similar on the two passes, however, the poleward region of downward field-aligned currents (opposite sense of magnetic field variation) is quite different. Whether this difference is due to spatial structure (the orbits are shifted by one hour in local time) or temporal variations is not known. Observations similar to these have led to suggestions by Potemra (Private Communication, 1981) that several existing small satellites equipped with magnetometers be launched into co-planar orbit.

As a final example let us consider measurements of structure and dynamics in the ionospheric plasma flow and electric field as obtained with the STARE radar system in northern Scandinavia. One four minute sequence of irregularity drift velocity patterns observed with the STARE radars is shown in Figure 6. Each small dot represents the location of a measurement and the length and direction of the line extending from each dot represents the magnitude and direction of the local drift velocity. By rotating the vectors 90° clockwise and equating 50 mV/m with a drift of 1000 m/s, one can obtain the local electric field strength. Drift velocity vectors are missing in the eastern portion of the field-of-view since it was not being examined at this time. Vectors are absent in the equatorward portion of the field-of-view because the streaming velocity between the electrons and ions was insufficient to excite the instability that produces the irregularities.

The four panels illustrate the drift velocity variations that occur on the trailing edge of an auroral phenomenon known as a westward travelling surge. Other studies using the STARE radars have identified the electric field variations occurring on the leading edge of such a surge (Inhester et al., 1981). It should also be noted that intense radar auroral clutter was observed in the region of equatorward drift at 2025 UT.

Other studies with the STARE radar system have provided numerous examples of impulsive drift velocity and electric field variations such as these or oscillatory variations associated with hydromagnetic waves incident on the E-region (e. g., Nielsen and Greenwald, 1978, 1979; Walker et al., 1979).

THE DUAL AURORAL RADAR NETWORK (DARN)

The STARE radar system is capable of imaging plasma drift velocities and electric fields over a region covering slightly more than 4° in latitude and 10° in longitude. While this begins to provide some large scale view of electric field dynamics, one could improve the spatial coverage considerably by developing a chain of radars across an extensive land mass such as the North American continent. In this case the coverage might be 7° in latitude and 90° in longitude.

One tentative concept for a North American radar network is shown in Figure 7. In this diagram 12 radar sites are seen spanning the northern tier of the conterminus United States, western Canada and Alaska. In practice one might have 10-15 sites. The viewing area of this system would be essentially coincident with the region of aurora shown in Figure 1. At each site transmission and reception will be achieved with a single bi-directional array of electronically-phased Yagi antennas. Phasing will be limited to directing the array lobe approximately 20° east and west of the magnetic meridian as shown in Figure 8. The two-way azimuthal beamwidth of this lobe will be approximately 6°. On-line processing equipment at each site will be designed to measure the intensity and Doppler shift of the signals scattered along each of these lobes with a range resolution of 30 km. The shaded area in Figure 8 represents the expected resolution cell of the DARN radars at a range of 1000 km.

It is envisioned that the DARN sites will be separated longitudinally by approximately 500 km. Thus, beams from neighboring sites will overlap to form the X-patterns shown in Figure 8. In the range interval extending from 400 km (assumed nearest range having good aspect angle sensitivity) to 1200 km (120 km altitude horizon) the two legs of the X are never separated by more than 250 km and their average separation is only 125 km. The latter distance is approximately equal to the diameter of the circle at 200 km altitude that is described by the beam of the Chatanika incoherent scatter radar when it performs a circular or three-position scan that is directed at 70° elevation angle. As is done in this common incoherent scatter operating mode, spatial uniformity of the electron drift velocity and ionospheric electric field over these dimensions will be assumed. Although STARE measurements have shown that this assumption is not always valid, it is often a very good approximation. For the DARN network the assumption is particularly good, since it only involves longitudinal variations in the electric field.

One can see in Figure 8 that neighboring X-regions are separated by a gap of 250 km and that they have a latitudinal extent of approximately 800 km (7°). The latitudinal resolution as defined by the assumed 30 km long radar pulse length is somewhat better than 0.3°. One can also see in the figure that, for any given latitude, the same component of the electron drift velocity is measured every 500 km in longitude. Thus, one could undoubtedly develop better mathematical models for the determination of longitudinal variations than the simple uniform field assumption described above.

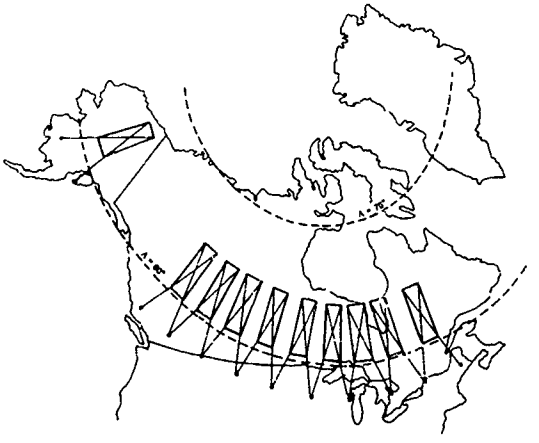


FIGURE 7 View of northern United States and Canada showing potential coverage of 12 station DARN network. The region between western Canada and Alaska might eventually be included by constructing radar facilities along the coast.

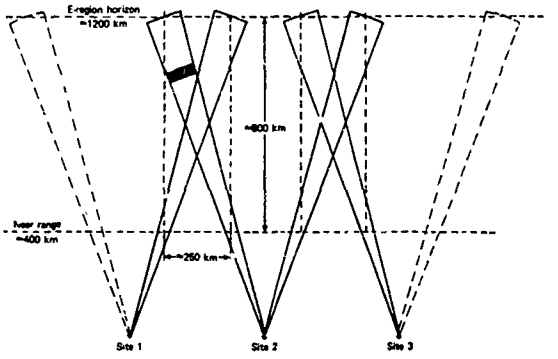


FIGURE 8 Plane view of beam intersection from three neighboring DARN radars. The analysis procedure assumes longitudinal uniformity of the electric field between the legs of each X.

Each radar in the DARN network will operate in the manner shown schematically in Figure 9. At the heart of each system is a microcomputer that will control the operation of the radar transmitter and receiver, control the direction of the antenna lobe, process the backscattered signal and transmit the processed data via a telephone link to a host computer. Operation of each station will be controlled from an on-line video monitor or from the host computer via the data link. In a like manner, the real time data from each station may be displayed on the video monitor or at the site of the host computer. Timing at each station will be provided by UTC-NBS time codes transmitted by the NOAA GEOS satellites.

Each of the DARN radars will use the single-pulse and double-pulse analysis modes that have been used in the STARE radars (Greenwald et al., 1978). Assuming that the microprocessor can handle the data stream, twenty-five single-pulse and double-pulse measurements will be made in each direction every second.

Backscattered signals will be processed on-line and the averaged data will be transferred continuously from each of the sites via a 1200 baud telephone line.

The temporal resolution of the DARN measurements is limited by our need to achieve sufficient statistical accuracy for the Doppler velocity estimators and by the amount of data that must be transmitted over the telephone lines. Of these two limitations, the latter will most likely be the more severe. For continuous data acquisition, it is unlikely that temporal resolution better than 20-30 seconds can be achieved. It is expected, however, that this resolution is more than adequate to meet the operational requirements of the system.

The operating characteristics of each DARN radar are summarized in Table 1.

TABLE 1

EXPECTED OPERATING CHARACTERISTICS OF A DARN RADAR

Operating Mode:	Pulsed-monostatic with bi-directional electronically-phased beam
Vertical Beamwidth (2-way):	≈ 6°
Horizontal Beamwidth (2-way):	≈ 6°
Operating Frequency:	Nominally 140 MHz
Pulse Power:	20 kW
Pulse Length:	200 us
Number of Ranges Analyzed:	27
Pulse Pattern Repetition Frequency:	Nominally 25 Hz each direction
Analysis Mode:	Single Pulse - intensity Double Pulse - mean Doppler

At the site of the host computer, it will be possible to display continually-updated plots of the intensity of auroral clutter across most of the North American auroral zone. These plots would be similar in appearance to the DMSP auroral image shown in Figure 1, however, they would be updated on 20-60 s time intervals. In addition, the Doppler data from the various station could be combined to yield the electric field and plasma drift velocity variations along each of the X-regions in Figure 8. This information can be further refined to yield a two-dimensional plot of the electrical potential variation across Canada. Such a plot might look similar to the example shown in Figure 10. Again, these electrical potential images would be updated on 20-60 s time intervals.

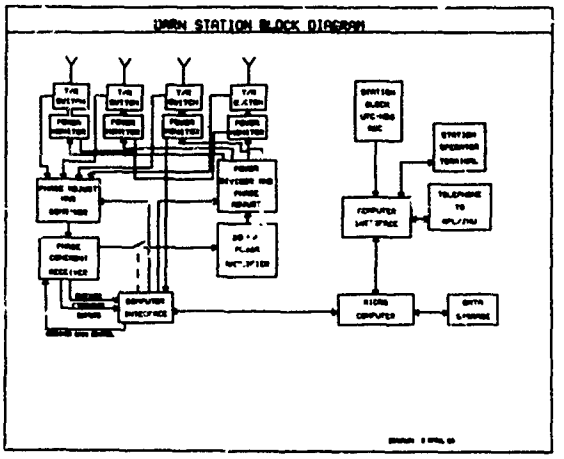


FIGURE 9 Block diagram showing elements of a typical DARN radar. Control of the radar and analysis of the backscattered signals is performed by the station microcomputer.

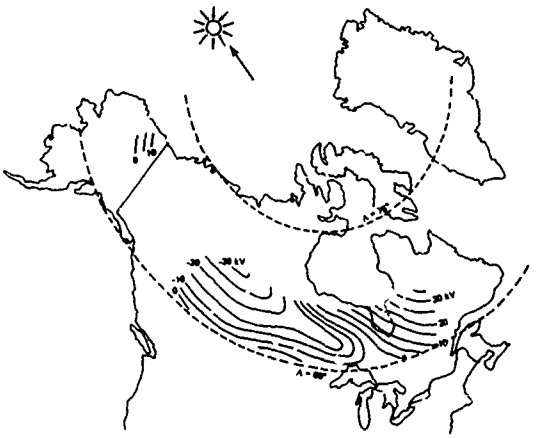


FIGURE 10 Example of instantaneous electrical potential distribution that might be derived from the DARN Doppler data.

BENEFITS OF DARN RADAR NETWORK

Given that a VHF radar backscatter network similar to the one described in this paper were established across North America, what might one expect as to potential benefits?

Firstly, such a system would act as an excellent monitor of ionospheric disturbances in the North American sector. Clutter maps indicating the intensity of E-region auroral backscatter would be displayed and updated continuously. Such maps would be especially useful in identifying regions in which surveillance radars might be affected by auroral clutter or in short term forecasting of periods in which radars might begin to be affected by auroral clutter.

The Doppler information provided by the DARN network might also be of use in evaluating the effect of auroral clutter on surveillance radars. Often these radars utilize Doppler discrimination wherein certain assumptions have been made as to the Doppler characteristics of the clutter. DARN will provide continuous monitoring as to whether the assumptions are valid.

Secondly, studies by Oksman et al. (1979) have indicated that auroral zone F-region irregularities often occur on the same field lines as E-region irregularities. Hence, the DARN images will provide some indication of regions from which one might expect to observe F-region clutter. Furthermore since F-region clutter produces scintillations in satellite-ground communications, one might be able to provide short term predictions of periods in which high latitude satellite-ground communications paths might be affected by this phenomenon.

Thirdly, imagery of the high latitude electric field and plasma drift is an important complement to DMSP auroral imagery. Together, they will better define the dynamic nature of large scale processes in the high latitude ionosphere. Information such as this is needed in order to develop reliable dynamic models of ionospheric variability. It is also needed for evaluating the magnitude of transent phenomena on radiowave systems.

Finally since measurements of electric fields and plasma drift in the high latitude ionosphere, reflect similar fields and drifts in the magnetosphere, ionospheric measurements may be used to reveal new information on large scale magnetospheric dynamics. Information such as this may be useful in evaluating the environmental change on satellite systems during periods of geomagnetic disturbance.

ACKNOWLEDGEMENTS

The author would like to thank T. A. Potemra for use of the TRIAD and MAGSAT data.

REFERENCES

Banks, P. M. and J. R. Doupnik, A review of auroral zone electrodynamics deduced from incoherent scatter observations, J. Atmos. Terr. Phys., 37, 951, 1975.

Cahill, L. J., Jr., R. A. Greenwald and E. Nielsen, Auroral radar and rocket double probe observations of the electric field across the Harang discontinuity, Geophys. Res. Letts., 5, 687, 1978.

Ecklund, W. L., B. B. Balsley and D. A. Carter, A preliminary comparison of F region plasma drifts and E region irregularity drifts in the auroral zone, J. Geophys. Res., 82, 195, 1977.

Greenwald, R. A., Studies of currents and electric fields in the auroral zone ionosphere using radar auroral backscatter, in Dynamics of the Magnetosphere, ed. S.-I. Akasofu, D. Reidel, Dordrecht, Holland, 1980.

Greenwald, R. A., W. Weiss, E. Nielsen and N. R. Thomson, STARE: A new radar auroral backscatter experiment in northern Scandinavia, Radio Sci., 13, 1021, 1978.

Gonzoles, C. A., M. C. Kelley, B. G. Fejer, J. F. Vickery and R. F. Woodman, Equatorial electric fields during magnetically disturbed conditions, 2. Implications of simultaneous auroral and equatorial measurements, J. Geophys. Res., 84, 5803, 1979.

Inhester, B., W. Baumjohann, R. A. Greenwald and E. Nielsen, Joint two-dimensional observations of ground magnetic and ionospheric electric fields associated with auroral zone currents. 3. Auroral zone currents during the passage of a westward travelling surge, submitted to J. Geophys. Res., 1981.

Meng, C. I., Electron precipitations and polar auroras, Space Sci. Rev., 22, 223, 1978.

Nielsen, E. and R. A. Greenwald, Variations of ionospheric currents and electric fields in association with absorption spikes during the substorm expansion phase, J. Geophys. Res., 83, 5645, 1978.

Nielsen, E. and R. A. Greenwald, Electron flow and visual aurora at the Harang discontinuity, J. Geophys. Res., 84, 5645, 1979.

Oksman, J., H. J. Moller and R. A. Greenwald, Comparisons between strong HF backscatter and VHF radar aurora, Radio Sci., 14, 1121, 1979.

Potemra, T. A., T. Iijima and N. A. Saflekos, Large-scale characteristics of Birkeland currents, in Dynamics of the Magnetosphere, ed. S.-I. Akasofu, D. Reidel, Dordrecht, Holland, 1980.

Rostoker, G., The auroral electrojets, in Dynamics of the Magnetosphere, ed. S.-I. Akasofu, D. Reidel, Dordrecht, Holland, 1980.

Walker, A. D. M., R. A. Greenwald, W. F. Stuart and C. A. Green, STARE auroral radar observations of pc5 geomagnetic pulsations, J. Geophys. Res., 84, 3373, 1979.

AN ADVANCED MISSION TO MAP THE WORLDWIDE TOPSIDE IONOSPHERE

H.R. Mathwich, D.E. Aubert, and A.F. Martz
RCA, Astro-Electronics,
Princeton, N.J.

K. Bibl and B. Reinisch
University of Lowell, Center
for Atmospheric Research,
Lowell, Mass.

Lt. D. Lewis
USAF, SD,
Los Angeles, CA

INTRODUCTION

The topside sounders of the Alouette and ISIS series have provided data leading to an understanding of the topside ionosphere which is of incalculable value. These data have defined the gross morphology via statistical analysis of tens of thousands of ionograms. Further an understanding of smaller scale-size phenomena has been achieved by studying in detail the anomalous properties of ionograms sounded in regions containing a disturbed ionosphere. Since the design and launch of these research sounders, radical improvements in the state of the topside sounding art have occurred. Advanced sounding techniques are available. Host spacecraft capable of supporting a mission containing many sensors have been designed. Sophisticated, rapid data processing techniques can now be designed using today's computers. It is now possible to configure topside sounding missions in which massive quantities of data can be taken with the data reduction accomplished with little or no human intervention. Possible missions include the sounding of data which is analyzed and fed to computer-based models of the earth's ionosphere, missions which continuously define certain distinct morphological features such as the mid-latitude or main trough or the equatorial anomaly and missions which study the small scale size irregularities. It is the purpose of this paper to describe the design of a mission of the first type - one in which perhaps 3000 to 6000 ionograms are sounded every day, the data reduced automatically and used to supply an ionospheric model (von Flotow, et al., 1978; Tascione, et al., 1979; Bent, 1975).

MISSION DESIGN

The total mission is illustrated in Figure 1. An advanced topside ionosonde orbits in a 3-axis stabilized non-dedicated spacecraft. The spacecraft orientation is kept fixed with respect to its velocity vector and nadir. This makes possible the use of an advanced sounder antenna. On-board data storage allows the continuous collection of data. The use of

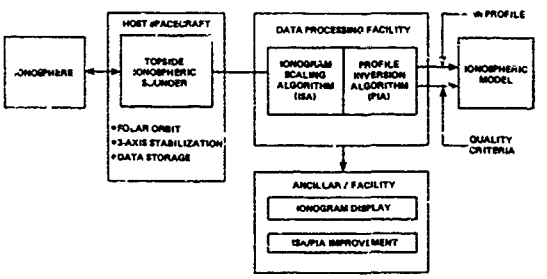


Figure 1. Mission Block Diagram

a near-polar constant-local-time orbit at a satellite height (hS) of approximately 800 km provides data sampling at all geomagnetic latitudes approximately once every 100 minutes. An earth-based Data Processing Facility (DPF) automatically scales the ionograms using an Ionogram Scaling Algorithm (ISA) and inverts the data to produce an electron density profile (Nh) in a Profile Inversion Algorithm (PIA). This data is fed to the ionospheric model along with quality criteria which define the degree to which the data is complete and accurate. To make such a mission feasible this on-line data processing must be completely automated. Ancillary facilities permit off-line interaction with the mission and improvement of the algorithms as experience accrues.

The use of a non-dedicated host spacecraft brings with it both advantages and disadvantages. A major advantage is mission cost-reduction, since the price of the bus, the on-board data storage facilities and earth-based data collection facilities is shared. The disadvantages include the presence of other sensors which may be sensitive to the pulsed r-f power output from the sounder, spurious sounder-receiver input signals from other spacecraft components and

the requirement of adding relatively long booms to the spacecraft for the ionosonde's antennae. Further, the shared spacecraft data-storage and playback facilities may possess a limited data-rate capability requiring the ionosonde to perform extensive data-compression.

The mission design to be discussed used the Defense Meteorological Satellite Program (DMSP) Block 5D2 spacecraft as the host spacecraft. An interface study placed the following limitations on the ionosonde:

Maximum Transmitter RF Power	= 30 Watts
Maximum Boom Length	= 17 Meters
Maximum Data Rate (R_T)	= 3168 bits/second

GLOBAL IONOSPHERIC MORPHOLOGY

We first present (in Figures 2 and 3) a simplified view of the gross morphology of the earth's ionosphere. This view is necessary to (a) determine the relationship between the important variables in the ionosphere and the sampling process provided by the ionosonde, and (b) in conjunction with an ionogram survey to be discussed, to determine mission and ionosonde parameters requiring definition.

The dominant variables which determine the gross morphology are geomagnetic latitude, local time of day (LT), time of year (or season) and sun activity. Other variables such as geographic position (with LT fixed) are of secondary importance.

Figure 2 illustrates the manner in which a polar orbiting ionosonde in a noon/midnight orbit would sound the ionosphere. There are three general regions - equatorial, mid latitude, and polar - each with day/night properties. The polar region illustrated is after Buchau 1973. The nightside mid-latitude or main trough is narrower than that illustrated (e.g., 2° to 5° wide).

Figure 3 illustrated the manner in which two satellites - one in a 12 hr./00 hr. orbit and one in an 08 hr./20 hr. orbit - would sample the ionosphere. The satellite tracks remain approximately fixed on the morphology as illustrated. These figures do not depict small scale size irregularities which can affect the quality of topside ionograms. To a first approximation the day-side equatorial and mid-latitude ionosphere possesses few such irregularities. The polar region and night-side equatorial region can possess many. The scintillation zone shown in Figure 3 is after Taur 1973.

IONOGRAM SURVEY

To properly design a mission of this type, it is essential that a broad sample of the types of ionograms

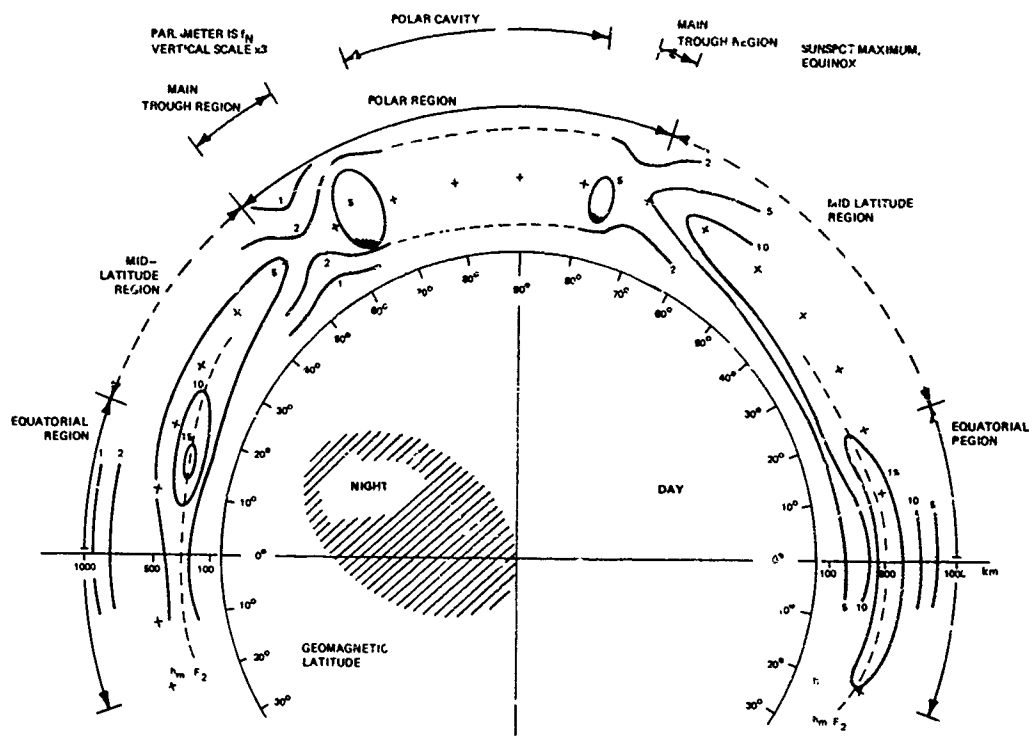
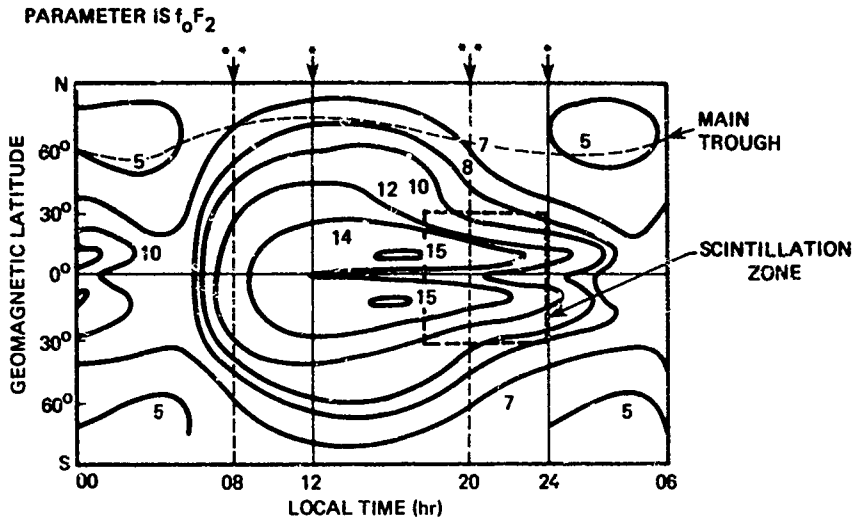


Figure 2. Global Ionospheric Regions



* NOON-MIDNIGHT ORBIT SAMPLING
** DAWN-DUSK ORBIT SAMPLING

After Johnson, 1965

AT EQUINOX DURING SUNSPOT MAXIMUM, 1947

Figure 3. Ionosphere Diurnal Characteristics

anticipated be obtained. This sample is required to determine sounder parameters such as frequency range, frequency step size, and antenna requirements and more importantly to establish a technical base for the design of the automated scaling and inversion algorithms. An ionogram survey has been completed in which some 500 ionograms sounded by ISIS-1 and ISIS-2 in approximately noon/midnight orbits were viewed. The ionograms were provided by the Communications Research Centre Ottawa, Canada. Ionograms from the following regions were studied:

Region	
Mid-Latitude Daytime Polar Nighttime	650 km $\leq$ hS $\leq$ 1800 km
Equatorial Daytime	hS = 1400 km; 580 km $\leq$ hS $\leq$ 800 km
Equatorial Nighttime	hS = 1400 km; 580 km $\leq$ hS $\leq$ 770 km

The survey pointedly attempted to study both good and disturbed ionograms with the widest possible variety of characteristics and problems.

Some results of this survey are as follows:

- A generic X-trace shape exists which might be of use in developing an ISA.

- Lower (e.g., 800 km) rather than higher (e.g., 1400 km) values of hS are to be preferred for this mission. Fewer problem ionograms result at the lower height due to decreased field-aligned propagation effects.

- The fraction of ionograms which might be deemed scalable using automated techniques is radically locale-dependent varying from near 100% in equatorial daytime to perhaps 10% in equatorial nighttime at high hS values.

- The ionosonde frequency range should be from 0.5 MHz to 19 MHz. Antenna efficiency can afford to decrease below approximately 3 MHz.

- A basic ISA strategy was established which consisted of requiring the following features: A resonance-region subroutine is required to establish f_0S and f_XS. Vertical-trace insensitivity is required to cope with earth-based-rf interference and nf_H resonances. O- and X-trace separation is highly desirable.

One atypically good ionogram is shown in Figure 4 for illustrative purposes. This ionogram (designated I-2) will be used in the remainder of this paper so that comparative results can be viewed. It is one of the best used in this study and is by no means representative of the ionogram quality required by the ISA's in order that successful scaling be achieved.

CANDIDATE IONOSONDE

In this section we describe an advanced topside ionosonde designed for this mission.

Recent advances in digital bottomside sounder technology now provides a basis for advanced topside sounder technology (Bibl and Reinisch, 1978). The following specific features are available and have demonstrated value.

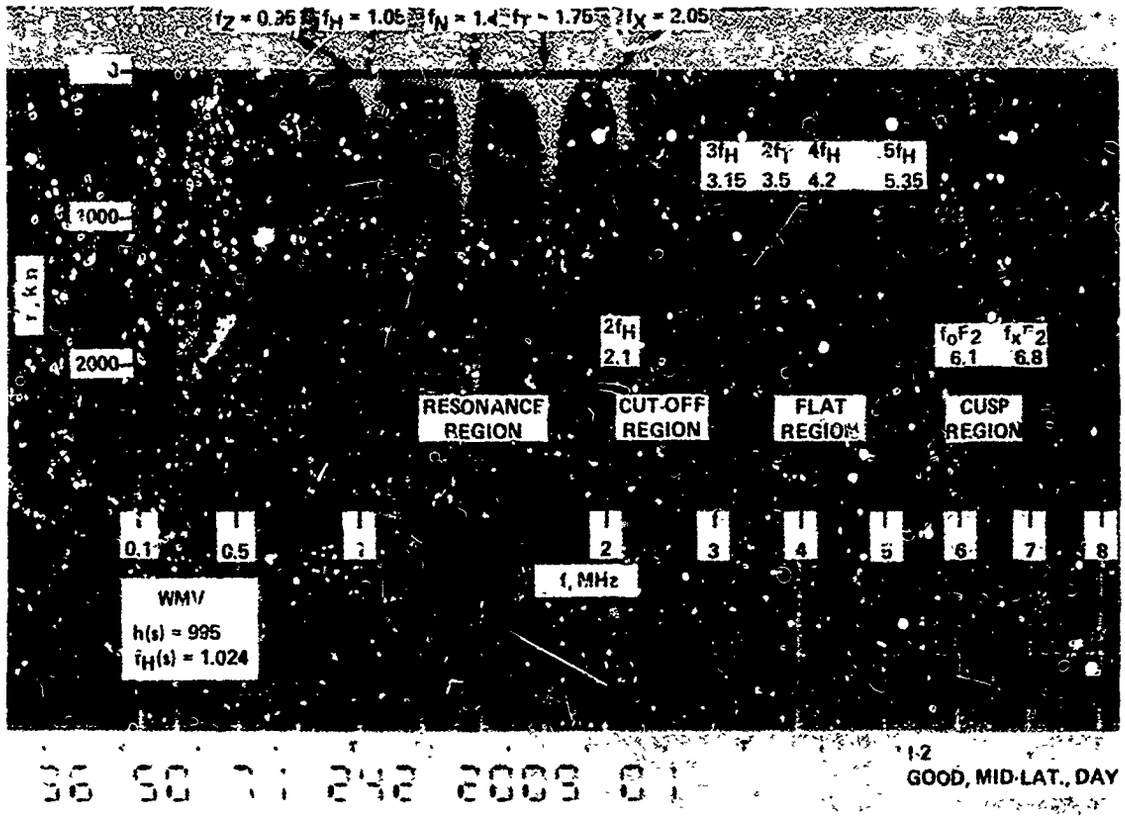


Figure 4. Good, Mid-Latitute, Daytime Ionogram (I-2)

- The use of a multi-chip, multi-pulse sounding signal which is coherently processed using a discrete Fourier transform (DFT). This provides signal processing gain, Doppler data, and reduction of the effects of earth interference.

- The use of the above type coherent signal processing has been found to extract coherent echo traces in the presence of certain small scale-size irregularities such as spread.

- The feasibility and value of separating the ordinary (O) from the extraordinary (X) signals has been established.

A block diagram of this advanced topside ionosonde is shown in Figure 5. A performance summary is shown in Table I. Typical ionogram times are shown in Table II. Using these it is estimated that great circle ionogram resolutions of about 1 degree will be obtainable in polar regions and 2 degrees in equatorial regions. The ionosonde start-frequency, stop-frequency and step size combinations can be programmed since many specific ionogram formats are available.

The analog processing and digital processing portions of this ionosonde have been constructed and tested and are now being used in a bottomside sounding test assembly. Extensive software simulation of the performance of the digital processing portion has been completed. The performance predictions closely check the performance measured in the hardware.

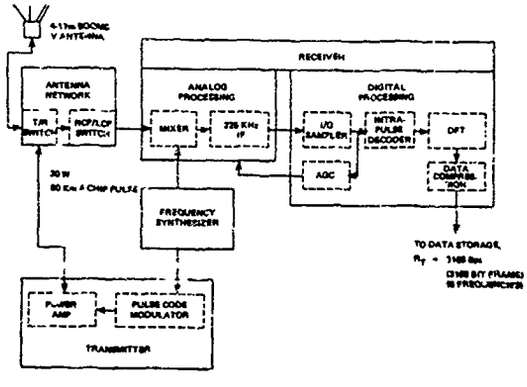


Figure 5. Ionosonde Functional Block Diagram

TABLE I

IONOSONDE PERFORMANCE SUMMARY

Parameter	Value
Frequency Range	0.5 to 19 MHz
Frequency Step Size	25, 50, or 100 KHz
Sounding Time/Frequency	0.125 Sec, Nominal
Ionogram Duration	10 to 40 Seconds 80 to 320 Frequencies 0.6 to 2.4 Great Circle Degrees
Apparent Range	125 to 1995 Km
Range Pixel	10 Km
PRF	200 and 263 Hz
Transmitter Power	30 Watts
Transmit Polarization	Linear
Receive Polarization	RCP and LCP
Processing Gain	16.8 dB
Telemetry Bit Rate	3168 bps

TABLE II

TYPICAL IONOGRAM TIMES

Start Frequency (MHz)	End Frequency (MHz)	Step Size (KHz)	Ionogram Time (Sec.)	Great Circle Traverse (Degrees)
0.5	10.95	50	25	1.5
0.5	20	100	12	0.7
5.0	18.95	50	35	2.1
5.0	18.5	100	17	1.0

In summary, this sounder possesses three unique features:

- Coherent Signal Proce ing
- RCP/LCP Tagging
- Extensive On-board Data Compression.

The manner in which these properties are achieved will now be briefly discussed.

Ionosonde Signal Design

The fact that the ionosonde must possess substantial signal processing gain (to permit the use of a low transmitter power) along with the requirement that ionograms be taken at sufficiently rapid intervals to accurately define the ionosphere requires that the ionosonde utilize a complex sounding signal. Complete description of this signal is beyond the scope of this paper. Briefly, each frequency is sounded using 27 pulses each consisting of four contiguous 133.3 micro second duration chips. Inter- and intra-pulse bi-phase coding is used to remove apparent range (r') ambiguities. Odd frequencies are sounded with PRF = 263 Hz and even frequencies with PRF = 200 Hz. The window structures used provide complete r' coverage from 125 km to 1995 km. About one-half the pulses are sounded with the receive antenna in a right circular polarized (RCP) configuration and the remaining half with the antenna in a left circular polarized (LCP) configuration. The use of this signal provides 16.8 dB of processing gain and RCP/LCP polarization tagging.

Extensive data compression is required to reduce the data rate to that acceptable to the telemetry link (3168 b/s). The Data Compression Algorithm (DCA) compares the amplitudes of the signals in each of 3 range windows, in each of 8 Doppler bins and each of 2 polarizations and selects the maximum value for transmission. This results in a data compression of 48:1. This process and the resultant ionogram format are illustrated in Figure 6.

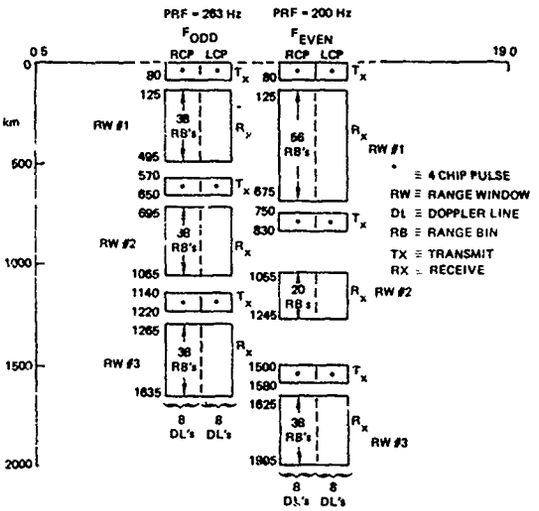


Figure 6. Ionosonde Sounding Signal

Ionosonde Antenna

The ionosonde antenna is worthy of special note. It possesses the following features:

- RCP/LCP separation is available.
- Single length monopoles successfully cover the required frequency range.
- Compatibility with the host spacecraft (e.g., the field-of-view requirements of other sensors) is attained.

The physical configuration is shown in Figure 7. The antenna consists of four 17 meter monopoles organized as two orthogonal V-type dipoles with an included angle of 100°. The use of the V-type configuration is its notable feature. A V-antenna has a pattern without lobing over a wide frequency range. The match properties are similar to those of a linear dipole of the same length. Satisfactory match, beam width and polarization separation properties are achieved over a 3 MHz to 19 MHz frequency range. The feasibility of achieving correct RCP/LCP tagging in the ionosonde has been clearly established. Conversion to O/X tagging on earth in the earth-based DPF is done using a simple algorithm based on knowledge of the position of the magnetic equator. The principles of this process are shown in Figure 8.

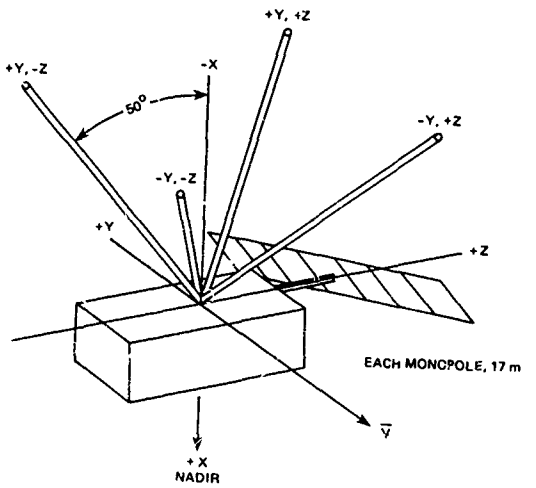


Figure 7. Sounder Antenna Configuration

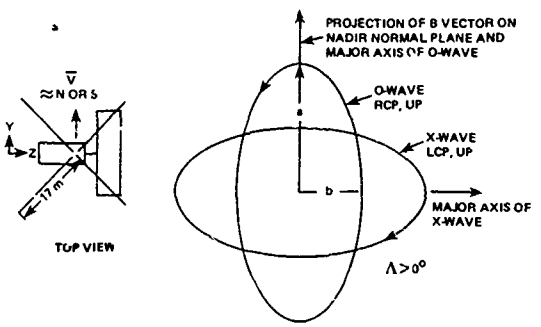


Figure 8. O/X Tagging Feasibility, Antenna and Ionospheric Relationships

During the transmit interval the two +Z monopoles are connected together as are the two -Z monopoles. Thus linear excitation of the X-type phenomena is achieved even near the magnetic equator. During the receive interval, the four monopoles are used to alternately create an RCP and then LCP circularly polarized receive antenna. Correct tagging will be achieved to within a few degrees of the magnetic equator.

IONOGRAM SCALING ALGORITHM DEVELOPMENT

The central problem in the mission design being discussed is the design and test of automated techniques to scale and invert the data delivered by the ionosonde. This must be done at the real-time data-reception rate using mini, not main-frame computers. A satisfactory tradeoff is to successfully scale a high enough fraction of all ionograms to satisfy the needs of the model keeping the process rapid and reliable by not demanding scaling of severly anomalous ionograms.

A first task is to provide digitized test data for the ISA/PIA development. To this end an Ionogram Reconstruction Facility (IRF) has been designed. This accepts as its input analog video ISIS-1 or -2 data as recorded on magnetic tape. These data are first digitized and formatted such that an ISIS -type ionogram exists. The absolute amplitude of the signal is preserved. Second, these data are then reformatted to have an r' pixel size (10 km) and frequency step size equal to that of the planned advanced digital sounder (e.g., 50 kHz or 100 kHz). Third, these data are compressed using the ionosonde's DCA. The output then simulates that to be expected from the ionosonde minus O/X tagging and Doppler data. A sample of IRF output before passage through the DCA is shown in Figure 9.

The central task of developing ISA and PIA techniques was first conservatively pursued assuming O/X tagging data would not be available. Two separate approaches now exist - one developed at the University of Lowell, Center for Atmospheric Research, Lowell, Mass. (ULCAR) (Reinisch and Huang, 1981) (Huang and Reinisch, 1981) and one at the RCA Laboratories, Princeton, N.J. (Waldman and Mathwich, 1980). Each possesses certain common features or subroutines such as the following:

- A geomagnetic field subroutine which calculates the values of f_H, B, and I as a function of spacecraft position.
- Thresholding to accept for processing only data which exceeds noise by an amount necessary to give adequate probability of signal detection and allowable probability of false alarm.
- A resonance-region subroutine which defines f_{OS} and f_{XS} thereby establishing a starting point on the echo traces and eliminating essentially all of the resonance-region data from subsequent processing.

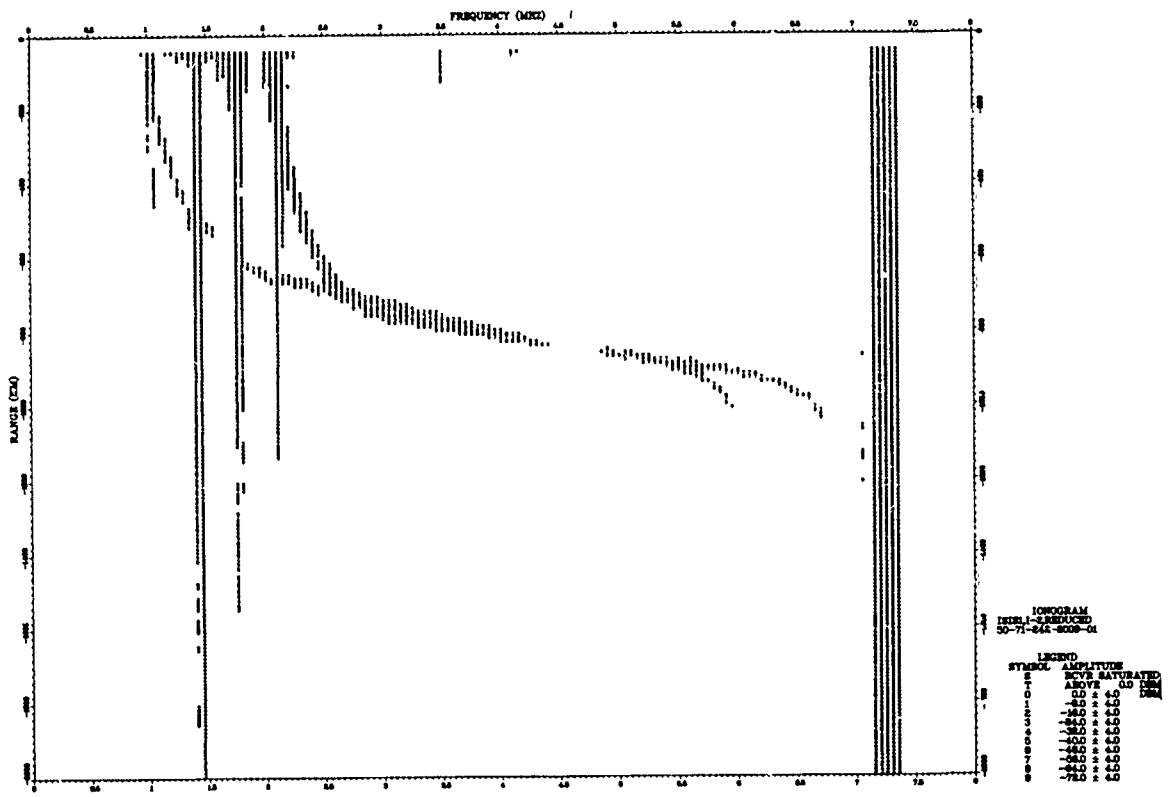


Figure 9. Ionogram From IRF

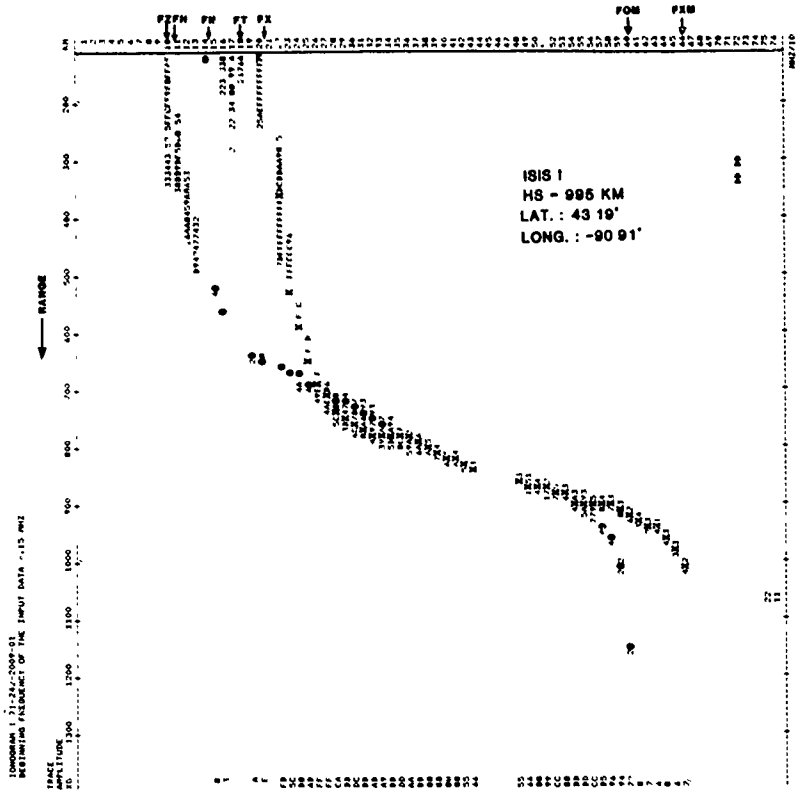
Aside from these, however, the two approaches were deliberately kept dissimilar. The ULCAR approach analyzes the amplitude distribution as a function of apparent range and frequency, separates the O- and X-trace pixels using a "separation line," and curve-fits directly to the electron density (Nh) profile using a novel profile inversion algorithm. Figure 10, the result of scaling I-2 (Fig. 4), shows the successful separation of the O- and X-traces. The PIA, using both the O- and X-trace points, calculated the profile shown as the solid line in Figure 11. The dashed curve resulted from applying the standard lamination inversion technique (Jackson, 1969) to the manually scaled ionogram. The agreement is excellent and, for this simple ionogram, use of only the O- or only the X-echoes would have resulted in essentially the same profile (see the O- and X-marks in Figure 11). For more complicated ionograms, however, the simultaneous use of both O- and X-echoes substantially improved the quality of the profiles.

This ULCAR ISA has been tested by applying to it digitized ISIS ionograms as produced in the IRF.

These ionograms, selected during the ionogram survey, included some which were chosen as automated-scaling objectives and some problem ionograms selected to deliberately overstress an ISA to determine its limits. This ISA successfully automatically scaled approximately 90% of the "objective" ionograms and, in addition scaled several for which success had not been anticipated.

The RCA ISA/PIA operates on the data in a global fashion. It makes use of curve-fitting techniques and extracts the X-trace which it fits to an algebraic equation. Profile inversion is done using modified versions of the Jackson algorithm (Jackson 1969). The current success rate for the RCA approach is approximately the same as that of the ULCAR approach. One X-trace result is shown in Figure 12 and the resultant Nh profile in Figure 13. The "manual" profile resulted from human scaling of the data-compressed ionogram. The total electron content (TEC)* for these two profiles for this good ionogram differ by only 1.2%. An average difference of about 10% has been measured using a larger ionogram sample.

Neither of the above currently available ISA algorithms makes use of the valuable O/X tagging data available from the ionosonde. Advanced algorithms of both types are now in development. It is anticipated that these will require substantially



(B) EXTRACTED O AND X TRACE POINTS (71-242-2009-01)

Figure 10. ULCAR ISA Result

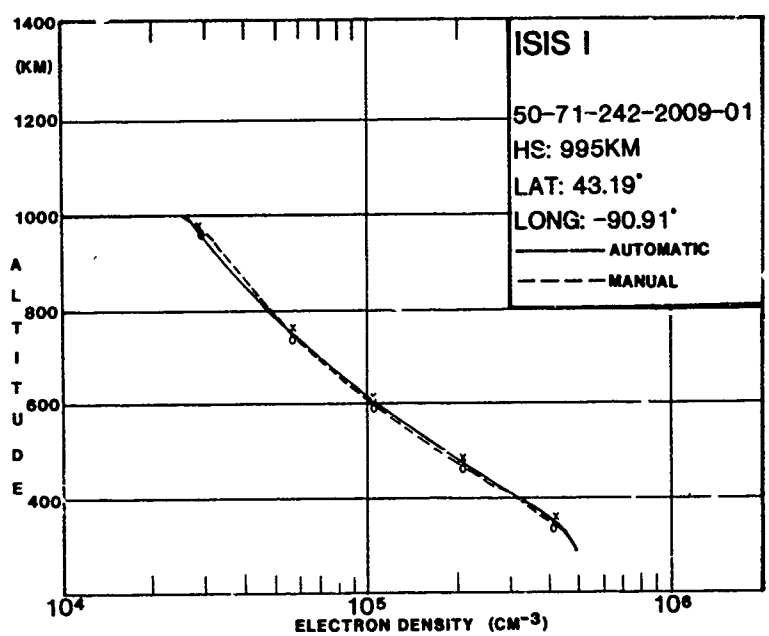


Figure 11. Electron Density Profile

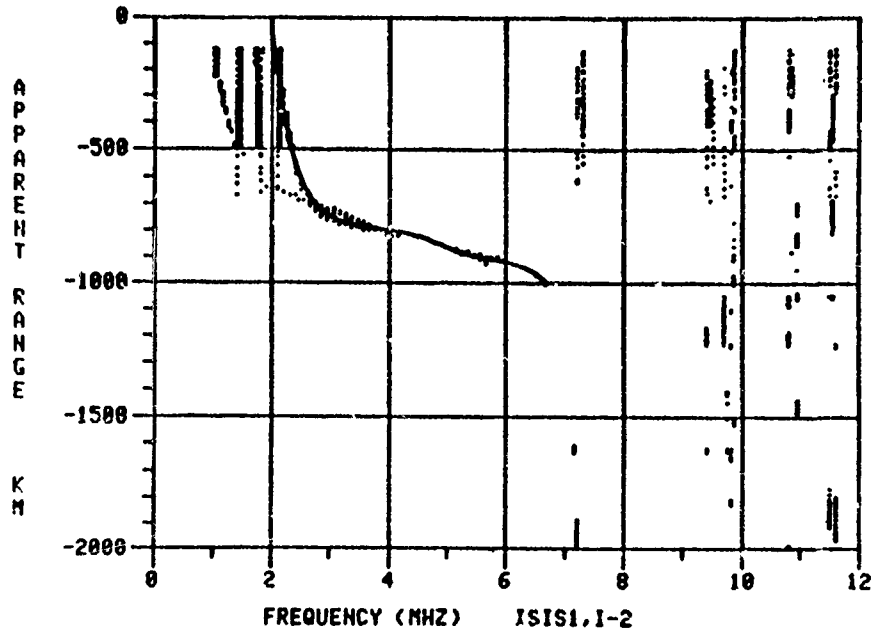


Figure 12. RCA ISA Result

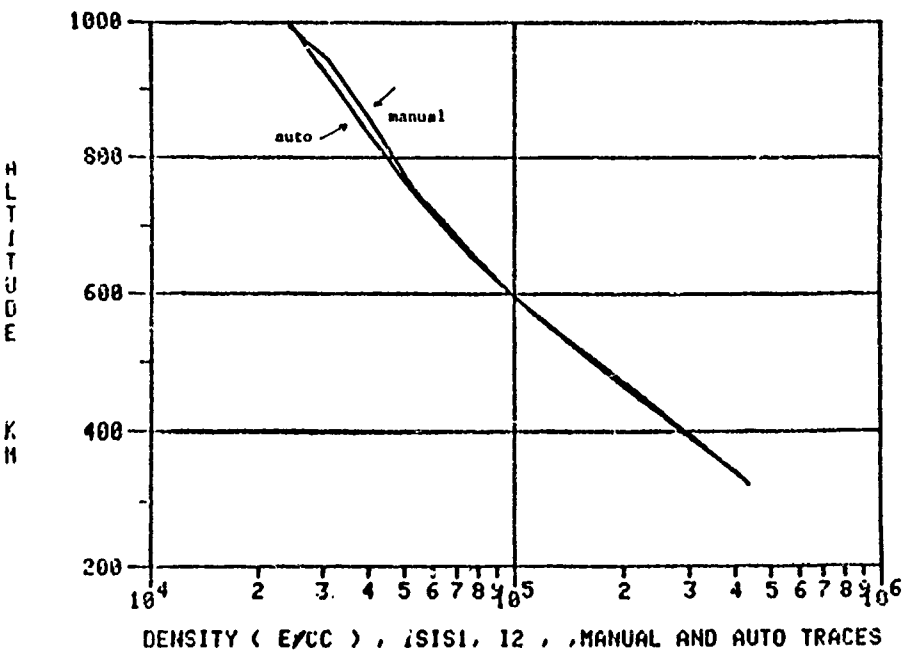


Figure 13. Electron Density Profile

simplified software and will provide even better performance.

CONCLUSIONS

In this paper we suggest that the current state of the art in advanced ionospheric sounders, host spacecraft and automated data analysis make possible new types of topside sounding missions which sound and analyze the data in real time using automated techniques. One such mission is designed which uses a topside sounder employing an inter- and intra-pulse coded sounding signal which is coherently detected in the sounder using a DFT. A novel antenna provides tagging of the RCP/LCP received pixels which can be used to O/X tag these pixels in the Data Processing Facility. The ionosonde sounds from a polar-orbiting spacecraft (the DMSP 5D2 spacecraft in this example) at an altitude of approximately 800 km. Surveys of existing topside ionograms have defined the range of characteristics expected in ionograms to be processed. Extensive on-board data compression reduces the telemetry link requirements to 3168 b/s. Automated ionogram scaling algorithms and profile inversion algorithms have been designed and tested using simulated data. Two ISA/PIA approaches using quite different techniques have both resulted in a high success rate in processing these simulated topside ionograms. Advanced ISA/PIA approaches are now being designed to better use the data sounded.

*Here TEC means only that integrated electron content between the lowest and highest Nh values.

Acknowledgements

Thanks are due to CRC personnel namely D. Muldrew, J. Whitteker, G. James and W. Campbell for their aid in providing ionogram data and in interpreting anomalous ionograms. H. Waldman of the RCA Laboratories, Princeton, N.J. designed the RCA ISA/PIA. H. Staras of the RCA Laboratories contributed a theoretical analysis of the O/X tagging task. J. Graebner of RCA, Astro Electronics provided theoretical performance predictions for the advanced topside ionosonde. D. Nelson of Aerospace, Los Angeles, CA provided valuable mission-design guidance.

Major portions of this mission design effort were funded by the U.S. Air Force, Space Division (SD) under contracts numbered F04701-78-C-0016 and F04701-80-C-0037.

References

Bent, R.B.; Llewellyn, S. K.; Nesterczuk, and Schmid, Paul E., "The Development of a Highly Successful World Wide Empiral Ionospheric Model and Its Use in Certain Aspects of Space Communications and World-Wide Total Electron Content Investigations," Proceedings, 1975 Symposium on the Effects of the Ionosphere on Space Systems and Communications.

Bibl, K. and Reinisch, B., "The Universal Digital Ionosonde," Radio Science. Vol. 13, No. 3, p. 519-530, May-June 1978.

Buchau, J; Pike, C. P.; Wong, M., "Detailed Specification of the Arctic Ionosphere and an Application to Three-Dimensional Raytracing," AFCRL-TR-73-0726, 27 November 1973.

Jackson, J. E., "The Reduction of Topside Ionograms to Electron-Density Profiles," Proc. IEEE, Vol. 57, No. 6, pp. 960-976, June 1969.

Johnson, F. S., "Satellite Environment Handbook," Stanford University Press, 1965.

Reinisch, B. W.; Huang Xueqin, "Automatic Evaluation of Topside Ionograms," URSI National Radio Science Meeting, Boulder, Colorado, 12-15 January 1981.

Tascione, Capt. T. F.; Flattery, Col. T. W.; Patterson, Maj. V. G.; Secon, Capt. J. A.; Taylor, Capt. J. W., "Ionospheric Modeling at Air Force Global Weather Center,": International Solar-Terrestrial Predictions; Proceedings and Workshop Program, Boulder, Colorado, 23-27 April 1979.

Taur, R. R., "Ionospheric Scintillation at 4 and 6 GHz,": Comsat Technical Review, Vol. 3, No. 1, pp. 145-163, Spring 1973.

von Flotow, Capt. C. S., "Ionospheric Forecasting at Air Force Global Weather Center," Proceedings, 1978 Symposium on the Effects of the Ionosphere on Space and Terrestrial Systems, 24-26 January 1978.

Waldman, H. and Mathwich, H. R., "Ionogram Scaling Algorithm," URSI, North American Science Meeting, Quebec, 2-6 June 1980.

Huang Xueqin, and B. W. Reinisch, ' Automatic Calculation of Topside Electron Density Profiles," URSI National Radio Science Meeting, Boulder, Colorado, 12-15 January 1981.

GLOBAL REAL-TIME IONOSPHERIC MONITORING

J. W. Wright
NOAA/ERL Space Environment Laboratory
Boulder, Colorado

ABSTRACT

An international global network of about 90 modern digital ionosondes is advocated to replace the 160 analog machines now active. Instruments following the Dynasonde concept employ a minicomputer for adaptive definition of data acquisition patterns and system control, for signal processing and echo recognition, for data processing and conversion to geophysical or radio communication parameters, and for data communications and interchange. About 90 practical observing sites on this planet suffice for vertical soundings of our ionosphere, and these same ionosondes may conduct bistatic oblique soundings between their 5 or 6 nearest neighbors (averaging 2500 km distant), resulting in about 320 locations of observation. The electron density distribution, lateral gradients or tilts, drifts, and Doppler measurements at each location are to be combined with similar data from the neighbor locations and each midpoint, to provide a real-time regional model. Applications in geophysics, radio communications, and disturbance monitoring are mentioned.

INTRODUCTION

The Present Ionosonde Network

There are about 160 ionosondes operating on the globe today (Figure 1). Only a handful of these are reasonably modern instruments, the great majority following design concepts developed in the 1940's and 1950's. They commit their measurement information to analog film in the form of ionograms. Only a small fraction (usually 25%) of the ionograms are reduced to numerical data. The stations function in a largely uncoordinated manner; it is usually only in a long-term (monthly-median) sense that this degree of global coverage can even be approached. Several regional programs (USA, USSR, France, India) represent and forecast large-scale ionospheric structure, e.g., Argo and Rothmuller (1979); Hatfield (1979). A global program (Thomson and Secan, 1979) is maintained by the AFGWC, based in part upon data from 44 widely-spaced ionosondes (some providing near-real-time parameters). However, this comparatively ambitious effort is considered data-starved by its authors. I wish to encourage a planned redistribution of these resources, leading to a smaller number of modern stations that, individually and collectively, will provide much more information (and do so more rapidly) than the present network.

The Role of the Ionosonde in Atmospheric Physics

Radio sounding by total reflection occupies a key position in ionospheric measurement: the plasma is a sensitive tracer of virtually all atmospheric processes underway in the ionosphere, and radio wave reflection provides a plasma probe of unparalleled sensitivity. High signal-to-noise ratios are obtained easily with systems of modest power and without recourse to long signal integration times. Radio frequency gives electron density directly at total reflection, and distances are measurable to about 1% of the radio wavelength, a few meters. Some of the measurements are useful directly, without conversion to geophysical parameters, as in the representation of long-distance radio communication conditions. They may also be inverted to a variety of atmospheric information, much of it available in no other practical way. The methods can provide information on all temporal scales, extending from the beginning of our awareness of the ionosphere a half-century ago, down through those (of solar or terrestrial origin) occupying fractions of a second. Information describing

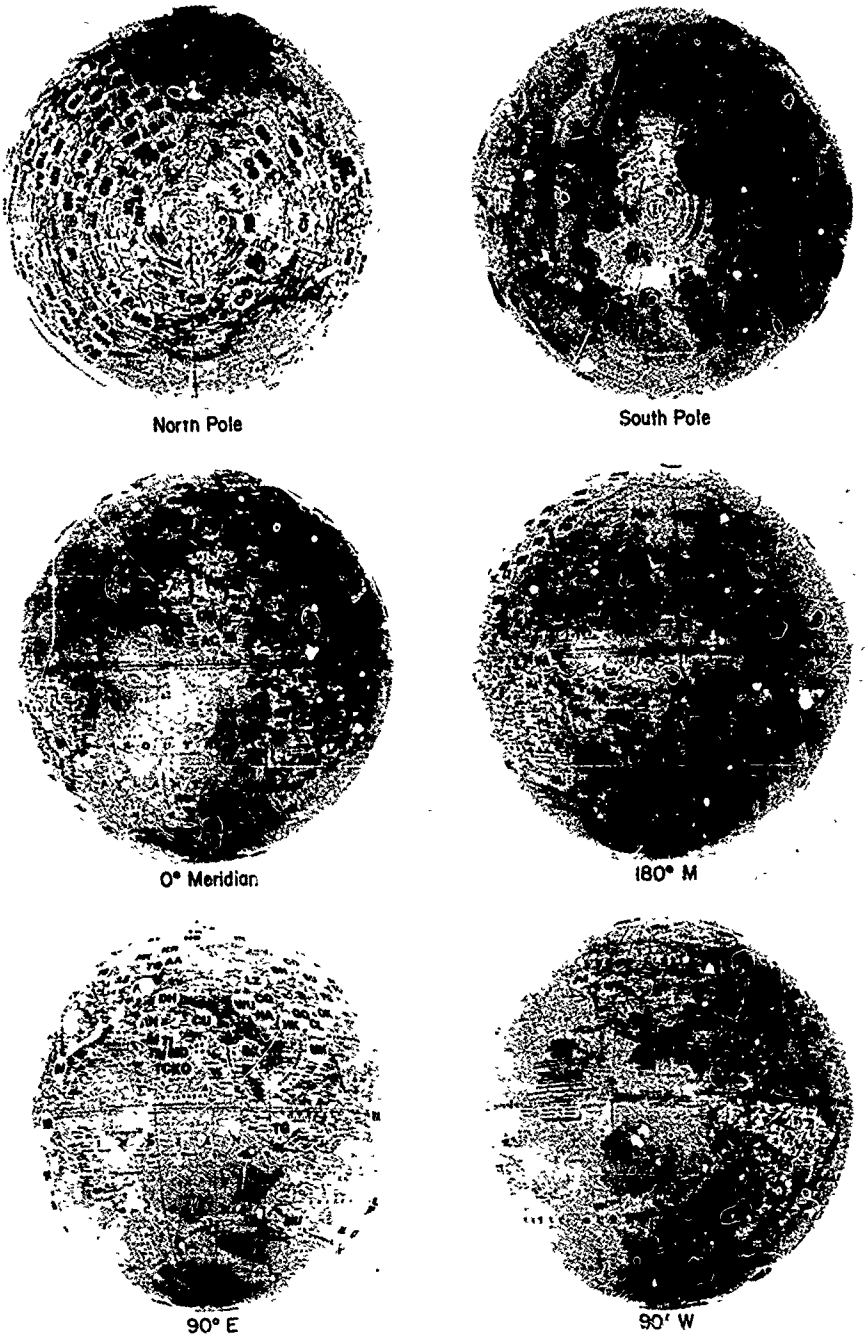


Figure 1. The present active ionosonde network, of about 150 conventional instruments. Stations are listed in Table 1. Station labels have 500-km radius and suggest the field of view afforded by usual antenna systems and instrument sensitivity; however, most conventional ionosondes cannot determine echolocation and thus confuse horizontal and vertical structure.

a wide range of spatial scales is accessible from a single observing location, typically starting with those comparable to the radio wavelengths used (.01-.3 km) and extending far beyond the first Fresnel zone scale (typically 3 km) to horizontal scales at least as large as the height of the ionosphere (100-300 km).

To attain the full benefit of total reflection sounding there are three inescapable requirements:

(1) The measurement system must have complete agility in time and frequency sounding patterns, and it must obtain a full description of each radio echo. This has only recently become possible, and it follows that the system must be programmable and digital.

(2) The measurements must be fully reduced to the geophysical information they contain; otherwise, when reduction is incomplete, the interpretation of each part suffers unnecessary ambiguity.

(3) Global (or at least regional) and continuous monitoring is necessary; first for obvious reasons identical to those for meteorology, and also--as suggested by this paper--for the temporal and spatial continuity of measurement that provides the necessary information by which each instrument can adapt its data acquisition pattern to prevailing conditions in the ionosphere.

In the following sections we summarize the demonstrated measurement capabilities of the "Dynasonde" class of ionospheric measurement systems (Wright, 1969; Wright and Pitteway, 1979 a,b), that we believe satisfy requirement (1), above. We also describe some natural extensions of these capabilities to provide oblique incidence, backscatter, and passive observations of the ionosphere, each contributing to increased lateral coverage. Finally, we attempt to suggest the size, deployment, modes of operation, costs, and benefits of a global ionospheric monitoring network satisfying requirement (3).

DYNASONDE MEASUREMENT CAPABILITIES

Hardware Aspects of a Modern Digital Ionosonde

The necessary hardware features of Dynasonde systems have been adequately described elsewhere (Wright, 1975, 1977) and need not be reviewed here. It is sufficient to show, with Figure 2, the functional layout of the latest such system as designed and assembled at NOAA-SEL by R. N. Grubb. The general-purpose computer, supported by the usual digital peripherals and by two microcomputers dedicated respectively to real-time signal processing and system timing management, are the essential features by which requirement (1) of our Introduction is achieved.

Geophysical Functions

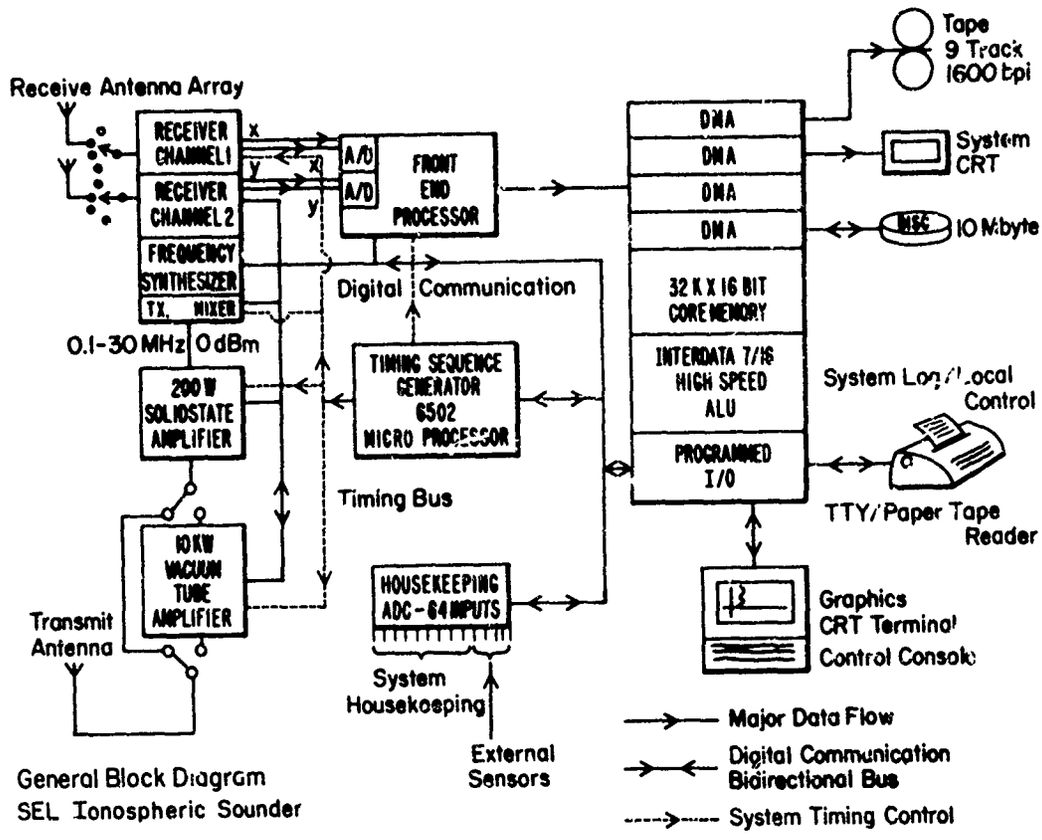
Figure 3 suggests four geophysical functions of a digital ionosonde:

Observables are the dependent variables which are measured for each ionospheric echo--effectively the group path and phase, and echo amplitude. These (or equivalent) quantities must be obtained as functions of the independent variables for space, time, and radio frequency. The frequency converts directly to plasma density, but the spatial localization of the echo is an involved inversion problem.

Derivables are quantities obtained by calculation from the observables using established theory and (if appropriate) statistically informative estimation procedures. The status of some of these data inversion procedures was discussed by Wright (1975a). They include methods for obtaining the electron density distribution from multifrequency vertical, oblique, and backscatter soundings; methods for locating and tracking discrete targets of medium scale; and methods for deriving ionospheric velocities.

Deducibles are almost all of the parameters, properties, and processes of the higher neutral atmosphere and ionosphere. Some of these (e.g., winds, conductivity) are almost directly related to observables and might be included in the derivable category. Others (e.g., eddy mixing, electric fields) are hardly observable by any direct means and must be deduced from their effects.

Comparables are quantities or processes that are measurable by means entirely independent of radio sounding. They are often important for validation of data acquisition and inversion methods, and the ionosonde may offer as much, if not more than it receives, in such comparisons.



General Block Diagram
SEL Ionospheric Sounder

Figure 2. Functional and hardware layout of the digital ionosonde currently assembled at NOAA, Space Environment Laboratory. Note the microprocessors dedicated to system timing and to real-time digital signal processing from two parallel quadrature receivers. Full system control for data acquisition, data processing, and user interaction is achieved through software in the general-purpose minicomputer.

The Need for Coordinated Measurements

The relationships among derivable and deducible properties are identical to the subject matter often termed theoretical modeling of the ionosphere and high atmosphere. Although particular links between boxes must often be developed in isolation, it is the degree of success of the entire pattern as a self-consistent system that expresses our state of knowledge. If important links are omitted in the modeling, or are inaccessible because of measurement system limitations, confidence in the entire structure suffers.

Second, specifically regarding total reflection sounding, we must emphasize the interdependent character of the observables in their relation to the derivables. Earlier sounding systems that, for example, measured ionospheric absorption, Doppler, direction-of-arrival, or drifts, at one frequency only, were exploratory and developmental efforts of pioneering value. However, it is only through doing all of these in concert that the modern digital ionosonde can be expected to perform adequately for the geophysical functions of Figure 3.

An extrapolation of this second point provides the main theme for the present paper. A very small number of digital ionosondes, however sophisticated, can not satisfy requirements for global monitoring of propagation, ionospheric, or geophysical conditions. But a threshold exists at which a moderate number of competent instruments, operating with real-time coordination, can do so. In addition to the vertical-incidence measurement capabilities just described, it is possible to use the

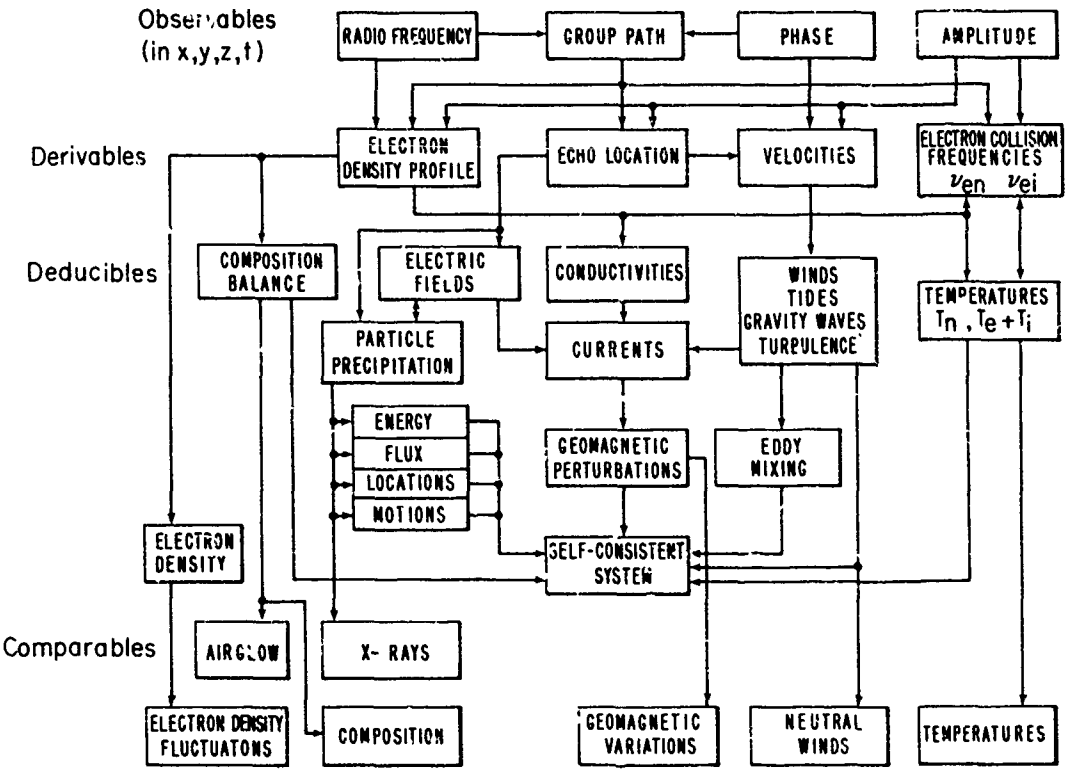


Figure 3. Geophysical functions of a digital ionosonde, showing relationships among the measurable quantities (observables) in ionospheric total-reflection radio sounding, and the geophysical properties of the ionosphere and atmosphere. Note the interdependence of the derivable and deducible properties on all of the observables.

same hardware and data processing facilities with relatively little modification for measurement of lateral variability in the ionosphere.

METHODS FOR LATERAL OBSERVATIONS

Accurate measurement of echo complex amplitude at a closely-spaced array of receiving antennas permits calculation of the direction-of-arrival of each echo (Paul et al., 1974; Wright and Pitteway, 1979b). If, in addition to the vertical-incidence transmitting antennas, there are available broad-band antennas producing low-elevation radiation in several azimuths, the same ionosonde can obtain a much greater lateral view by a variety of well-established or potentially valuable methods:

Measurement of the zenith and azimuth angles-of-arrival during ionogram or adaptive soundings provides information on the nature and magnitude of ionosphere tilts and horizontal gradients. Doppler information may provide an extra dimension of discrimination when multiple ray paths are known or suspected (Pfister, 1971; Brownlie et al., 1973).

Passive observations of "transmitters of opportunity" might provide useful supplementary information concerning lateral gradients and traveling disturbances, at the cost of data processing only. For remote transmitters of known schedule, location, and frequency stability, the digital ionosonde can measure the presence, field strength, Doppler, and angle-of-arrival of the signal. A minimum of about 0.5-sec observation per frequency should be sufficient for this, divided equally between the same two orthogonal pairs of receiving antennas that are used for vertical sounding.

With transmitting antennas that favor a particular azimuth and an intermediate zenith angle, oblique backscatter at frequencies exceeding the overhead critical frequency may be observed. About all that is seen by systems of ordinary power is the so-called leading-edge backscatter, the envelope of skip-distance or minimum-time-delay scatter, and even this is seldom available beyond half-paths of about 1500 km. Within the observable range, however, the trace provides nearly direct information on the variation of maximum density with distance from the observing location. Computer simulation and the development of backscatter data analysis is being pursued actively by the University of Illinois (DuBroff et al., 1978); backscatter measurements have been used to map foF2 within a radius of 1500 km by Hatfield (1970).

Bistatic oblique sounding over 1-hop paths provides a means for probing the ionosphere at the path midpoint. If transmission and reception in both directions is performed, the absolute group-path delay is the mean value of the two observed echo delays measured from the local transmission times. The only timing accuracy required is that which is sufficient to assure frequency synchronization and is of the order of the repetition interval. Inversion of oblique ionograms to midpoint N(z) profiles is not fundamentally different from the vertical sounding problem (Chuang and Yeh, 1977).

CONSIDERATIONS AFFECTING NETWORK DEPLOYMENT

Idealizations

The ratio of the Earth's surface area (radius R) to that observable from a single site (radius r) is $4\ R^2/r^2$. If r is taken to be 1000 km, 162 sites are required; 113 for r = 1200 km; and 72 for r = 1500 km. A satisfactory plan may be based upon the icosahedron by dividing each edge by three and adding one vertex near the center of each equilateral face. This produces isosceles triangles in each face, as shown by the sketch at lower right in Figure 4. If the Earth were divided in this way, 92 station locations would result, with three characteristic interstation distances: 2351, 2598, and 2722 km. These distances are all comfortable for bistatic oblique sounding of the F region. In addition to the 92 observation sites for vertical observations, the contiguous-pair midpoints provide 260 additional measurement locations. Each of the 12 icosahedral vertices "sees" five interpolated sites at equal azimuths of 72°, while each of the interpolated sites sees six sites at 60° azimuths. The locations are represented in Mercator projection in the main part of Figure 4.

It has appeared less than obvious to us that a global network of this size could conduct scheduled bistatic soundings (in a minimum cycle of six azimuths each) between pairs of stations without contention for the same observing period in the cycle. Each sounder must select the appropriate transmit-receive antenna for the azimuth of its neighbor, without being required to serve another azimuth during the same period. The small numerals of Figure 4, at the midpoints of each station pair, suggest the extent to which this may be accomplished. Contention appears at six of the 260 midpoints (wherever the number of midpoints in a closed loop is odd), but otherwise an efficient and compatible schedule for bistatic and backscatter soundings seems attainable.

New Standard Parameter Products From A Modern Network

We propose to initiate studies illustrated for N(h) profile data by the bottom panel of Figure 5. We assume that the profile has been obtained from virtual height-frequency data at frequencies that have been chosen adaptively by the ionosonde so that the desired detail is available in the profile. The figure defines a number of parameters with the following properties:

(1) They characterize the profile sufficiently well that the profile itself can be reproduced from them to an acceptable accuracy.

(2) They may each be summarized statistically into mean values and standard deviations (and possibly higher-order moments for suitable samples).

(3) In their statistically-summarized form, they permit reversion to an acceptable mean profile. This procedure is analogous to that proposed by Dudeny (1978), except that in our case the parameters are literally properties of the N(z) profile; in theirs, [illegible] ttempted to construct a profile using the old standard parameters [illegible] ariety of assumptions.

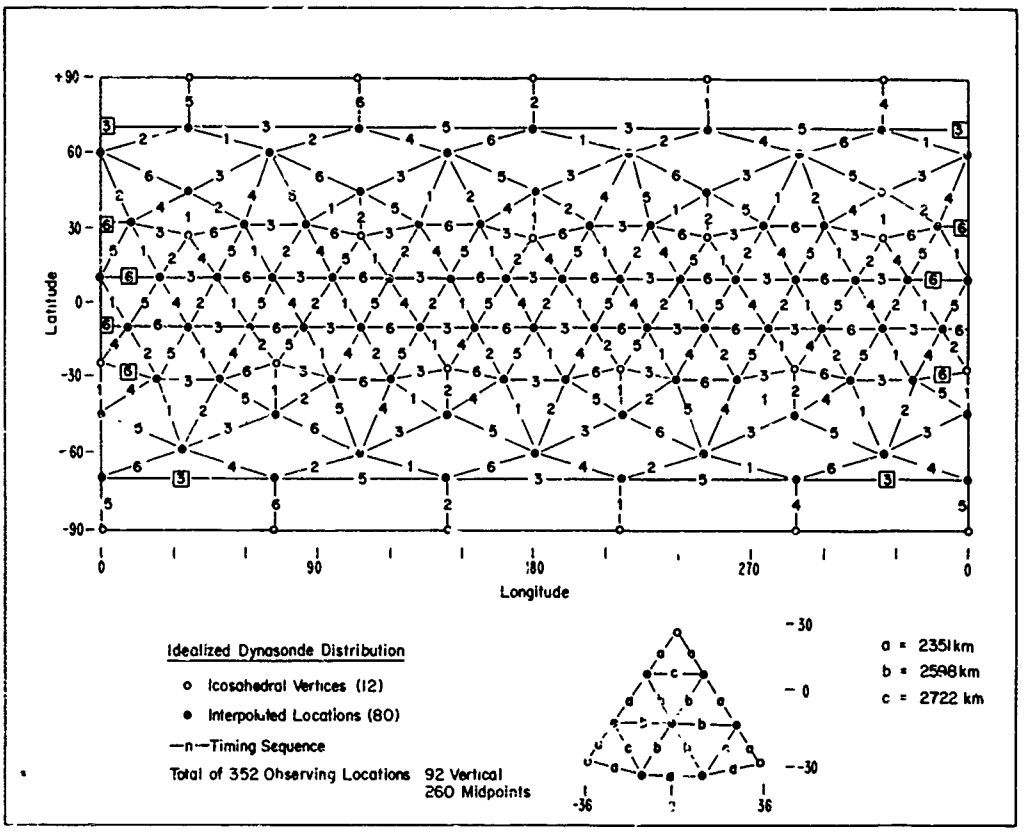


Figure 4. An idealized station deployment based upon the twelve vertices of an icosahedron (o), each of the 20 faces of which are further divided (•) into nine nearly equal isosceles triangles. Small numbers between sites suggest a possible schedule of bistatic soundings between adjacent station pairs in a cycle of six.

A similar approach would be necessary for other kinds of information gathered by the system.

Volume of Information From a Modern Network.

The modern network will not yield an unmanageable volume of data.

Suppose a modernized set of standard parameters were devised, somewhat along these lines, to summarize the height variations of electron density, $\Delta N/N$, $V_{x,y,z}$, and, say 3 or 4 other parameters. We then obtain about 80 parameters per recording, consisting of N(z), velocities, etc., at (say) 10 heights. Since at their source these are supplemented by some time-rate-of-change information, let us assume that an interpolation for intermediate times is satisfactory, yielding 160 bytes of data each 10 minutes. This is equivalent to about 10^7 bytes/year/station. If we assume that each bistatic sounding path also yields 10^7 bytes/year, then, at least, the 92 stations must archive for a total of 352 locations; the accumulation is 3.2×10^7 bytes/year/station. A 9-track tape at 6250 bpi can hold about 1.6×10^8 bytes, assuming 20% record gaps. Thus each station would fill only 0.2 tapes/year, or 18 tapes globally per year.

But in fact, it is desirable that each station archive data that describe the region around the station. Regions of three different sizes can be imagined:

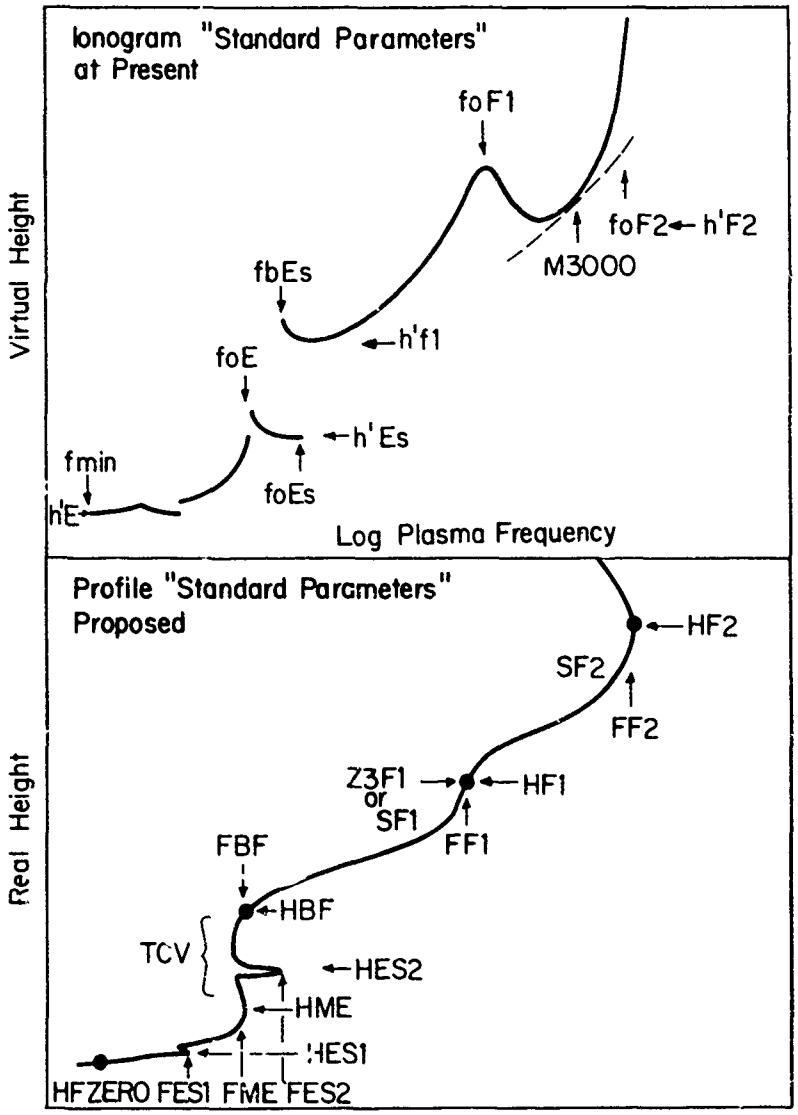


Figure 5. Top: The 11 standard parameters manually digitized from analog ionograms, exchanged internationally and archived in hourly tabulations at World Data Centers. Bottom: Some standard parameters that could serve to describe ionospheric electron density profiles and that could be summarized statistically.

i) The circle, of about 1300-km radius, from each station to the midpoints of the oblique paths to the neighboring stations; or

ii) The circle, of about 2600-km radius, to the neighboring stations; or

iii) Total global sharing of all available data, at all stations.

In the extreme case (iii), each of the 92 stations would accumulate reduced data at the rate of 18 tapes/year, but this means, globally, 92 copies of the same information, and, while manageable, seems quite unnecessary. If the intermediate case (ii) is considered, we have five or six copies of this information (depending on the number of nearest stations), so that only about one tape/year/station is required for regionally complete archives.

THE DEPLOYMENT OF A PRACTICAL MODERN NETWORK

Figure 6 provides global views of an optimum network of modern ionosonde instruments, based upon the requirements and technological opportunities reviewed above. At about 99 practical locations, we envision instruments capable of providing all of the detailed vertical incidence sounding information of the advanced digital ionosonde, plus additional information on lateral structure and gradients near each site, plus bistatic soundings between nearby pairs of sites.

Typical interstation distances in Figure 6 are about 2500 km, although many (e.g., Dublin-Prague) are less than 2000 km, and some, still useable for oblique soundings, are between 3000 and 4000 km (Seattle-Hawaii; Aricebo-Azores).

A list of the station locations of Figures 1 and 6 is given in Table 1. The 156 traditional stations, and 89 locations of the modern plan, have 39 places in common. Of the 24 mandatory island locations, five are currently instrumented. Eleven more of the 89 were once occupied by ionosondes and have become inactive. This overlap of the present network with our plan is intentional, reflecting our presumption that an existing station has a motivated administration behind it, experienced personnel, and a practical site. We should not wish our suggestion of optimum locations to express any deeper political considerations, and we hope that it is clearly understood that any rearrangement of locations that preserves approximately the suggested station density is equally satisfactory.

CONCLUSIONS

It is evident from the foregoing that we differ categorically with the views held by some (e.g., Bibl and Reinisch, 1978) that "the most important requirement for a monitoring ionosonde is low cost, since only a dense network of sounders can produce a global picture of the ionosphere" and that "two sets of ionosondes are needed, one for research and one for monitoring purposes."

We must emphasize however that there is no need for absolute uniformity among the digital ionosonde instruments comprising our proposed network. Stations offering occasional or frequent research facilities (or an ionosonde research center, in the manner of incoherent scatter centers such as Arecibo) will of course need relatively elaborate systems. Very isolated stations (gap-fillers) will need only a minimal system, to which additional hardware could be interfaced temporarily if the station becomes central to a geophysical campaign.

Some particular comments are applicable to educational and government administrations that have become accustomed to the operation of their own ionosonde and would find that it is not needed in our plan:

(1) Clearly an economy of radio sounding activity is desirable on grounds of efficient use of the radio spectrum and minimizing radio interference; these considerations alone may eventually make our network suggestions imperative.

(2) National administrations in Europe realized at the beginning of large-scale space research that a proliferation of individual national centers could be a weakening policy, as compared with the cooperative establishment of ESRO The same considerations prevail in the field of high-energy particle accelerators (Amaldi, 1979). A similar approach is worth consideration for ionospheric sounding centers.

(3) Modern communications permit large numbers of widely separated users to enjoy the use of one central, large computing facility, almost as if it were theirs alone. This can apply as well in multiple access to an ionospheric sounding facility, to obtain the latest data or a summary, or to take active control of a share of the ionosonde's measurement potential. Surely these possibilities, incorporating the advanced measurement ability of the center itself, must outweigh present advantages of operating one's own analog ionosonde.

(4) At an intermediate level we consider the educational institution which can fill a gap and can identify ionospheric research as an important specialization. To such institutions the modern ionosonde offers a small but complete data processing center, in addition to its primary sounding functions.

The most essential ingredient of our plan is again administrative, and as geophysicists rather than administrators we are unable to provide it. We refer to the planning, funding, and coordination required to bring about a network of the kind and capability described in this paper, on a global or regional scale in a reasonable time. Such a network is virtually certain to evolve eventually anyway, say within the next 20 years, because of the progressive obsolescence of present ionosondes and the continuing need for monitoring ionospheric weather. But we suggest that if a few national or multinational efforts (e.g., by NATO) were coordinated internationally by URSI, a regional demonstration network could be realized within five years. The merits of expanding or duplicating the demonstration network could then be readily assessed.

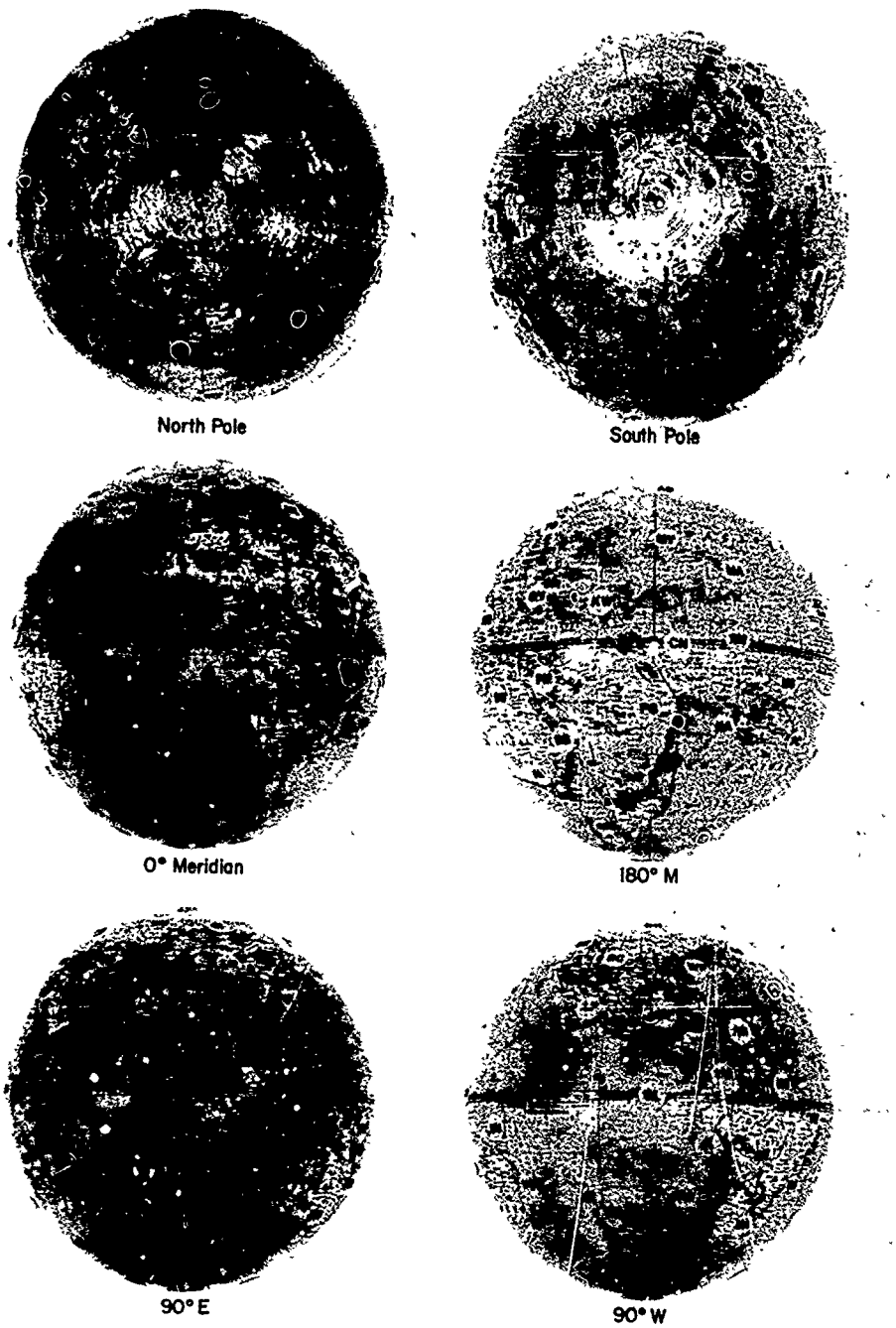


Figure 6. Suggested deployment of an optimum network, of about 90 modern digital ionosondes, also listed in Table 1. Each station would maintain an accurate three-dimensional local model applicable to a vertical cone of about the label radius, 500 km, at the F2 peak. Backscatter, bistatic oblique sounding, and other methods would maintain observations at about 250 intermediate locations.

Table 1. CURRENTLY ACTIVE IONOSONDE STATIONS AND A SUGGESTED OPTIMUM NETWORK

Map Code	Present Stations	Map Code	Suggested Stations	Map Code	Present Stations	Map Code	Suggested Stations
AB	Aberystwyth (Wales)			CV	Cape Zevgari (Cyprus)	CT	Capetown (S.Af.)
AC	Accra (Ghana)						
		AD	Adak (U.S.A.) P,M			CA	Carnarvon (Aust.)
		AL	Adelaide (Aust.)	CG	Changhan (China)		
				GH	Christchurch (N.Z.)		
AH	Ahmedabad (India)					CD	Chokurdakh (U.S.S.R.)
AK	Akita (Japan)						
AA	Alma Ata (U.S.S.R.)					XM	Christmas Isl. (U.K.) P,D,M
		AN	Angola	CH	Churchill (Can.)	CH	Churchill
		AK	Ankara (Turkey)				
						CS	Cocos Isl. (Aust.)
AR	Arıcebo (U.S.A.)	AR	Arıcebo				
				CO	College (U.S.A.)	CO	College
AI	Argentine Isl. (U.K.)	AI	Argentine Isl. M	CP	Concepcion (Chile)	CP	Concepcion
		AE	Ascension Isl. (U.K.) M	CG	Chongqing (China)		
AS	Ashkabad (U.S.S.R.)	AS	Ashkabad				
AT	Athens (Gr.)			DK	Dakar (Senegal)	DK	Dakar
AU	Auckland (N.Z.)	AU	Auckland				
						DS	Dar es Salaam (Tanz.)
		AZ	Azores (Port.) M			DW	Darwin (Aust.)
				DH	Delhi (India)	DH	Delhi
BK	Bangkok (Thai.)	BK	Bangkok	DT	DeBilt (Neth.)		
						DG	Diego Garcia Isl. (U.K.) D,M
BB	Baudouin (Belg.)	BB	Baudouin				
				DI	Dixon (U.S.S.R.)		
BN	Beiging (Peking) (China)			DJ	Djibouti (Fr.)	DJ	Djibouti
				DE	Dourbes (F.)		
GE	Belgrano (Argent.)					DN	Dublin (Ire.)
BE	Beograd (Yug.)			DU	Dushanbe (U.S.S.R.)		
BO	Billerica (U.S.A.)						
		BG	Bogota (Col.)			EI	Easter Isl. (Chile) D,M
						FG	Fiji (N.Z.) P,M
BM	Bombay					FA	Fort Archambault (Chad)
BC	Boulder (U.S.A.)	BC	Boulder				
		BV	Bouvet Isl. (Nor.) D,M	FZ	Forteleza (Braz.)	FZ	Forteleza
BB	Bribie Island (Aust.)			FM	Fort Monmouth (U.S.A.)		
BR	Brisbane (Aust.)	BR	Brisbane	FR	Freiburg (Ger.)		
						GL	Galapagos (Ecuad.) D,M
BU	Budapest (Hung.)						
BA	Buenos Aires (Argent.)	BA	Buenos Aires	GY	Garchy (Fr.)		
				GO	Godhavn (Den.)		
CU	Calcutta (India)			GS	Goose Bay (Can.)		
CI	Campbell Isl. (N.Z.)			GK	Gorky (U.S.S.R.)		
				GR	Grahamstown (S.Af.)		
CB	Canberra (Aust.)						
		CN	Canton (China)	GZ	Graz (Aust.)		
				GQ	Guangzhon (China)		
CF	Cape Parry (Can.)					GM	Guam Isl. (U.S.A.) P,M
CE	Cape Schmidt (U.S.S.R.)						

(Key: (P) = previously instrumented; (M) = mandatory location; (D) = difficult logistically).

Table 1. CURRENTLY ACTIVE IONOSONDE STATIONS AND A SUGGESTED OPTIMUM NETWORK (cont.)

Map Code	Present Stations	Map Code	Suggested Stations
HB	Halley Bay (U.K.)	HB	Halley Bay
HA	Haik'ou (China)		
HN	Hanover (U.S.A.)		
HE	Hermanus (S.Af.)		
HS	Highgate Springs (Aust.)		
HO	Hobart (Aust.)		
HK	Hong Kong (U.K.)		
HU	Huancayo (Peru)		
IB	Ibadan (Nig.)	IB	Ibadan
		HT	Camp Heurtin (Fr.) D,M
IR	Irkutsk (U.S.S.R.)	IR	Irkutsk
IS	Istanbul (Turk.)		
JI	Jicamarca (Peru)	JI	Jicamarca
JO	Johannesburg (S.Af.)		
JR	Juliusruh (Ger.)		
KL	Kaliningrad (U.S.S.R.)		
KR	Karaganda (U.S.S.R.)		
KG	Kerguelen Isl. (Fr.)	KG	Kerguélen Isl. M
KB	Khabarovsk (U.S.S.R.)		
		KY	Khaybar (Saudi Arabia)
KV	Kiev (U.S.S.R.)		
KT	Kiruna (Sw.)		
KO	Kodiakanal (India)		
		KU	Kourou (Fr.Guiana)
		KW	Kwajalein Isl. (U.S.A.) P,M
LA	Lancaster (U.K.)		
		LC	Lanchow (China) P
LN	Larnion (Fr.)		
LZ	Lanzhou (China)		
		LP	La Paz (Mex.)
LT	Leicester (U.K.)		
LD	Leningrad (U.S.S.R.)		
LI	Lindau (Ger.)		
		LB	Lisbon (Port.)
LY	Lycksele (Sw.)		
		MQ	Macquarie Isl. (N.Z.) P,M
MD	Madras (India)		
MG	Magadan (U.S.S.R.)	MG	Magadan
MN	Manila (Phillipines)		
		MR	Marion Isl. (S.Af.) P,M
		MS	Marquesas Isl. (Fr.) M
		MT	Mato Grosso (Braz.) D
MA	Maui (U.S.A.)	MA	Maui M
ML	Manzhouli (China)		
		MM	Mauritius M
MW	Mawson (Aust.)		
		MP	Merida (Mex.)
MX	Mexico City (Mex.)		
		MY	Midway (U.S.A.) M
MZ	Miedzeszyn (U.S.S.R.)		
MH	Millstone Hill (U.S.A.)		
		MQ	Mindanao (Phillipines)
MI	Mirny (U.S.S.R.)	MI	Mirny
MO	Moscow (U.S.S.R.)	MO	Moscow
MU	Mundaring (Aust.)		
MK	Murmansk (U.S.S.R.)		
		MC	Muscat (Oman) D,M
NR	Nairobi (Kenya)		
NQ	Narssarssuaq (Den.)	NQ	Narssarssuaq
NI	Norfolk Isl. (Aust.)		
NO	Norilsk (U.S.S.R.)		
NR	Novokazalinsk (U.S.S.R.)		
NS	Novosibirsk (U.S.S.R.)		
NU	Nurmijarvi (Fin.)		
		OM	Omsk (U.S.S.R.)
OK	Okinawa (Jap.)		
OT	Ottawa (Can.)		
OU	Oagadougou (Fr.)		
		PO	Palermo, Sicily (It.)
PK	Patrick AFB (U.S.A.)		
			Pitcairn Isl. (U.K.) P,D,M
PA	Point Arguello (U.S.A.)		
PT	Poitiers (Fr.)		
		PQ	Prague (Czech.)
PS	Port Stanley (U.K.)	PS	Port Stanley
		PR	Post Maurice Cartier (Alg.)
RA	Raratonga (N.Z.)	PA	Raratonga M
RB	Resolute Bay (Can.)		
		RJ	Rio de Janeiro (Braz.)
RO	Rome (It.)		
RV	Rostov (U.S.S.R.)		

Table 1. CURRENTLY ACTIVE IONOSONDE STATIONS AND A SUGGESTED OPTIMUM NETWORK (cont.)

Map Code	Present Stations	Map Code	Suggested Stations
SH	Sachs Harbor (Can.)		
SD	Salerkhard (U.S.S.R.)		
SY	Salisbury (Rhod.)	SY	Salisbury
QM	Sanae (S.Af.)		
		SA	San Ambrosio Isl. (Chile) D,M
BZ	San Jose (Braz.)		
SN	San Juan (Arg.)		
SB	Scott Base (U.S.A.)		
SM	Søndrestrømfjord (Den.)		
		SE	Seattle (U.S.A.) P
SU	Seoul (Korea)		
SL	Slough (U.K.)		
SO	Sodankyla (Fin.)		
SQ	Sofia (Bulg.)		
SG	South Georgia (U.K.)	SG	South Georgia
ST	South Uist (U.K.)		
PO	South Pole (U.S.A.)		
		BI	Stanleyville (Bangui)
SV	Sverdlovsk (U.S.S.R.)		
SW	Syowa (Jap.)	SW	Syowa
TT	Tahiti (Fr.)		
	Taipei (Taiw.)	TT	Taipei
TG	Tangerang (India)	TG	Tangerang
TQ	Tashkent (U.S.S.R.)		
TB	Tbilisi (U.S.S.R.)		
TA	Terre Adelie (Fr.)	TA	Terre Adelie
TH	Thule (Den.)	TH	Thule
TC	Thumba (India)	TC	Thumba
TI	Tiruchchirappalli (India)		
TX	Tixie Bay (U.S.S.R.)		
TO	Tokyo (Jap.)	TO	Tokyo
TK	Tomsk (U.S.S.R.)		
EB	Tortosa (Spain)		
TL	Townesville (Aust.)		
TW	Trelew (Argent.)		
		TT	Tristan da Cunha (U.K.) D,M
TR	Tromso (Nor.)	TR	Tromso
TS	Tsumeb (S.W.Af.)		
TU	Tucuman (Argent.)		
TZ	Tunguska (U.S.S.R.)		
UP	Uppsala (Sw.)		
UA	Ushaia (Arg.)		
VA	Vanimo (New Guinea)		
VO	Vostok (U.S.S.R.)		
WK	Wakkanai (Jap.)		
WP	Wallops Isl. (U.S.A.)	WP	Wallops Isl.
WS	White Sands (U.S.A.)		
WU	Wuch'ang (China)		
YA	Yakutsk (U.S.S.R.)		
YG	Yamagawa (Jap.)		
		SA	Yuzhno Sakhalinsk (U.S.S.R.)

REFERENCES

Amaldi, U. (1979), Particle accelerators and scientific culture. CERN 79-06. CERN-European Organization for Nuclear Research, Geneve.

Argo, P. E., and I. J. Rothmuller (1979), PROPHET, an application of propagation forecasting principles. In Solar-Terrestrial Predictions Proceedings, I: Prediction Group Reports. R. F. Donnelly, ed. U.S. Dept. of Commerce, NOAA/ERL, p.312.

Bibl, K., and B. W. Reinisch (1978), The universal digital ionosonde. Radio Sci. 13, p. 519.

Brownlie, D. D., L. G. Dryburgh, and J. D. Whitehead (1973), Measurement of the velocity of waves in the ionosphere: a comparison of the ray theory approach and diffraction theory. J. Atmos. Terr. Phys. 35, p. 2147.

Chuang, S. L., and K. C. Yeh (1977), A method for inverting oblique sounding data in the ionosphere. Radio Sci. 12, p. 135.

DuBroff, R. E., N. V. Rao, and K. C. Yeh (1978), Backscatter ionogram inversion by ray-tracing methods. RADC-TR-78-86.

Dudeny, J. R. (1978), An improved model of the variation of electron concentration with height in the ionosphere. J. Atmos. Terr. Phys. 40, p. 195.

Hatfield, V. E. (1970), Derivation of ionospheric parameters from backscatter data. In: Ionospheric Forecasting, AGARD Conf. Proc. #49.

Paul, A. K., J. W. Wright, and L. S. Fedor (1974), The interpretation of ionospheric radio drift measurements--VI. Angle-of-arrival and group path (echolocation) measurements from digitized ionospheric soundings: The group path vector. J. Atmos. Terr. Phys. 36, p. 193.

Pfister, W. (1971), The wave-like nature of inhomogeneities in the E-region. J. Atmos. Terr. Phys. 33, p. 999.

Thomson, R. L., and J. A. Secan (1979), Geophysical forecasting at AFGWC. In: Solar-Terrestrial Predictions Proceedings, I: Prediction Group Reports. R. F. Donnelly, ed. U.S. Department of Commerce, NOAA/ERL, p. 350.

Wright, J. W. (1969), Some current developments in radio systems for sounding ionospheric structure and motions. Proc. IEEE 57, p. 481.

Wright, J. W. (1975), Development of systems for remote sensing of ionospheric structure and dynamics: functional characteristics and applications of the "Dynasonde". NOAA ERL SEL 206, U. S. Dept. of Commerce, NOAA/ERL.

Wright, J. W. (1977), Development of systems for remote sensing of ionospheric structure and dynamics: The Dynasonde data acquisition and dynamic display system. Special Rep., U.S. Department of Commerce, NOAA/ERL.

Wright, J. W., and M. L. V. Pitteway (1979a), Real-time data acquisition and interpretation capabilities of the Dynasonde. 1. Data acquisition and real-time display. Radio Sci. 14, p. 815.

Wright, J. W., and M. L. V. Pitteway (1979b), Real-time data acquisition and interpretation capabilities of the Dynasonde. 2. Determination of magnetoionic mode and echolocation using a small spaced receiving array. Radio Sci. 14, p. 827.

☆ U.S. GOVERNMENT PRINTING OFFICE : 1982 0 - 367-001 : QL 3