AD-A233 797

Radio Science Meeting

Editor-in-Chief
John M. Goodman

Associate Editors

Publication

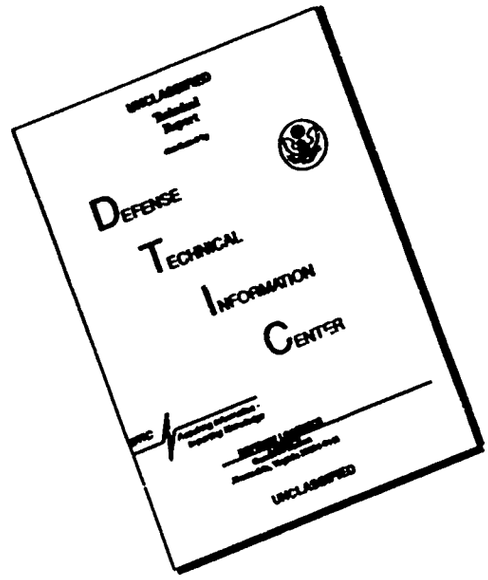
Issued by
Ionospheric Effects Symposium

May 1989

Published by
Army Center for Communications Systems

In Cooperation with the

DISCLAIMER NOTICE



THIS DOCUMENT IS BEST QUALITY AVAILABLE. THE COPY FURNISHED TO DTIC CONTAINED A SIGNIFICANT NUMBER OF PAGES WHICH DO NOT REPRODUCE LEGIBLY.

The Effect of the Ionosphere on Radiowave Signals and System Performance

Editor-in-Chief

John M. Goodman

Assoc. Editors

J. A. Klobuchar
R. G. Joiner
H. Soicher

Based on
Ionospheric Effects Symposium

1-3 May 1990

Jointly Sponsored by the

Naval Research Laboratory
Office of Naval Research
Air Force Geophysics Laboratory
and the
Army Center for Communications Systems

In Cooperation with the

Defense Nuclear Agency
Institute for Telecommunications Sciences (of NTIA)
Voice of America (of USIA)
Defense Communications Agency
and the
Naval Ocean Systems Center

SYMPOSIUM ORGANIZING COMMITTEE AND STAFF

Co-Chairmen:	Dr. John M. Goodman (NRL) Mr. John A. Klobuchar (AFGL) Dr. Haim Soicher (USA-CECOM) Mr. R. Gracen Joiner (ONR)
Other Members of the Steering Committee:	Dr. Patrick Crowley (DCA) Dr. Charles Rush (NTIA) Mr. George Lane (VOA) Major Michael Crawford (DNA) Dr. Juergen Richter (NOSC)
Coordinator:	Mr. Frank D. Clarke
Secretarial Staff:	Mrs. Rhea A. Smithson Mrs. Phyllis A. Kehres

Session Chairmen	Session No.	Title
Dr. John Goodman	1	Topical Reviews
Dr. Patrick Crowley	2	High Latitude Phenomena
Mr. John Klobuchar	3-1	Scintillation Effects
Dr. Charles Rush	3-2	HF Radiowave Effects
Major Michael Crawford	4-1	Radar System Effects
Mr. George Lane	4-2	HF, Meteor, and Misc. Systems
Dr. Haim Soicher	5-1	TEC and Modeling
Dr. Juergen Richter	5-2	Sounders, Probes and Radars
Dr. Haim Soicher	6-1	TEC and Global Positioning System
Mr. Gracen Joiner	6-2	Longwave Systems and Modification

SPONSORING AND COOPERATING AGENCIES

Sponsoring Activities

Naval Research Laboratory
Office of Naval Research
Air Force Geophysics Laboratory
U.S. Army Center for Communications Systems

Cooperating Activities

Defense Nuclear Agency
Institute for Telecommunications Sciences (of NTIA)
Voice of America (of USIA)
Defense Communications Agency
Naval Ocean Systems Center

PREFACE

The series of Triennial Ionospheric Effects Symposia which were inaugurated in 1975 have brought together a good mix of working scientists and radiowave propagation specialists on the one hand and system planners and DoD managers on the other. Although scientists and propagation specialists have appeared to dominate the events in terms of percentage participation and papers presented, the symposia would not have been successful without the support of the sponsoring activities and cooperating agencies and the vigorous participation of their representatives.

That which began as a singular Naval Research Laboratory (NRL) affair on a cold snowy January period in 1975 has become a government-wide conference which is held every third spring in the Washington, D.C. area. It retains its DoD roots with the primary symposium chairmen being representatives from Navy, Army and Air Force. Nevertheless, the Organizing Committee is drawn from other agencies including the Department of Commerce and the United States Information Agency. Although most of the attendance is drawn from U.S. citizens, a growing international flavor is observed. This is an indication of increasingly global views held by the scientific community as well as elements of various military establishments within the alliance.

The symposium welcome was given by Dr. Timothy Coffey, Director of Research at NRL, in behalf of the host agency. A copy of Dr. Coffey's remarks are provided herein.

We were privileged to have Dr. Thomas P. Quinn, Principal Deputy Assistant Secretary (C3I) of the Secretary of Defense, as the symposium keynote speaker. Dr. Quinn has had a considerable influence on ionospheric research activities within DoD over the years, specifically as it has related to C3I issues. He was a keynoter for the very first IES conference and, in his current address, he discussed the new environment including technical changes which have occurred, the political climate, and budget constraints. Dr. Quinn's address is contained within this volume.

On the evening of 2 May 1990, the traditional IES banquet was held, and we were treated to a fascinating talk by Dr. Gerald Hawkins, an internationally known author on topics of a wide-ranging nature. Dr. Hawkins is credited for developing the field of archaeoastronomy, and his after-dinner speech "Mysterious Effects and the Improbable Truth" was thoroughly enjoyable to a mixed audience of scientists and spouses.

The technical program for IES '90 was organized into eight sessions with two being held on the first day, and four being accommodated on the second and third days through the use of parallel sessions. The topics themselves concerned the full range of ionospheric research, related radio propagation and systems studies, and special interest areas such as ionospheric modification. The opening session was comprised of review papers on selected topics to provide attendees with a general overview of the more detailed technical papers to appear in later sessions. Session themes may be of interest and are provided on the page adjoining this one.

The Organizing Committee for IES '90 was comprised of the conference co-chairmen representing the sponsoring agencies and five other representatives from cooperating activities. Each member of the committee served as a session chairman as an additional duty. The designated session chairmen are listed in the frontis material along with an identification of the conference coordinator and supporting staff. Operational aspects of the conference were managed and executed by the Bendix Field Engineering Corporation (BFEC) under contract to NRL. Under this arrangement, Mr. Frank D. Clarke served as the conference coordinator.

The staff associated with IES '90 was directed by Mrs. Rhea Smithson, and she was ably assisted by Mrs. Phyllis Kehres and others assigned by the Space Science Division office. Mrs. Kehres handled the day-to-day operation and was largely responsible for assembling material for the preprint document (distributed at the symposium) as well as this book. This activity was not a trivial one and I would like to acknowledge Mrs. Kehres for a job exceedingly well done. Mrs. Smithson and her staff deserve a major portion of the credit for how well the conference was conducted. Others involved in this effort included Mrs. Jane Clarke, Messrs. Tom Priddy and Larry Quinn for visual assistance, and Mrs. Dora Wilbanks for preprint production and copying services.

Finally, speaking for the other co-chairmen and conference organizers, I would like to express my appreciation to all of the attendees and speakers for their participation in IES '90. It is hoped that this volume, not unlike those developed from proceedings of previous IES conferences, will be a useful reference for research and citation purposes.

JOHN M. GOODMAN
Editor-in-Chief

CONTENTS

Preface iii
John M. Goodman

Welcoming Address xi
Timothy P. Coffey

Keynote Address xiii
Thomas P. Quinn

SESSION 1

Some Recent Trends in Ionospheric Data Management at World Data Center-A 1
K. Davies, R. Conkright

Computerized Ionospheric Tomography 12
J.R. Austen, T.D. Raymund, C.H. Liu, S.J. Franke, J.A. Klobuchar, J. Stalker

Ionospheric Modification by Chemical Releases and High Power Radio Waves 24
P.A. Bernhardt, W.A. Scales, M.J. Keskinen, H.L. Rowland, L.M. Duncan

Forecasting Morphology and Dynamics of F-Layer Irregularities 39
Jules Aarons

Atmospheric Gravity Waves and Traveling Ionospheric Disturbances: Thirty Years of Research 46
R.D. Hunsucker

SESSION 2

The Trough in the Daytime F Layer: A Macroscopic Effect of Ionospheric-Magnetospheric Convection 60
J.A. Whalen

E-Layer Due to Pure Proton Aurora 69
B. Basu, J.R. Jasperse

Space-time Structure of Auroral Radio Absorption Events Observed with the Imaging Riometer at South Pole 76
J.K. Hargreaves, D.L. Detrick, T.J. Rosenberg

X-Ray Imagery of Energetic Electron Inputs to the Atmosphere at High Latitudes 82
W.L. Imhof, D.W. Datlowe, E.E. Gaines J. Mobilia, R.R. Vondrak, H.D. Voss, R.M. Robinson

Irregularity Anisotropy and Drifts from Polar Cap Scintillations During the Current Solar Maximum 92
Su. Basu, S. Basu, E. Costa

Goose Bay Radar Observations of Earth-Reflected Atmospheric Gravity Waves in the High-Latitude Ionosphere 100
J.M. Ruohoniemi, R.A. Greenwald, K.B. Baker, J.C. Samson

Computer Simulations of Plasma Structure, Dynamics, and Turbulence in the High Latitude Ionosphere with Magnetospheric Coupling 108
M.J. Keskinen, J.D. Huba, S.T. Zalesak, J.A. Fedder, H.G. Mitchell, P. Satyanarayana

SESSION 3-1

Scintillation Near the F-Layer Trough Over Northern Europe 118
L. Kersley, S.E. Pryse, C.D. Russell

Coherence Scales of Wavefield During Propagation Through Naturally Disturbed Ionosphere in the Polar Cap, Auroral, and Equatorial Regions 124
S. Basu, Su. Basu, R.C. Livingston

Theoretical Scaling Laws for the Spectrum of Density Irregularities in the High and Low Latitude Ionosphere 133
M.J. Keskinen

A General Channel Model for RF Propagation Through Structured Ionization 142
R.A. Dana, L.A. Wittwer

System Issues Related to Satellite Communications in a Nuclear Environment 152
P.A. Kullstam

SESSION 3-2

A Review of the History of Field Strength Models in PROPHET 160
D.B. Sailors

HF Noise Environment Models 171
C.M. Keller

Probability Distribution of Received Interference Levels in the HF Band 181
H.E. Nichols, D.J. Gooding

HF Channel Occupancy and Subband Congestion -- The Other User Interference Problem 189
R.H. Stehle, G.H. Hagn

Transfer Functions and Pulse Distortion for an Ionospheric Reflection Channel with Embedded Random Irregularities 201
T.W. Rand, K.C. Yeh

Knowledge-Based Fading Channel Analyzer for Adaptive Radio Links 209
J.S. Yalowitz, F.G. Baird, J.M. Morrison

Propagation Effects in Extended Random Media 219
L.J. Nickisch

SESSION 4-1

The Impact of the Solar Cycle on Over-the-Horizon Radar Systems 229
J.S. Schleher

VHF/UHF Radar Scintillation Effects on ALTAIR 237
D.L. Knepp, H.L.F. Houpis

An Ionospheric Propagation Prediction Method for Low- and Mid-Latitudes 247
H.L.F. Houpis, L.J. Nickisch

Effects of Ionospheric Scintillation on DPCA Clutter Rejection Performance 257
J.T. Reinking, T.N. Moody, D.L. Knepp

UHF Radar Backscatter Enhancements from the Topside Ionosphere 266
K.M. Groves, M.C. Lee, J.C. Foster

Ionospheric Effects Related to the Performance of an Orbital Debris Radar System 273
D.F. Bishop

SESSION 4-2

Updated Climatological Model Predictions of Ionospheric and HF Propagation Parameters 281
M.H. Reilly, F.J. Rhoads, J.M. Goodman, M. Singh

Upgrades for Efficient 3D Ionospheric Ray Tracing - Investigation of HF NVIS Effects 290
M.H. Reilly

Angle of Arrival Observations of HF Radio Waves 300
Z.R. Jeffrey, L.F. McNamara

HF/VHF Propagation Resource Management Using Expert Systems 313
D.P. Roesler

Frequency Selection and Interference Effects for Meteor Burst Systems 322
D.J. Cohen

Effects of Absorption on High Latitude Meteor Scatter Communication Systems 329
J.C. Ostergaard, J.A. Weitzen, P.A. Kossey, A.D. Bailey, P.M. Bench, S.W. Li, J.R. Katan, A.J. Coriaty, J.E. Rasmussen

SESSION 5-1

Evaluation of Six Ionospheric Models as Predictors of TEC 343
L.D. Brown, R.E. Daniell, Jr., M.W. Fox, J.A. Klobuchar, P.H. Doherty

A Global Ionospheric Conductivity and Electron Density (ICED) Model .. 351
R.E. Daniell, Jr., D.T. Decker, D.N. Anderson, J.R. Jasperse, J.J. Sojka, R.W. Schunk

Ionospheric Errors at L-Band for Satellite and Re-Entry Object Tracking in the Equatorial Anomaly Region 360
W.A. Pakula, J.A. Klobuchar, D.N. Anderson, P.H. Doherty

The Latitudinal Extent of the Equatorial Anomaly 366
J.A. Klobuchar, D.N. Anderson, P.H. Doherty

A Comparison of Ionospheric Slab Thickness at Widely Spaced Stations in Middle Latitudes and Low Latitudes 378
K. Davies, X.M. Liu

The Variation of Ionospheric Slab Thickness During Geomagnetic Storms 388
M.W. Fox, M. Mendillo, P. Spalla

Reduction of Ionospheric Effects on Polarization Measurements During Satellite Tracking 396
C. Yoo, M.C. Lee

SESSION 5-2

Ionospheric Observations in Western Europe 401
F.J. Gorman, H. Soicher

Applications for High Accuracy Digital Ionosonde Data 410
A.K. Paul

Oblique-Incidence-Sounder Measurements with Absolute Propagation Delay Timing 418
M. Daehler

Probing the Ionosphere with the University of Illinois Sounder 427
K.H. Lin, W.L. Schlegl, K.S. Yang, K.C. Yeh

Determination of Daytime Midlatitude Electron Density Profiles from Limited Real Time Measurements 436
D.T. Decker, J.M. Retterer, J.R. Jasperse, D.N. Anderson, F.J. Rich, J.C. Foster

Using Ionosondes for Global Mapping of Thermospheric Meridional Winds 442
K.L. Miller, P.G. Richards, D.G. Torr

A Survey of Polar Cap F-Region Electron Densities Measured by the Sondrestrom Radar 452
R.M. Robinson

SESSION 6-1

Comparison of GPS and Incoherent Scatter Measurements of the Total Electron Content 460
A.J. Coster, E.M. Gaposchkin, L.E. Thornton, M. Buonsanto, D. Tetenbaum

Measurements of Transionospheric Effects Using Signals from GPS 470
G.J. Bishop, J.A. Klobuchar, S. Basu, J.R. Clynch, D.S. Coco, C. Coker

Ionospheric Delay Measurements Using GPS Satellite Signals 481
J.D. Holmes, R.L. Meyer

Ionospheric Predictions - II 491
N.C. Gerson

SESSION 6-2

Status of the Naval Ocean Systems Center's Long Wave Propagation Capability 505
J.A. Ferguson

3-D Modeling of Subionospheric VLF Propagation in the Presence of Localized D-Region Perturbations Associated with Lightning 515
W.L. Poulsen, T.F. Bell, U.S. Inan

VLF/LF Remote Sensing of Transient and Localized Disturbances of the Lower Ionosphere Associated with Lightning Discharges 523
U.S. Inan

Observations of the Characteristics of Propagation of VLF Signals During Meridian Transits by SCO X-1 533
J.H. Beall, K.S. Wood, F.J. Kelly

HF Propagation Through Actively Modified Ionospheres 542
P.E. Argo, T.J. Fitzgerald, J.H. Wolcott, D.J. Simons, S. Warshaw, R. Carlson

Ionospheric Modification Experiments with Satellite-Borne and Ground-Based Transmitters 550
M.C. Lee

Wideband Noise Observed at Ground Level in the Auroral Region 561
R.F. Benson, M.D. Desch

Author Index 569

WELCOMING ADDRESS - IES '90

Dr. Timothy Coffey
Director of Research
Naval Research Laboratory
Washington, DC

Good Morning! I am very pleased to welcome you to the Sixth Ionospheric Effects Symposium. The theme of this symposium is the Effect of the Ionosphere on Radiowave Signals and System Performance. The broad interest in this topic is evident from examining the symposium sponsorship. The symposium is sponsored by the Naval Research Laboratory, the Office of Naval Research, the Air Force Geophysics Laboratory, and the Army Center for Communications Systems, and is presented in cooperation with the Defense Nuclear Agency, the Institute for Telecommunications Sciences, the Voice of America, the Defense Communications Agency, and the Naval Ocean Systems Center.

As I reviewed the symposium program I found that many of the topics are similar to those addressed in the first Ionospheric Effects Symposium held in 1975. It became clear, however, on reviewing the actual papers in the proceedings, that the level of sophistication has increased greatly since 1975. The understanding of the microscale processes and ionospheric instabilities, for example, has progressed significantly. The use of large-scale numerical models, and even the use of global models for studying system performance, are becoming relatively routine. I believe this particular series of symposia can take a lot of credit for having brought these communities together over the past fifteen years. As we look towards the next fifteen years, I believe we will see very significant progress in the area of global monitoring of the ionosphere through the use of techniques, such as ultraviolet imaging of the ionosphere. These future developments, combined with the groundwork which has been and is being laid by the community here present, I believe, will lead to still more sophisticated exploitation of the ionosphere for enhanced systems performance and enhanced systems capability. On behalf of the Naval Research Laboratory and the symposium sponsors, I wish you the greatest success in this undertaking.

"IONOSPHERIC RESEARCH AND C3, A STATUS REPORT"

KEYNOTE ADDRESS - IES '90

Dr. Thomas P. Quinn
Principal Deputy Assistant (C3I)
Department of the Secretary of Defense
Washington, DC

Good Morning! Let me add my welcome to the Sixth International Ionospheric Effects Symposium. As some of you may recall, I spoke at the First Ionospheric Effects Symposium held in 1975, and I believe it was attended by many of the same people that are here today. When John Goodman and I discussed the current meeting, I re-examined the talk I gave at the previous one, and then I called John and said, "You know, that's still pretty much a valid set of comments." So let me review for a moment what I said then.

At that time, satellites were just coming into their own, and we had in fact just established an operational capability with UHF satellites. There had been predictions by many people that high frequency (HF; 3-30 MHz) was on the way out, but as you know that never did happen. I mentioned all the advantages of satellite connectivity, as well as the drawbacks and it's interesting that some of them are still with us today. Scintillation was one that we mentioned then and it still is of concern to us. Another important factor is the expense of satellite systems, although if you are not too concerned about extremely capable warfighting performance, you can get by with cheaper terminals and obviously cheaper systems. But when you try to make a system resistant to the gamut of electronic attacks as well as physical attacks, you run into some extremely expensive undertakings as we have discovered with the MILSTAR system. So, as I suggested to you back in 1975, the switch to satellite communications, although it has revolutionized the way we conduct our command, control, and communications (C3) business, is not a total cure-all and we still have some important shortcomings to overcome. Another important step in the progress we have made is that results of ionospheric and upper atmospheric research are finding their way into more applied arenas than was the case a decade or so ago. In fact, looking at the proceedings of the previous conference on this topic, the fifth conference held in 1987, I note that most of the papers were devoted to examining the effects of the ionosphere on systems and on various attempts to do something about it.

Again, looking back at the inaugural conference, it's really hard to imagine that fifteen years have gone by since this group began discussing the effects of the ionosphere on systems. The fact that the conference is still on-going and doing well and I still seem to be around (and both of us are expanding) must mean that we are doing something right. As I said, the first conference was held in 1975, this is the sixth one, and the way we do business has changed considerably over that period. Let's review how much change has really occurred.

I think if you go back a little earlier than 1975, the time when I was involved in ionospheric research, the motivation then was mostly to support HF

communications. There was some concern for the propagation of other radio frequencies through the ionosphere because of its effect on those signals and in some cases using those effects to infer ionospheric behavior. In some cases where radio signals were propagated to and from rockets and satellites, there was interest in correcting for ionospheric effects. There were a lot of uncertainties in the models that were used then to predict both the ionospheric parameters and the propagation parameters. I think a great deal of progress has been made in both of those areas and, if you look at the types of work being reported here, we are now shifted considerably to the applied end of the spectrum as opposed to concentrating on fundamental or basic research activity. That is not to say that all of the basic problems have been solved or that our understanding of the ionosphere and its response to many influencing forces is totally understood. But with the advent of extremely powerful computers and simulators that are now available and the time and effort of people like yourselves, we have certainly come a long way to narrowing the degree on those uncertainties. We can now apply model results with much more confidence. It is also interesting that the area of simulation has greatly matured and we now can study the effect on systems without actually having to operate the systems in the presence of the ionosphere itself. Of course much of the earlier work had difficulty in just dealing with a massive data reduction problem. That has largely been overcome. Just doing the propagation calculations in enough detail to have confidence in the predictions has been considerably improved. We seem to have an understanding of such effects as scintillation and have moved toward developing systems, coding schemes, and modulation techniques that are capable of operating in the presence of scintillation or at least in reducing the negative effects of scintillation. So we have moved from an emphasis on attempting to understand the physics and chemistry of the ionosphere and its interaction primarily with high frequency radio waves to where we have a reasonable understanding of the physics and chemistry and are working toward applying the results of research to designing systems. Because of the tremendous advances in technology, both computers and microcircuits, we can now apply coding and modulation techniques to minimize the negative effects of the ionosphere. Just look at the list of sponsors here and one observes that a great deal of interest is being expressed by the military communities in engineering the results of ionospheric research.

On the other hand, although we are probably asymptoting the goal of understanding the behavior of the ionosphere under various diurnal conditions, solar cycle effects and so on, graduate students need not worry that thesis topic material will diminish. Be assured that each new revelation also raises a new question and there will remain much about the ionosphere to study far into the future. Just look at the weather and lower atmospheric research community whose programs continue to expand. And they make direct observations and measurements. I would guess that at the twelfth conference on Ionospheric Effects, the parameters that describe the behavior of the ionosphere will be much better refined. But I would also guess that by then EHF systems will be commonplace and most effects will be routinely taken into account. But there will no doubt also be things we haven't discovered yet that will occupy the attention of the then ionospheric research community. Great progress is being made, but it's slow.

Since I have nothing technical to add to our understanding of today's subject and you really have come together for technical discussions, I want to move away from that topic and talk a little bit about where we have come in the last decade in C3 systems and where we are headed.

When we entered the decade of the 80's, we had a number of severe deficiencies in the C3 business. Strategically, we were faced with the issues of survivability and endurance. There was inadequate anti-jam performance and our systems were unable to operate through a high altitude electromagnetic pulse (HEMP) event. In the tactical area, we were plagued with interoperability problems, between our air, ground, and naval systems. Our worldwide backbone communications systems were rapidly aging and we also had an interoperability problem both nationally and across the NATO alliance. In the early days of the Reagan administration that situation changed, or at least the attitude to do something about that situation changed. New command, control, communications and intelligence (C3I) research, development, and production became a cornerstone of the defense strategic modernization program. In fact, C3I from then on consistently enjoyed the highest priority within the general defense budget. That gradually resulted in our developing and procuring sophisticated, highly capable systems that are now being fielded with the U.S. forces and, in some cases, with NATO. The systems that resulted from the attention and priority given during that period went a long way towards remedying the deficiencies of the previous decade. Among the strategic C3 systems that were fielded during that era were the Ground Wave Emergency Network (GWEN) radio relay network, designed to help solve the HEMP problem.

The development of LF/VLF Miniature Receive Terminals installed on our strategic bombers improved communications from our airborne command posts. The ability to communicate with the strategic submarine force will be improved as the aging EC-130 TACAMO aircraft are replaced with the newer, more enduring E-6A aircraft. The Military Strategic and Tactical Relay (MILSTAR) satellite communications system, making use of advanced technology and survivable ground-based control stations, will combat degradation from scintillation and blackout and provide significant anti-jamming capabilities. And, finally, the deployment of the Automated Emergency Processing & Dissemination System (AEPDS) will significantly reduce the time required for messages to reach the executing strategic forces.

The capability, survivability and endurance of our command centers have been greatly improved. Our National Military Command Center (NMCC) and its alternate have been upgraded by providing protection against electromagnetic pulse effects, by modernizing their power systems and by improving their ability to rapidly receive, process, and transmit information. Our National Emergency Airborne Command Post (NEACP) aircraft have been hardened against electromagnetic pulse (EMP) effects and they have had their communications systems upgraded. Similarly, both the fixed and airborne command centers of our nuclear capable Commanders-in-Chief have been upgraded.

On the tactical front, new interoperable systems such as the Single Channel Ground and Airborne Radio System (SINCGARS) and Mobile Subscriber Equipment (MSE) enhance interoperability among the maneuver forces. Also, interoperability will be enhanced through the fielding of the Joint Tactical

Information Distribution System (JTIDS), the U.S. Air Force's Modular Control Equipment (MCE) and the U.S. Marine Corps' Tactical Air Operations Modules (TAOM).

Tactical sensors are also being improved to provide commanders with more powerful "eyes and ears." Additional E-2C Hawkeye aircraft are being fielded to provide early warning of approaching enemy air and surface units. The E-3 Airborne Warning and Control System (AWACS) aircraft are now being used to provide a deep look at both the air and surface environments, and upgrades have begun to enable it to respond effectively to the future threat. Our current air surveillance radar is being replaced with the AN/TPS-70 which has an Ultra Low Sidelobe Antenna system for resistance to hostile jamming. Selection of a new ground based sensor for the Forward Area Air Defense (FAAD) system is also well underway.

Improvements to our navigation and position-fixing capabilities will be provided by the NAVSTAR Global Positioning System (GPS). It provides precise three-dimensional position, velocity, and time data for weapons platforms. It will replace TACAN, LORAN, OMEGA and other less effective and more costly systems. The Position Location and Reporting System (PLRS) is in production to provide Army and Marine Corps commanders with automatic real-time location of their forces on the battlefield. The Enhanced PLRS, which includes a data communications capability. and the JTIDS together comprise the Army Data Distribution System (ADDS). This provides the data communications necessary for FAAD command and control (C2).

Efforts to improve interoperability with our NATO allies have been boosted by the Nunn Initiative. Efforts underway within this program include the Post 2000 Architecture, the NATO Identification System (NIS), and the Multifunctional Information Distribution System (MIDS). The Post 2000 Architecture is intended to insure that future communications systems developed by members of the Alliance will be fully interoperable. The NIS will result in an advanced aircraft identification system designed to prevent fratricide. MIDS is the NATO version of the U.S. JTIDS Program that will link the air defense weapons and C2 facilities TO forces. NIS and MIDS are part of our participation in a much larger NATO ..ct, known as the Air Command and Control System (ACCS), designed to upgrade the overall air defense C3I capability of the Alliance by integrating offensive and defensive air operations. Finally, we are participating in an initiative to improve the sharing of sensitive tactical intelligence data called the Battlefield Information, Collection and Exploitation System, or BICES.

Without doubt, the watershed changes now taking place within Eastern Europe will have a profound effect on the structure and priorities of the U.S. forces well into the next century. Clearly, barring some unforeseen circumstance, the U.S. will be fielding a much smaller, more mobile force, capable of bare-base operations with a view towards low intensity operations in the lesser developed areas of the world. Operations of this nature will put a premium on advanced C3I capabilities, and we therefore expect that C3I will continue to enjoy a privileged position within the Department of Defense. However, this position must be taken within the context of generally declining defense budgets. We

will be rigorously disciplining our requirements and identifying promising technologies early in the development process. We will be doing more with less.

In the context of the foregoing, I would like to say a little about the budget and our planning process for the future. Historically, we have submitted a five-year defense plan to the Congress based on an assumption of real spending growth in the out years. We have changed that this year! For the first time we have submitted a budget that projects real reduction in defense spending in the years immediately ahead. The biggest change in terms of force levels has to do with the developments in Europe. That's the one place where there is justification for thinking anew, specifically with regard to the forces stationed there and those in the United States committed to the reinforcement of Europe. With that in mind, the Secretary of Defense provided the services with fiscal guidance for this year consistent with the changed political and military environment.

The first budget submitted by the new Secretary last Spring and approved by the Congress would cost $317 billion today in budget authority. However, we are not asking for $317 billion in the current budget; the request is for $295.1 billion, a reduction of $22 billion for FY 1991. That represents a cut of 2.6% in real terms. For the remaining years through 1994, there is a trend of a 2% drop in real dollar spending. If you consider the budget submitted in January 1989 and then look at the revision made by President Bush for 1990 through 1994, the gap between the two comes to approximately $231 billion. That is the amount that has been cut out of the five year defense plan since Secretary Cheney took office last spring.

Another way to look at the budget is the defense share of the gross national product. During the peak Vietnam years, that was 9.1%. During the Carter years it was 4.7%, rising to 6.3% during the peak Reagan years. The 1991 budget sent to Congress is at 5% of GNP, and, if we follow the 2% per year reduction plan for defense spending, by 1995 we will be at 4% of GNP, the lowest level of defense spending as a product of GNP since before World War II.

So, as I said before, we are going to do more with less and the challenge is how to accommodate that. In the C3I area, a main frustration for all of us is the inability to exploit the tremendous technical breakthroughs occurring almost daily by capturing them in system designs that will greatly assist military commanders to better do their job. We know how to design new systems and improve existing systems to give them manyfold increases in performance capability. But the development acquisition bureaucracy prevents rather than facilitates getting these new capabilities into the field before they are obsolete. I wish I had the answer to that problem but I don't. However, I do observe that the same problems seem to occur in almost all information handling system development programs. The requirements are not adequately defined at the outset. And that's not that people don't try, but sometimes it's just not possible to anticipate how a system really will be used. What people thought were the requirements are not really what the user wants. A possible remedy is to provide the user with a modicum of baseline capability, let him use it and determine what it should do that it doesn't, and then feed that information back to the developer. This technique has been referred to by various names: build-a-little/test-a-little, being one; the current approach is called evolutionary

acquisition. It means what it says. Get the user involved in the acquisition at a very early stage and provide him with a limited prototype system as soon as possible. Then allow it to evolve to meet the user's needs as determined in the field and not envisioned in a Pentagon office. There obviously have to be boundaries on this open-ended approach and we're aware of that. You have to know when you're finished. You have to be able to test the system to see that you are getting what you paid for and that it does perform under stressed environments. But these are constraints that can be accommodated by reasonable people. So we are pushing for this approach in acquiring C3I systems in the future and have included provisions for this in the latest version of DoD Direction 5000.1, the bible for defense acquisition.

Another part of the problem has to do with overwhelming layers of management. Too many people who can say no. The Defense Management Review dealt with that problem and some very positive results have occurred. One of the earliest efforts within the Defense Management Review was to reduce the number of regulations, directives, and other inhibiting constraints that are in the way of getting the job done. I forget the exact number, but hundreds of directives have either been eliminated or combined with others or simplified so that only regulations that really are essential will remain in a few months and contractors should find it easier to do business with the Department of Defense. Then if we can get some of the energy and attention that had been devoted to coping with that kind of bureaucracy and convert it into true production energy, we will have improved the efficiency significantly. In addition to removing the obstacles in the way of getting from a concept to acquiring hardware and software systems, the acquisition management chain has been streamlined. A new management chain is now in effect in each of the Services and in the Office of the Under Secretary of Defense for Acquisition. In case there is someone here who hasn't heard of this new arrangement, I will describe it in its simplest form. The Under Secretary of Defense for Acquisition is the Chief Executive Officer and the single person responsible to the Secretary of Defense for all acquisition matters. The program managers have been given much more responsibility and authority to do their jobs, and their reporting chain has been extremely simplified. A program manager now reports to a program executive officer who in turn reports to the Service Acquisition Executive. The Service Acquisition Executive reports directly to the Under Secretary of Defense for Acquisition, Mr. John Betti. So there are only two people between the program manager and the Chief Executive Officer for Acquisition in the Department of Defense. This is a major change in the way of doing business and, as you might expect, because it is quite different and does tread on some existing turf, we would expect some growing pains in getting it fully established and working. However, I might say that all of the people involved in putting this new system into effect from the Secretary of Defense on down are dedicated to making it work and will be watching it very carefully over the next year or so that it will probably take to see any major impact.

So we have a new management arrangement, much streamlined and hopefully much more efficient and effective, and it couldn't have come at a better time. As I said earlier, the new team is going to have to deal with a markedly different budget environment than has existed for the past eight years. The Defense budget has enjoyed a period of growth for the past eight years or so and our particular part of that budget has been equally well treated. The C3 part of

the budget has increased from about six percent in the early 1980's to well over seven percent of the defense budget as of last year. A great deal has been accomplished as a result of that support, much of which is coming to fruition now or will soon come to fruition in terms of systems actually being deployed. However, there will be a flattening out of that growth curve and in fact we will see either zero growth or slightly negative real growth, beginning with the current budget. So the challenge of the new management team and people like yourself is to do more with less, a phrase that is becoming overused perhaps, but we must recognize reality and react accordingly.

So we are involved in a very dynamic era for the Department of Defense. The threat is changing and, thankfully, in the right direction. That means the fiscal environment is also changing and rightfully so. Our challenge is to adapt to these changes without diminishing the effectiveness of our military forces. With the help of groups such as this and the right attitude on all our parts, I'm sure we're up to that challenge. Thank you for your kind attention.

AD-P006 265

Some Recent Trends in Ionospheric Data Management at World Data Center-A

Kenneth Davies
Space Environment Laboratory
National Oceanic and Atmospheric Administration
Boulder, Colorado 80303

Ray Conkright
National Geophysical Data Center
National Oceanic and Atmospheric Administration
Boulder, Colorado 80303

Introduction

During the past decade or so, a number of changes have taken place in the management of ionospheric and related data in the World Data Center-A at the National Geophysical Data Center in Boulder, Colorado. In this paper some of these changes are discussed on the following topics: (1) advances in computer technology, (2) some pitfalls in data usage, (3) types of data, and (4) examples of available data.

1. Advances in Computer Technology

In recent years there has been a shift from using large, centralized computers to using personal computers for digital data storage and retrieval. This change is driven by advances in less expensive but faster IBM-compatible PCs with compact data storage (e.g., CD-ROM) and user-friendly utility programs for data display and analysis. Researchers with little or no programming experience can, for the first time, handle large data sets (500 megabytes or larger) on their relatively inexpensive personal computers ($3,000-$3,000) and can use free or inexpensive utility programs to manipulate and analyze large amounts of data.

These new methods of data storage and retrieval make the large ionospheric data bases on CD-ROM readily accessible to users who are not expert in retrieval of data from large, mainframe computer systems. In earlier times, the *de facto* philosophy forced by lack of funds and technology was to store all the data in the format in which it was received from various sources. The Data Center did not attempt quality control, test for data retrievability, or review the documentation received with the data. The lack of basic documentation often made it difficult for a user to adequately assess the characteristics of the data, e.g., time used (UT or LT), scales, station location, and equipment used to gather the data. Important facts known to the data collectors were not passed on to prospective users.

Today the situation has changed, as incoming data are promptly transferred to intermediate, high-density storage media and processed for storage on optical discs. Emphasis is now placed on complete documentation of each disc; users will find it easier and faster to deal with the data. Data sets are being put to a "20-year test." To pass this test a data set must be usable and understandable by an investigator 20 years after its creation, without reference to external documentation.

91-09658

91

2. Some Pitfalls in Data Usage

2.1 Data Users

Many users are relative novices in the collection, processing, and display of data. They generally assume that numbers appearing in tabulated data reflect natural phenomena, but this is not always the case. For example, a trace may be wrongly identified or, in phase measurement systems, the number of cycles can be in error. In some methods, extensive data processing may be interposed between the raw data and the numerical output and may involve questionable or invalid assumptions or introduce systematic errors.

2.2 Primary and Secondary Users

We define a primary user as the individual or group that makes and analyzes the measurements, whereas a secondary user is a person or group that uses data deposited in a data center by other investigators. Primary users have access to original records, equipment parameters, and calibrations, and have knowledge of the advantages and disadvantages of the technique used. Secondary users, in contrast, are at the mercy of the "experts," who provided their measurements and processed the records before sending numerical data to a data center.

Even the most conscientious worker can make an error in the measurement of an ill-defined quantity, e.g., $f_{\min}$. As a consequence, within the ionosonde community, the use of qualifying symbols has emerged to indicate the degree of confidence a scaler has in a particular numerical value (Piggott and Rawer, 1978). Such a symbol may have considerable utility in the determination of a monthly median. For example, it is more useful to know that an off-scale value was actually below a lower limit or greater than an upper limit than to simply know a blank space in the data record.

While these qualifying codes have been used extensively in the scaling of analog data (e.g., ionograms), corresponding codes have yet to be developed for automatic scaling procedures such as those for scaling ionograms (e.g., see Reinisch, 1986; Gilbert and Smith, 1988; and Reinisch et al., 1990).

2.3 User Education

An educated user community is essential for the proper application of ionospheric data because of the great spatial and temporal variability of the ionosphere. Ionospheric quantities such as $N_m F_2$ can vary by orders of magnitude from day to night, with sunspot number, and with geographical location. Hence, a user should know what data set is needed, its reliability, the quantity involved, the meaning of terms, etc. However, data centers often receive requests such as, "send all the data you have from all the stations." It would be impractical to fill that request as stated. Not only would the data overwhelm the recipient, but sending it might well tax the resources of the data center. It is much easier to provide scaled hourly values already in digital form than to supply the corresponding originals (e.g., ionograms). Ionogram scalings provide examples of the difficulties involved in transferring data from one form (photograph) to numerical tabulations. The basic idea underlying ionogram scaling is to provide data sufficient for a user to visualize the original record to some extent. However, from a set of standard characteristics, the tabulated data cannot be an accurate description of the ionogram except, possibly, under the simplest conditions. In the presence of horizontal gradients, multiple layers, and spread echoes, the tabulated values give an inadequate description. Thus, for detailed study of a particular situation, there is no substitute for an examination of the original records by an interested, knowledgeable individual.

3. Types of Data

3.1 Accurate Data Versus Representative Data

Users need to be aware of the difference between the accuracy and representativeness of ionospheric data. The ionosphere changes with time; a particular technique may be capable of measuring a certain characteristic with high accuracy, but this is of little consequence if that characteristic has changed in value during the course of the measurement and/or between successive measurements. For example, consider the critical frequency of the F_2 layer.

A modern ionosonde can measure the critical frequency with high precision, e.g., 0.01 MHz, but a number with this precision is neither spatially nor temporally representative of the F_2 layer. Constantly changing ionospheric structure may cause a critical frequency at a location 10 km away to differ by about 1 MHz. There is little value in measuring a critical frequency to 0.01 MHz to obtain a monthly mean, or median in the case where the standard deviation of the data for that hour is on the order of 1 MHz. While high accuracy may be important in some special campaigns, secondary users are usually interested in representative values.

For most data-archiving purposes, representative data are more important than highly accurate data, especially when comparisons are to be made with similar data taken at different times and different locations. For example, a rocket or satellite sounder may make a highly accurate measurement at a given location and time, but that single measurement can hardly be considered representative. Some techniques involve measurements on only a few days or a single day per month (e.g., incoherent scatter). Generally, such measurements are good only for the time and location when/where they are made. Somewhat similarly, a satellite (topside) sounder does not return to the same geographical location and local time for several months. This makes it difficult to use their soundings as broadly representative data.

This problem is important in the intercomparison of data sets and in comparisons between data and the outputs of models. Usually, both numerical and physical models are driven by average inputs. Users of ionospheric models must, therefore, be aware that the models may not represent conditions at a particular place and time, and this fact should have played an important role in documenting the model's validation procedure.

3.2 Real-Time Data and Archived Data

In recent years there has been an increasing demand for "real-time" data, which are available shortly after the measurements are made. These data are valuable to users who must make decisions or corrections to systems as a result of ionospheric disturbances. Such users include radio operators and users of satellite systems affected by atmospheric drag. Real-time data sometimes are recorded and made available later through data centers; while some of these data are highly reliable, it is important that the user understand the difference between these real-time data and archive-quality data.

Because real-time data have to be available shortly after measurement, important interpretation and editing procedures may have been omitted. The case of total electron content measurements illustrates this point. Such measurements depend on phase and, hence, phase slippage or cycle ambiguity can lead to errors, especially at night. These cycle errors can only be corrected by the intervention of someone with specialized knowledge of prevailing ionospheric conditions and/or long-term behavior. Thus a user of real-time data should be aware of the uncertain quality of these data.

4. Examples of Digital Ionospheric Data Available from the World Data Center-A for STP

4.1 Solar-Terrestrial Indices

The types of data available from the WDC-A for STP and of interest to ionospheric users are given in Table 1. These include solar and geomagnetic data for correlative studies. The long-term variations of some of the more useful solar and geomagnetic indices, such as smoothed sunspot number and magnetic index, are shown in Figure 1.

For long-term ionospheric predictions, several indices are available that are based on ionospheric (F_2) data and expressed as the equivalent sunspot number that gives the observed average value of $f_o F_2$ for a number of ionospheric sounding locations. Such an index is the IG, an ionospherically based index for use with the existing data in CCIR Report 340-5 (CCIR 1986). Noontime monthly median values of observed $f_o F_2$ observed at 13 middle-latitude stations throughout the world are used to determine the value of the sunspot number

Table 1
Solar-Terrestrial Physics

Data Categories				
Solar	**Interplanetary**	**Magnetospheric**	**Ionospheric**	**Geomagnetic**
Sunspot Numbers	Solar Wind Plasma	Solar X-Rays	Ionosonde Operations	Activity Indices
Mass Ejections	Magnetic Fields	Energetic Particles	Ionograms • Digital • Analog	Magnetograms • Digital • Analog
Film Images	Cosmic Rays	Magnetic Variations	Scaled Parameters	Major Storm Listing
Radio Emissions	Solar Constant	Auroral Imagery	Fields and Currents	Models
Flares	Lyman-Alpha	Total Energy	Models	Maps
Publications: SGD, UAG	Sector Boundaries	Satellite Anomalies	INAG Bulletin	Indices Bulletin

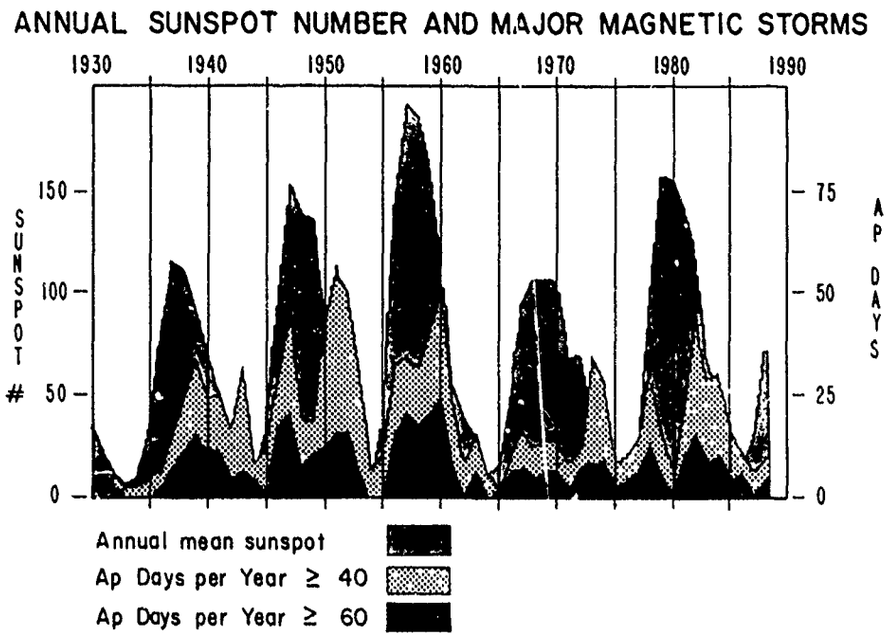


Figure 1. Long-term variations of solar and geomagnetic characteristics.

required in Report 340 to give the observed F_2 critical frequency. The value of the sunspot number thus determined is designated as the value of IG for the month in question. The IG is plotted in Figure 2 along with the monthly sunspot number. The overall agreement is good, except near the peak of the 1955–1964 cycle, when the F_2 layer critical frequency appears to "saturate." The 12-month smoothed value of IG is denoted by IG_{12}; it can be compared to the 12-month smoothed value of sunspot number R_{12}. Although the long-term averages of these indexes show good correlation, large differences can occur in individual indexes as seen in Figure 2.

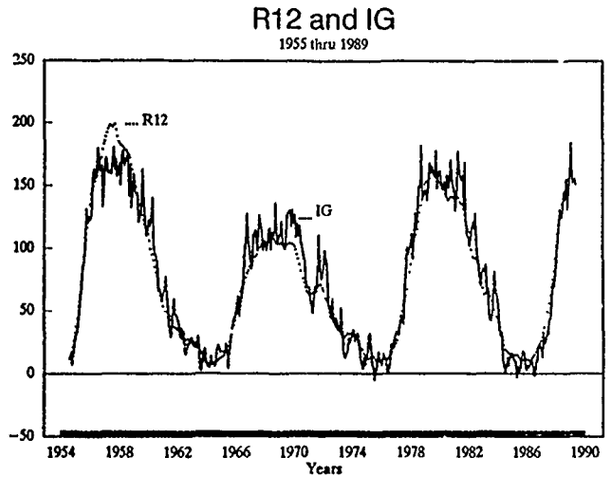


Figure 2. IG and R (smoothed).

Geomagnetic indices, available in publications, on tapes, on floppy disks, and CD-ROMs, are important resources that are provided by the Data Center. Figure 3 shows the long-term variation of the aa index, SSN, and number of flares. Note the difference in phase. The flare count does not seem to have a dependence on the magnitude of the sunspot cycle. Thus, there is more than one "11-year cycle."

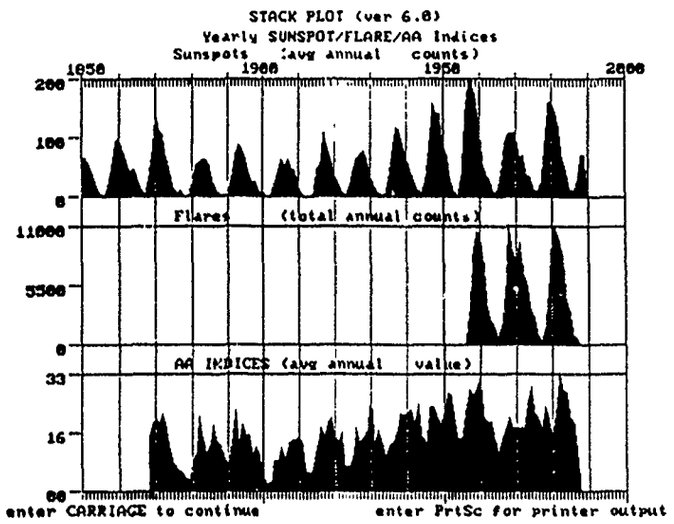


Figure 3. Long-term variations of geomagnetic data.

4.2 Maximum Electron Density

The maximum density of the F_2 layer, $N_m F_2$, is one of the most useful ionospheric parameters. It is derived from the critical frequency of the F_2 layer $f_o F_2$ by the relation: $N_m F_2 = 1.24 * 10^{10} * (f_o F_2)^2$ where $N_m F_2$, is in electrons per m^3 and $f_o F_2$ is in MHz. Hourly $N_m F_2$ data are available at some locations for over 50 years. One example is the monthly noon and midnight $N_m F_2$ near Washington, DC, as seen in Figures 4(a), 4(b), and 4(c). The following features are evident in these figures:

(1) There is a marked sunspot cycle variation of $N_m F_2$.

(2) The seasonal variation of $N_m F_2$ at sunspot minimum shows a peak in winter and a minimum in summer; in contrast, at sunspot maximum the seasonal peaks tend to occur around the equinoxes and the minima in summer. This is more clearly seen in Figures 4(b) and 4(c). Note that around sunspot maximum the summer noon value may be less than the summer midnight value.

(3) The seasonal variation of the midnight values show summer peaks and winter minima with high values near the peaks of the sunspot cycles and low values near the minima of the sunspot cycle.

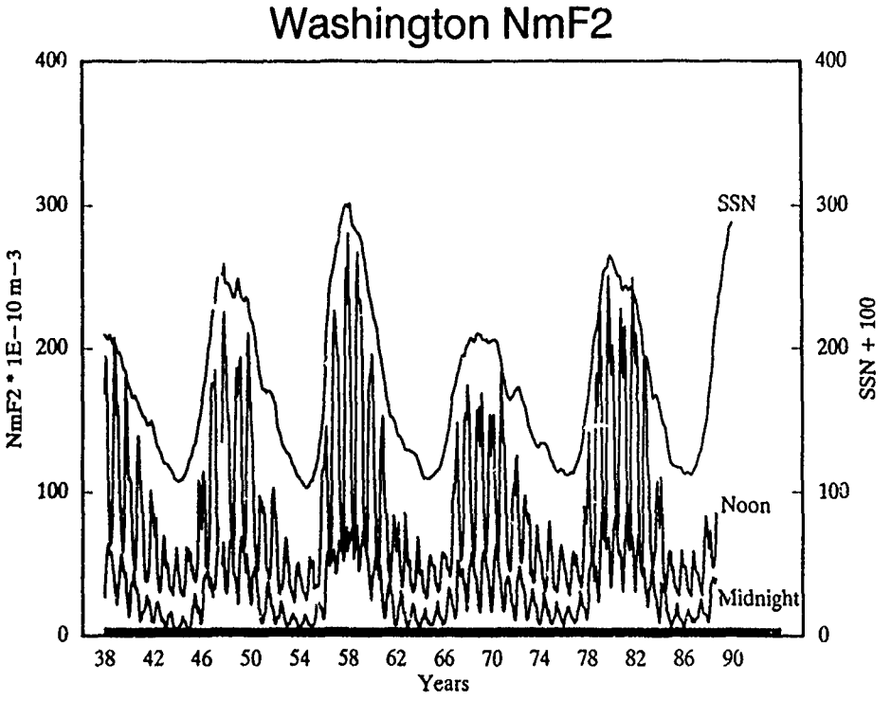


Figure 4(a). $N_m F_2$ near Washington, DC, 1938–1989.

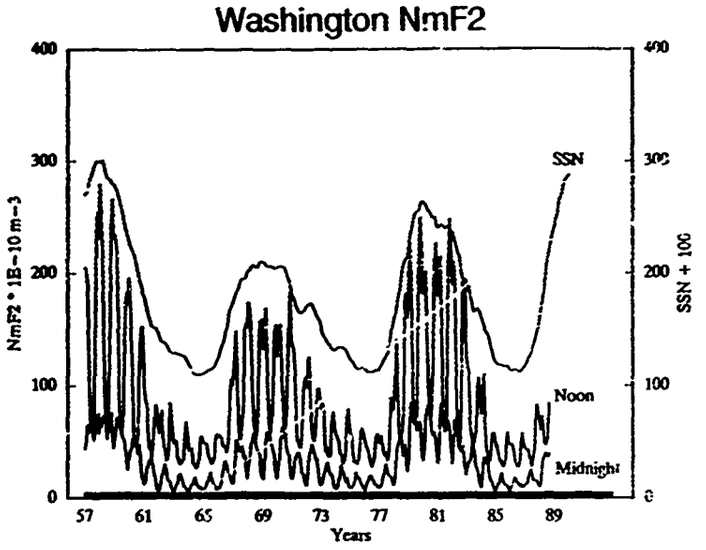


Figure 4(b). $N_m F_2$ **near Washington, DC, 1957–1987.**

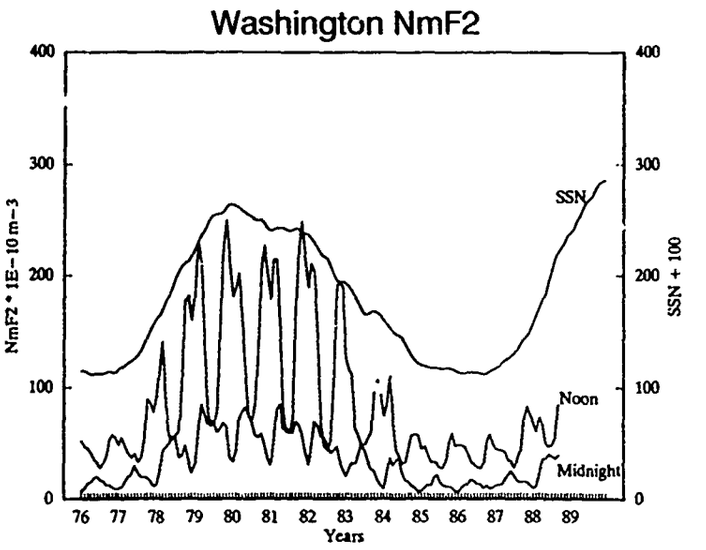


Figure 4(c). $N_m F_2$ **near Washington, DC, 1976–1987.**

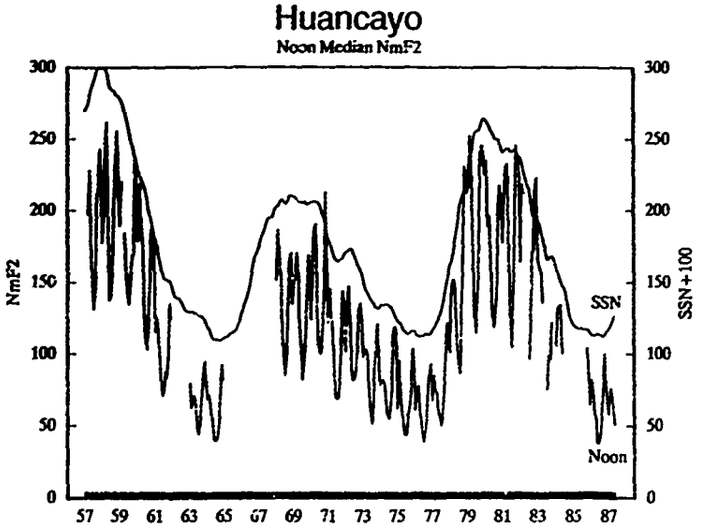


Figure 5(a). $N_m F_2$ at Huancayo, Peru, 1957–1987.

The dependence of $N_m F_2$ on sunspot number varies with geographical (and geomagnetic) location. Figure 5 shows the long-term variation of monthly median $N_m F_2$ at the magnetic dip equator. Gaps in figures are caused by missing data.

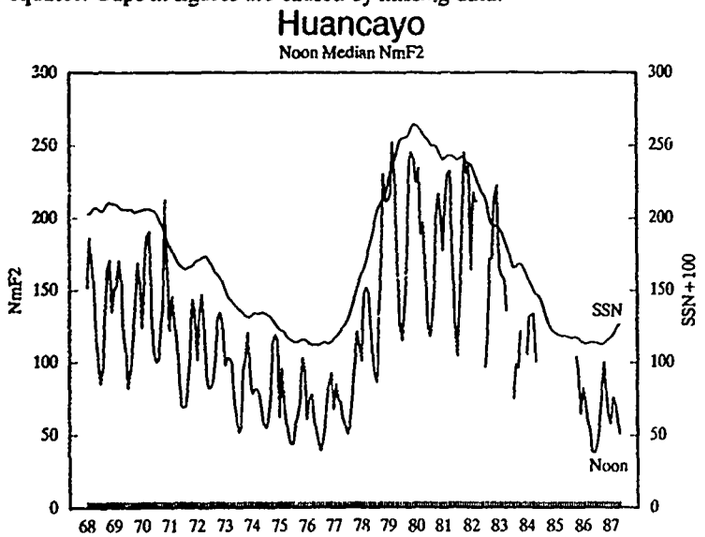


Figure 5(b). $N_m F_2$ at Huancayo, Peru, 1968–1987.

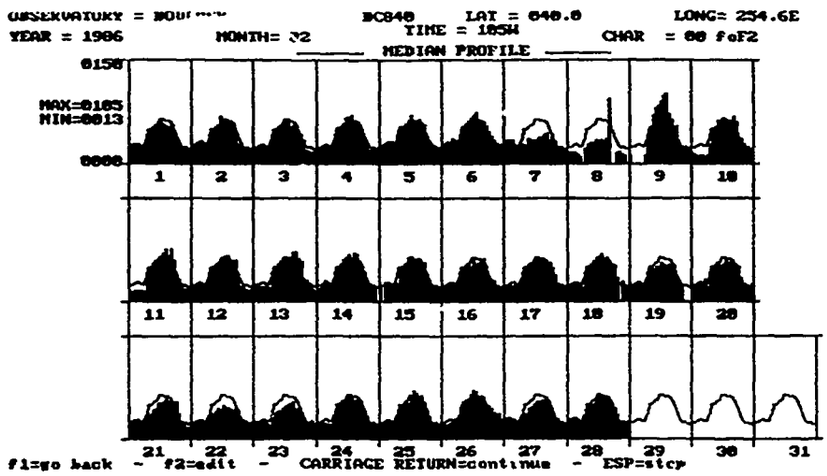


Figure 6. Hourly f_oF_2 plot, Boulder, CO, February 1986.

A particularly useful utility program is one that enables hourly data for a whole month to be displayed on screen. An output of this program (see Figure 6) shows Boulder hourly f_oF_2 values for February 1986. This display is useful, for example, in the identification of unusual events such as the storm on February 7, 8, and 9, and for the location and correction of possible erroneous values (e.g., data entry errors).

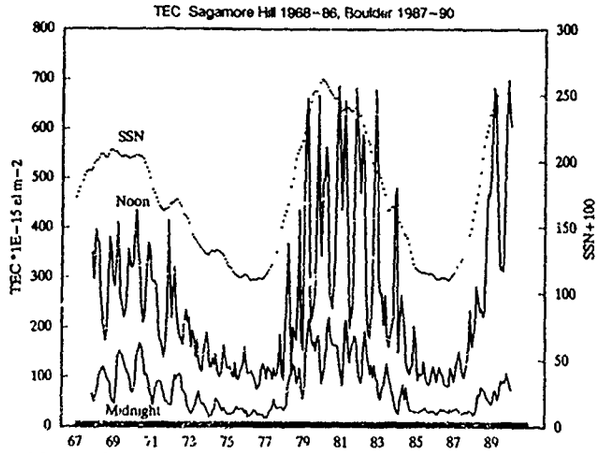


Figure 7. Monthly median electron contents at Sagamore Hill, MA (1967–1986) and at Boulder, CO (1987–1989).

4.3 Total Electron Content

Another important ionospheric characteristic is the total electron content in a 1-square-meter vertical column above a point (Davies, 1990, Chapter 8). Monthly median noon and midnight values of this quantity (obtained from Faraday rotation) are available from Sagamore Hill, Massachusetts, from November 1967 through September 1986; similar data are available from Boulder, Colorado, from January 1967 through December 1989. These data (Figure 7) again show the long-term solar control of the ionosphere. The sunspot cycle and seasonal variations are similar to those of $N_m F_2$; e.g., peaks occur in the noon values near equinoxes near sunspot maximum.

4.4 Sudden Ionospheric Disturbances

Most sudden ionospheric disturbances (SIDs) are caused by increases in D-region electron density due to enhancements in (0.1 nm to 1 nm) solar x-rays. Data on the occurrences of all types of SID (Davies, 1990, Chapter 9) are available from the monthly publication, *Solar-Geophysical Data* (SGD), by the National Geophysical Data Center. The monthly SID totals are shown in Figure 8. A clear solar-cycle dependence is evident; however, the relatively small number of occurrences in 1949–1964 largely reflects the relatively small number of observers. The correlation of SID numbers is somewhat better with the background solar x-ray flares (0.1 nm to 1 nm) than with sunspot number (Wagner, 1988). It must be remembered that the SIDs result from bursts of x-rays rather than from background emissions.

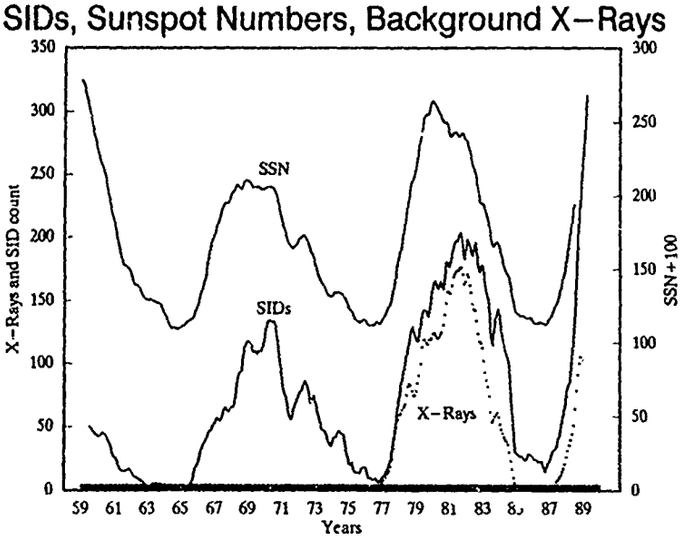


Figure 6. Smoothed monthly count of all types of SIDs occurring 1959–1989, smoothed sunspot numbers, and smoothed background solar x-ray fluxes.

Conclusions

We have sought in this paper to illustrate some of the changes that are taking place in providing solar-terrestrial data to customers of the World Data Center–A in Boulder, Colorado. Special attention has been paid to ionospheric data, to show new approaches to data processing as a result of advances in (small) computer technology, as well as improvements in data quality and ease of data retrieval. Emphasis is now on quality control.

Production of the NGDC Ionospheric CD–ROM (a data base of over 60 megabytes that can be manipulated by an individual researcher on a personal computer) would have been impossible without the recent shift of emphasis to quality control and the advances in small-computer technology.

Another important development in ionospheric data is the time extension of various data sets. Early ionospheric prediction systems were based on one or two solar cycles, and it was assumed that the behavior during this/these cycles would be typical. As the data bases have expanded, it has become clear that there are changes from cycle to cycle in the response of the ionosphere (and the geomagnetic field) to a given solar index.

With the advent of modern digital ionospheric sounders, new approaches and rules must evolve to deal with the large amount of data as well as the problems introduced by the replacement of skilled human scalers by computer (pattern recognition) techniques (Gilbert and Smith, 1988).

Errors in real-time data are of particular concern. Because of the short time scales required by the original data users, these data may not receive sufficient expert editing, and they may be inadequately calibrated for archival purposes.

References

CCIR, Propagation in Ionized Media: Recommendations and Reports of the CCIR, 1986, VI, Int. Telecom. Union, Geneva, 1986.

Davies, K., Ionospheric Radio, Peter Peregrinus, London, 1990.

Gilbert, J.D., and R.W. Smith, A comparison between automatic ionogram scaling system ARTIST and the standard manual method. *Radio Sci.*, **23**, 968, 1988.

Piggett, W.R., and K. Rawer, URSI Handbook of Ionogram Interpretation and Reduction, Report UAG–23A, World Data Center–A, U.S. Dept. of Commerce, Boulder, Colorado 80303, 1978.

Reinisch, B.W., New techniques in ground-based ionospheric sounding and studies, *Radio Sci.*, **21**, 331, 1986.

Reinisch, B.W., R.W. Gamache, and L.G. Bossy, Ionospheric characteristic for IRI in real time, *Adv. Space Res.*, **10(8)**, 25, 1990.

Wagner, W.J., Observations of 1–8 Å solar x-ray variability during solar cycle 21, *Adv. Space Res.*, **8(7)**, 67, 1988.

AD-P006 266

COMPUTERIZED IONOSPHERIC TOMOGRAPHY

Jeffrey R. Austen
Department of Electrical Engineering, Tennessee Technological University
Box 5004, Cookeville TN 38505

Timothy D. Raymund, C. H. Liu, and Steven J. Franke
Department of Electrical and Computer Engineering, University of Illinois at Urbana-Champaign
1406 W. Green St., Urbana IL 61801

John A. Klobuchar
Ionospheric Physics Division, Air Force Geophysics Laboratory
Hanscom AFB MA 01731-5000

J. Stalker
Physics Research Division, Emmanuel College
Boston MA 02115

91-09659

Abstract

In this paper the background of computerized tomography (CT) and its application to the ionosphere is reviewed. CT techniques, using only total electron content (TEC) data, can be used to reconstruct a two-dimensional image of the electron density in the ionosphere. The limitations of this technique are discussed and examples showing the limitations and capabilities are presented. Simulation results for two applications are presented: imaging the high latitude trough, and the correction of tracking radar range rate errors. Some possible extensions of the technique are presented.

Introduction

The mathematical foundation for Computerized Tomography (CT) was laid in 1917 by Radon [Radon, 1917] but was not utilized until 1970 [Gordon et al., 1970]. Since that time CT has revolutionized medical imaging and has been applied to several geophysical imaging problems; e.g., [Dynes and Lytle, 1979; Munk and Wunsch, 1979; Fleming, 1982]. The application to ionospheric imaging was made in 1988 [Austen et al., 1988]. In this paper the current status of Computerized Ionospheric Tomography (CIT) is presented. In the first section the theory of CT, as applied to ionospheric imaging, is reviewed. The next section shows the limitation of this technique when used with a "standard" geometry and gives example reconstructions for several cases. Next, an application of this technique to imaging the ionospheric trough, using a quite realistic model, is presented. The sensitivity of this technique to noise is investigated in the next section and then an application to the correction of range-rate data from a radar is given. Finally, the possibility of using data from other satellite systems in addition to the "standard" TEC data is explored.

Theory

An easily measured parameter of the ionosphere is the total electron content (TEC), which is the number of free electrons in a column of unit cross-sectional area

$$TEC = \int_p N(s)ds \qquad (1)$$

where N(s) is the electron content per unit volume, p is the propagation path between the transmitter and the receiver, and ds is the differential path length. The TEC values are not corrected for the angle between the path and zenith direction (the "Secant χ" correction) but are the "slant TEC" values. TEC measurements can be made in any direction; when vertical TEC values are mentioned they will be specifically labelled as such. Geometric optics is assumed for all cases, i.e. it is assumed that the carrier frequencies used to measure the TEC are high enough so that the path p is a straight line joining the transmitter and receiver.

TEC measurements supply the data input for the CT algorithms. The particular method used to determine TEC is not important but it should be noted that the true value of TEC is necessary, i.e. the so-called "constant of integration" must be determined by some other means if it is not directly measured.

Computerized tomography is the process of reconstructing an image from a series of its projections. CT principles and applications are presented in Herman [1980] and Kak and Slaney [1988]. The algorithms used for CT can be divided into two classes: transform methods and finite series expansion methods. They are reviewed by Lewitt [1983] and Censor [1983], respectively. Although the transform methods are not used for CIT, their analysis can provide some insight on the limitations of the CIT technique.

Examination of a simplified version of the direct Fourier inversion technique will provide some insight into the relationship between the sampling geometry and the properties of the reconstructed image. The sampling geometry is shown in Figure 1. The object density is represented by the function f(x,y) and each ray is described by coordinates s and θ (note that these are not polar coordinates) and has value p(s,θ). In the ionospheric imaging case this is the TEC value. Using the projection-slice theorem (see, e.g., Herman [1980])

$$\hat{p}(R,\theta) = \hat{f}(R\cos\theta\ ,\ R\sin\theta) \qquad (2)$$

where $\hat{p}$ is the 1-dimensional Fourier transform of p with respect to the first variable and $\hat{f}$ is the 2-dimensional Fourier transform of f. The meaning of this theorem can be illustrated by an example using the parallel-projection geometry. Let θ be fixed at some angle $\bar{\theta}$ and p(s,$\bar{\theta}$) be known for all s (see Figure 2a). The function p(s,$\bar{\theta}$) represents the line integrated density of the object over an infinite set of parallel lines through the object at angle $\bar{\theta}$. $\hat{p}(R,\bar{\theta})$ is the Fourier transform of p(s,$\bar{\theta}$) and by the projection slice theorem is equal to a line segment of the 2-dimensional Fourier transform of the object density at angle $\bar{\theta}$ through the origin (Figure 2b). If p(s,$\bar{\theta}$) is determined for all angles, $\bar{\theta}$, the function $\hat{f}$ is completely determined and, by 2-dimensional inverse Fourier transform, the function f can be computed. Using this theorem the relationship between the sampling and the output can be seen. Figure 3a shows an idealized case of CIT, assuming flat earth and uniformly spaced stations. As the satellite moves overhead the ray paths between the satellite and stations trace out a fan-like shape at each station. The minimum angle of the rays is determined by the minimum elevation angle at which the stations are capable of operating; there are no rays in the horizontal direction. This corresponds to limited-angle sampling of the image space and the corresponding known regions in 2-dimensional Fourier space are shown in Figure 3b. In this Figure it is clear that there are areas of Fourier space that are not sampled at all, specifically the k_y axis which corresponds to the vertical profile information in the CIT case. In order to recover the vertical profile it is necessary to sample along the k_y axis which requires horizontal rays; however, these rays are not available when using only ground-based TEC receivers (Figure 3a). When using a curved-earth geometry there is some information about the profile in the data, but the information is still insufficient to completely determine the profile without the use of *a priori* information.

The transform techniques are well known and there is much similarity between these techniques and those used in synthetic-aperture radar and very-long-baseline interferometry. As yet, these techniques do not seem to be suitable for application to the ionospheric tomography problem. This is because in CIT the geometry is irregular, there is a large amount of missing data, and there may be *a priori* information that can be used to aid the reconstruction process.

In the finite-series expansion methods the tomography problem is discretized at the beginning of the development and equations are written for the discrete system This process yields a large set of linear equations which can then be solved to determine the image. The solution of a large set of linear equations is a well-studied area of mathematics. The development of this method is presented by Austen et al [1988] but will be reviewed here.

Let f(r,ϕ), where r and ϕ are coordinates in a two-dimensional polar coordinate system (polar coordinates are used here for notational convenience), represent the electron density in the ionosphere and let p_i be the path of ray i. The measured parameter, y_i, is the TEC along path p_i and there are N_p paths,

$$y_i = \int_{p_i} f(r,\phi)ds + e_i^n \qquad i = 1,...,N_p \qquad (3)$$

where e_i^n represents the error due to noise and inaccuracies in measurements. The actual electron density, $f(r,\phi)$, can be approximated by using a set of N_b basis functions $\{b_j(r,\phi)\}_{j=1}^{N_b}$

$$f(r,\phi) = \hat{f}(r,\phi) + e^b(r,\phi) = \sum_{j=1}^{N_b} x_j b_j(r,\phi) + e^b(r,\phi) \tag{4}$$

where x_j is the weighting coefficient for $b_j(r,\phi)$ and $e^b(r,\phi)$ is the error due to the finite number of basis functions used to approximate the function f. Combining the above two equations and defining a two-dimensional matrix that depends only on the choice of paths and basis functions,

$$y_i = \sum_{j=1}^{N_b} A_{ij} x_j + e_i \tag{5a}$$

or, in matrix notation,

$$\mathbf{y} = \mathbf{A}\mathbf{x} + \mathbf{e} \tag{5b}$$

and

$$A_{ij} = \int_{p_i} b_j(r,\phi)\,ds \tag{6}$$

where e_i is the sum of the error terms, $e_i = e_i^b + e_i^n$, and e_i^b is the result obtained by integrating the error term $e^b(r,\phi)$ over the ith path. This matrix equation is the basis for the finite-series expansion technique. This equation is solved to determine **x** and then equation (4) is used to determine the function, f.

In many cases it is convenient to divide the region being imaged into small rectangular elements called pixels. Then a natural choice for the basis functions is equal to the set of pixels; with this choice the weighting factors x_j can be interpreted as the average density of the object in the jth pixel. For this case the definition for the set of basis functions is

$$b_j(r,\phi) = \begin{cases} 1, & \text{if } (r,\phi) \text{ is inside the jth pixel} \\ 0, & \text{otherwise} \end{cases} \tag{7}$$

There are many other possible basis function sets such as sine and cosine functions. In this paper it is assumed that approximately rectangular pixels are used as the set of basis functions (Figure 4). When using this set of basis functions, it should be noted that the values of **A** are all non-negative and the value of A_{ij} represents the length of the ith ray through the jth pixel. In the ionospheric imaging case the values of **x** (electron density) are non-negative and **y**, being the integral of electron density, are also non-negative. Once the basis functions are chosen and the ray paths are known, (6) can be evaluated to obtain **A**. The problem is then reduced to solving (5) where **x** is the unknown and **y** is the known data, subject to some constraint on **x** and the error term **e**.

Although (5) is a standard set of linear equations there are two characteristics which should be noted. In the ionospheric case the number of pixels and the number of rays can each easily exceed one-thousand. This corresponds to an **A** matrix of over one-million elements. Even though the matrix is large, it is relatively sparse, only a few percent of the entries are non-zero and this percentage decreases as the size of the matrix increases. The non-zero entries, though, are not arranged in a pattern that can be made use of [Censor, 1981].

There are many different techniques for solving (5). They can be broken into two general classes: non-iterative and iterative. Non-iterative techniques employ a formula to compute **x** directly from **y** and **A**; for example, if **A** is square and non-singular then $\mathbf{x} = \mathbf{A}^{-1}\mathbf{y}$. Iterative solutions begin with an initial guess for the solution, $\mathbf{x}^{(0)}$, and then iteratively refine the solution where, for the kth iteration, $\mathbf{x}^{(k+1)} = f(\mathbf{x}^{(k)}, \mathbf{y}, \mathbf{A})$ where f is some function. At some point the iterations are stopped and the current result, $\mathbf{x}^{(k+1)}$, is declared the solution. In many cases the set of equations is over-determined and inconsistent and therefore an exact solution does not exist.

Non-iterative solution techniques directly compute a solution to the equation. It would be impractical, even if a unique solution was known to exist, to attempt to solve the equations by some common method such as Gaussian elimination, but there are a few direct solution techniques that are useful. Two possibilities are discrete backprojection and singular-value-decomposition. A useful variation on discrete backprojection is presented in Raymund et al., [1990].

Iterative solution techniques have generated much interest due to their simplicity and computational efficiency. Although some of the algorithms are heuristic, there are many which have mathematical proofs of their convergence properties. Herman and Lent [1976] and Gordon [1974] provide a good review of iterative algorithms. Several algorithms have been investigated for CIT: algebraic reconstruction technique (ART), simultaneous ART (SIRT), and multiplicative ART (MART). The authors have found that MART performs best.

The MART algorithm examines each ray, in turn, and determines the error ratio, the true value of y_i divided by the value of y_i computed from $\mathbf{x}^{(k)}$; it then updates the x values in a multiplicative fashion to reduce the error. This error is then distributed over all the x values so that the pixels in which the ray has a longer path have a larger correction factor. Because of the multiplicative correction, instead of additive, in the "standard" geometry case (e.g. Figure 4) where there is no information about the profile contained in the data, the profile of the initial guess is preserved through the iterations. This algorithm was first mentioned by Gordon, et al. [1970], and it was proven to converge to the maximum entropy solution by Lent [1976]. The initial guess is $\mathbf{x}^{(0)} = \mathbf{1}$. For the kth iteration:

$$x_j^{(k+1)} = x_j^{(k)} \left(\frac{y_i}{\sum_{j=1}^{N_b} A_{ij} x_j^{(k)}} \right)^{\lambda_k A_{ij}} , \quad j=1,\ldots,N_b \tag{8}$$

where $i = (k \bmod N_p)+1$ and $0 < \lambda_k \leq 1$. The equations must be scaled so that $A_{ij} \leq 1$ for all i,j. This algorithm is computationally much more time-consuming than ART because of the exponentiation. When applied to CIT, this algorithm is used with an initial guess containing some information about the structure of this ionosphere rather than the one given above.

There are several geometrical considerations for the reconstruction problem. Obviously all pixels must have at least one ray through them and all rays must go through at least one pixel. For the ionospheric case there are many transmitter positions and few receiver positions. For the tomography technique to work well it is necessary to maximize the information content of equation (5). To do this the set of rays associated with each receiver should cover as many pixels as possible. This results in a spatial sampling requirement: the spacing of the rays must be small enough so that adjacent rays to the same receiver go through the same or adjacent pixels. For the type of geometry shown in Figure 4 the most stringent sampling requirement occurs for the highest altitude pixels directly above a receiver. At this location the horizontal distance between rays must be equal to or less than the horizontal dimension of the pixels. The sampling period at the receiver is

$$\text{sampling period} < \frac{\text{width of pixel}}{\text{velocity of satellite}} \frac{\text{height of satellite}}{\text{height of highest pixel}} \tag{9}$$

Another assumption made is that the region being observed does not change significantly during the time of observation. A pure motion, without evolution of the features, can be compensated for if the direction and magnitude are known, but an evolution of the features during the time of observation will result in a degradation of the reconstructed image.

Reconstruction of Profiles

In this section the CIT technique is applied to the reconstruction of electron density profiles without irregularities. The geometry is based on a hypothetical network of TEC receivers in Europe. It consists of 13 receivers on a north-south line at 10° east longitude with a spacing of 2.5° in latitude (approximately 275 km) between receivers. The most southern receiver is located at latitude 40° N and the most northern receiver is at 70° N. A network similar to this, but using fewer receivers, has been implemented and is described by Leitinger et al. [1984]. The satellite is assumed to be in a circular polar orbit at an altitude of 1000 km; at this altitude the satellite velocity is 7.35 km/s. The ionosphere to be imaged ranges from 39.9° N to 70.1° N latitude (approximately 3500 km), just beyond the farthest receivers, and in height from 100 km to 900 km above the surface of the Earth. The top and bottom heights of the imaged region are chosen so that it contains all of the region where there is significant electron density. This region is divided into a total of 1820 pixels, 20 pixels in the vertical direction and 91 pixels in the horizontal direction; they are approximately square with dimensions 40 km by 40 km. Because a spherical geometry is used the pixels are actually sections of annuli instead of rectangles and the horizontal dimensions vary slightly with altitude. The TEC data is sampled at 1.5 times the minimum required sampling rate at zenith, resulting in a horizontal sampling distance of 28.6 km at the satellite altitude, or a sampling period of approximately 3.9 seconds. This geometry is shown in Figure 4.

Realistic quantities of noise have been added to the simulated TEC data used for this example. The receivers are assumed to provide absolute TEC measurements (i.e. the "constant of integration" is determined) to an accuracy of 10^{15} el./m^2 RMS and relative measurements (i.e. sample-to-sample noise) to an accuracy of 10^{14} el./m^2 RMS. For all stations exact time

synchronization is assumed and a minimum elevation angle of 15° is assumed to be required for reception of the satellite beacon.

Three models will be presented in this section, all of them variants of a Chapman profile. The background profile is modeled as a single Chapman layer where N_{max} is the maximum density, z_{max} is the height of maximum density, and H_{scale} is the scale height. The parameters z_{max}, H_{scale}, and N_{max} are, in general, functions of latitude and longitude. The first model is a simple profile with parameters $N_{max} = 10^{12}$ el./m^3, $H_{scale} = 75$ km, and $z_{max} = 350$ km. Figure 5 shows the model and two reconstructions, each with a different initial guess. Note that the reconstruction is equal to the initial guess in this case – the profile cannot be determined from the data alone. The second model is a tilted version of the first. The model and a reconstruction is shown in Figure 6. Note that again the profile cannot be determined. It should be emphasized that this is due to incomplete data, specifically the lack of data corresponding to rays which are parallel and nearly-parallel to the surface of the earth; it is not a fundamental limitation of the CT technique. By obtaining additional data from more than one orbiting satellite or incorporating data from other sources, such as a sounder, this weakness can be overcome.

The next model has a wave-like perturbation to the Chapman profile. The perturbation is a sinusoidal variation in peak height with a wavelength of 6° latitude (approximately 710 km) and a peak variation of ±10% of the nominal value. Figure 7 shows the model and the reconstruction for this case; the initial guess for this reconstruction is the unperturbed Chapman profile. When a reasonable initial guess is given, variations in the profile are reconstructed, even though the profile itself cannot be reconstructed. CIT is also able to correctly reconstruct perturbations in the amplitude and scale height of a Chapman layer profile (not shown). Figure 7 also shows that the vertical TEC computed from the reconstructed image is much better than that obtained by combining the secant χ corrected TEC data from each individual receiver.

Trough Reconstruction

A reconstruction of a realistic high latitude trough model was carried out using simulated TEC data supplied by Klobuchar and Stalker [Raymund et al., 1990]. The data were supplied without the original electron density model being known to the other authors (JRA, TDR, CHL, and SJF), who used them to reconstruct the image of the electron density. After producing the reconstruction the original model electron density was revealed and compared with the results from CIT. The model covered the height range from 100 to 1000 km altitude above the earth's surface and a latitude range from 40° to 90°N in a north-south plane. Five TEC receivers were used and an orbiting satellite at an altitude of 1000 km was the signal source. The five stations were placed at the same longitude and at 57.0°N, 61.0°N, 65.0°N, 69.0°N, and 73.0°N latitude. Deliberate numerical errors in calculating the simulated data provided noise levels approximately equal to those expected in real TEC data. Figure 8 shows the TEC received at the five stations; the roughness in the curves is due to the noise. These five TEC curves from the five receiving stations are plotted as functions of the satellite positions and they are equivalent to the time series of TEC records received at the corresponding stations. These curves contain all the information of the ionosphere to be reconstructed and were the only data used for the reconstruction.

MART, using a backprojection-Chapman initial guess [Raymund et al., 1990], was used for the reconstruction. For comparison purposes, the reconstructed image is plotted between 59°N and 71°N, though the reconstruction image ranged between 57°N and 73°N and from 50 to 950 km altitude. The image was divided into 50 vertical pixels by 100 horizontal pixels, yielding an image resolution of roughly 20 km by 20 km. Figure 9 shows the geometry, including the grid of the image (small squares) and the ray paths between the satellite and the stations (fans).

The parameters of the Chapman layer function were found by using an iterative technique described in Raymund et al. [1990]. This technique involves producing trial reconstructions and varying the parameters of the Chapman profile while monitoring the rates of convergence. Using this method H_{max} was estimated to be 300.0 km and H_{scale} was found to be 65.0 km. The backprojection-Chapman initial guess was then constructed using these parameters.

This initial guess is used as the starting point for the iterative application of the MART algorithm. Figure 10 shows the image after 10 iterations with a relaxation of 0.01. The small black squares at the bottom denote receiver positions. This image can be compared with the actual data. Figure 11 shows a contour plot of the model from which the TEC data was generated; this model is an idealization of an actual ionospheric mapping made by the Chatanika incoherent scatter radar from 0543 to 0557 UT. The radar mapping of the high-latitude ionosphere is shown in Figure 12 (after Figure 2c in Weber, et al., [1985]). This mapping covers the range from approximately 80 to 700 km, and 59.7° to 70.5°N latitude, or 59.5° to 71.0°N corrected geomagnetic latitude. The latitude range shown is roughly that shown in Figs. 10 and 11. In Figure 12, note the boundary blob north of the narrow trough and the slight horizontal gradient, which are simulated well in the model. Comparing Figs. 10 and 11, we note that the reconstructed image shows good agreement in electron density levels when compared to the contour of the model, though the peak levels are not quite as high in the reconstruction. This can be attributed to the current limit on the number of pixels. An increased number of pixels lowering the size of individual pixels, would make the features easier to resolve. The trough has been very accurately located in the reconstruction. Some degree of the horizontal gradients in the original model was also reconstructed. The diagonal ripples in the reconstruction are artifacts of the reconstruction.

Effects of Noise

To demonstrate the performance of this technique in the presence of noise several reconstructions have been performed. For these reconstructions the geometry is the same as shown in Figure 4. The model consists of a tilted Chapman layer containing Gaussian depletions. The background has parameters $N_{max} = 10^{12}$ el./m^3, $H_{scale} = 75$ km, and $z_{max} = 350$ km at 60° N latitude with a slope in z_{max} of +5 km per degree north. This background is multiplied by Gaussian depletions with a scale size of 100 km and a maximum depletion of 90% (i.e., the minimum value is 10% of the background). These depletions are all located at altitude 350 km and are spaced horizontally so that there is one every 5° latitude (approximately 590 km). The model is shown in Figure 13.

For the reconstructions the initial guess is set equal to the actual background (tilted Chapman layer) and the MART algorithm is used. Figure 14 shows the reconstruction with no noise added to the data. There are two types of noise: measurement noise which is uncorrelated from sample to sample, and "constant of integration" noise which is perfectly correlated for all rays at the same receiver; both forms of noise are modelled as Gaussian. In Figure 15 the normally expected amount of noise has been added to the data. The measurement noise has an RMS value of 10^{14} el./m^3 and the "constant of integration" noise has an RMS value of 10^{15} el./m^3. Figure 16 shows the result with ten times the normal noise added to the data: 10^{15} el./m^3 measurement noise and 10^{16} el./m^3 "constant of integration" noise.

From this example it is seen that CIT suffers very little degradation in the presence of the expected amount of noise and the technique still performs well, although some degradation is seen, when the noise is increased by a factor of ten.

Application to Doppler Tracking Radar

A potential application of ionospheric CT is to estimate the corrections to the range rate of a Doppler radar used for tracking targets through the ionosphere. In tracking orbital targets, the Doppler frequency shift of the radar is used to calculate the radial velocity component of the target. The true range rate based on both the measured Doppler shift and the TEC is [Raymund et al., 1990]

$$\frac{dR}{dt} \cong \frac{-\Delta f\, c}{2f} + \frac{40.31}{f^2} \frac{d}{dt} \int_R N \, dR. \qquad (10)$$

Using the reconstructed image of the ionosphere, the second term in (10) can be calculated and potentially used to correct measurements of the radial velocity.

To demonstrate this application, a simulated ionosphere model shown in Figure 17 is used. The reconstructed image using six receivers is shown in Figure 18. Consider a radar tracking an orbiting body through this ionosphere as shown in Figure 19. To compute the range rate correction for a radar operating at 150 Mhz through this ionosphere, the second term in (10) was calculated using the original model ionosphere (Figure 17), and is shown in Figure 20. This is the true effect of the intervening ionosphere on the range rate (relative to free space propagation). To compare how well the reconstructed ionosphere can predict the range rate correction, a similar computation of the second term in (10) was made based on the reconstructed image (Figure 18) and is shown as the second, more jagged curve in Figure 20. These computations were made using numerical integration along the ray path, and then central differencing for d/dt. Note that except for the extreme values, the curves match fairly well, indicating that reconstructed images of the ionosphere can be used to make corrections to the Doppler radar range rate measurements.

Potential Extension of the Technique

Using several orbiting satellites in concert it may be possible to collect data which would augment the data collected by ground stations. If the necessary data can be obtained the electron density, including the profile, can be completely determined. One possibility is the use of the proposed Earth Observing System (EOS) platform in conjunction with the Global Positioning System (GPS) satellite constellation.

The EOS platform will be a nearly polar orbiting platform at roughly 778 km altitude and will have an orbital period of 98 minutes. GPS satellites are located at roughly 20,000 km altitude and appear to move slowly when viewed from the earth. The orbital elements have been slightly simplified to make the orbits ideal. Other assumptions are that the earth is perfectly spherical, the fore EOS antenna will be able to acquire a rising GPS satellite instantly, and ray-bending is negligible.

For the map Figure (Figure 21), the two circular maps at the top are the North pole region (on the left) and the South pole region (on the right). A Mercator projection of the earth is on the bottom. The polar plots cover from the pole to 30 degrees of latitude away from the pole. The EOS-GPS links and the EOS orbit are projected on the earth; only the part of the link that is below the altitude of the EOS is shown. The EOS sample time (that is the time between consecutive links to GPS) is about 56 seconds. There are a total of 678 links. The sub-orbital path cuts north-south above the Atlantic ocean and north-south above South America and Australia. The orbit also passes very nearly over both poles. Note that it appears that some links cross

directly over the satellite orbit (for example over the North pole); these actually are two links, one going to GPS off to one side of EOS, the other going to GPS on the other side of EOS. These links provide the broad global coverage that would be of particular interest for 3-D tomography applications. Also note that some of the links travel almost along the EOS orbit path (for example from Australia southward). These north-south links are also of particular interest for the 2-D tomography application.

Summary

The above examples show that the CIT technique has a promising future. Many simulations have been performed and the results demonstrate that CIT can be used to generate a two-dimensional image of the electron density in the ionosphere. However, simulations are never perfect (or perhaps too perfect!), thus the next step in the development of this technique is to test it with real data. In order for the test to be valid it must be conducted such that there is another source of data with which the CIT results can be compared. Establishing a line of TEC receivers near an incoherent-scatter radar site would be one method of obtaining the necessary data.

It has been shown, both theoretically and through simulations, that with only the data available from ground-based TEC stations the electron density profile cannot be determined. Possible methods to overcome this weakness are to use *a priori* data, such as that from a ground-based ionospheric sounder, or to establish a TEC station in orbit, making use of the EOS or GPS platforms.

Further work is also needed in the areas of basis functions, initial guesses, incorporation of motion into the algorithm, and in the number and placement of TEC stations for optimum results.

Acknowledgements

JRA, TDR, CHL, and SJF were supported by NSF grant ATM88-02860.

References

Austen, J. R., S. J. Franke, and C. H. Liu, Ionospheric Imaging Using Computerized Tomography, *Radio Sci., 23,* (3), 299-307, 1988.

Censor, Y., Row–action methods for huge and sparse systems and their applications, *SIAM Rev., 23,* (4), 444–466, 1981.

Censor, Y., Finite series-expansion reconstruction methods, *Proc. IEEE, 71,* (3), 409–419, 1983.

Dines, K. A., and R. J. Lytle, Computerized geophysical tomography, *Proc. IEEE, 67,* (7), 1065–1073, 1979.

Fleming, H. E., Satellite remote sensing by the technique of computed tomography, *J. Applied Meteorology, 21,* 1538-1549, 1982.

Gordon, R., A tutorial on ART, *IEEE Trans. on Nucl. Sci., NS-21,* (3), 78-93, 1974.

Gordon, R., R. Bender, and G. T. Herman, Algebraic Reconstruction Techniques (ART) for three-dimensional electron microscopy and X-ray photography, *J. Theor. Biol., 29,* (3), 471–481, 1970.

Herman, G. T., *Image Reconstruction From Projections,* Academic, Orlando, Fla., 1980.

Herman, G. T., and A. Lent, Iterative reconstruction algorithms, *Comput. Biol. Med., 6,* 273-294, 1976.

Herman, G. T., A. Lent, and S. W. Rowland, ART: Mathematics and Applications, *J. Theor. Biol., 42,* (1), 1-32, 1973.

Kak, A. C., and M. Slaney, *Principles of Computerized Tomographic Imaging,* IEEE Press, New York, NY, 1988.

Leitinger, R., G. K. Hartmann, F.-J. Lohmar, and E. Putz, Electron content measurements with geodetic Doppler receivers, *Radio Sci., 19,* (3), 789–797, 1984.

Lent, A., A convergent algorithm for maximum entropy image restoration, with a medical X-ray application, SPSE International Conference on Image Analysis and Evaluation, 1976, Toronto, Canada, Proceedings, 249–257.

Lewitt, R. M., Reconstruction algorithms: Transform methods, *Proc. IEEE, 71,* (3), 390–408, 1983.

Munk, W., and C. Wunsch, Ocean acoustic tomography: A scheme for large scale monitoring, *Deep Sea Res., 26A,* 123–161, 1979.

Radon, J., Über die Bestimmung von Funktionen durch ihre Integralwerte längs gewisser Mannigfaltigkeiten, *Berichete Saechsische Akademie der Wissenschaften, 69,* 262-277, 1917.

Raymund, T. D., J. R. Austen, S. J. Franke, C. H. Liu, J. A. Klobuchar, and J. Stalker, Application of Computerized Tomography to the Investigation of Ionospheric Structures, to be published in *Radio Sci.*

Weber, E. J., R. T. Tsunoda, J. Buchau, R. E. Sheenan, D. J. Strickland, W. Whiting, and J. G. Moore, Coordinated measurements of auroral zone plasma enhancements, *JGR, 90,* (A7), 6497-6513, 1985.

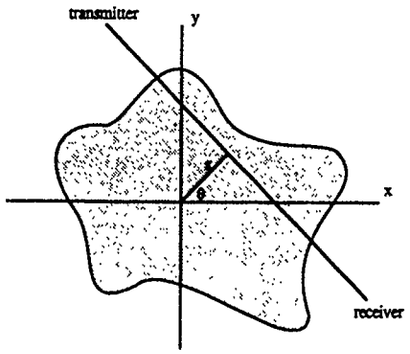


Figure 1. Geometry for CT sampling. Note that the pair (s,θ) is not polar coordinates.

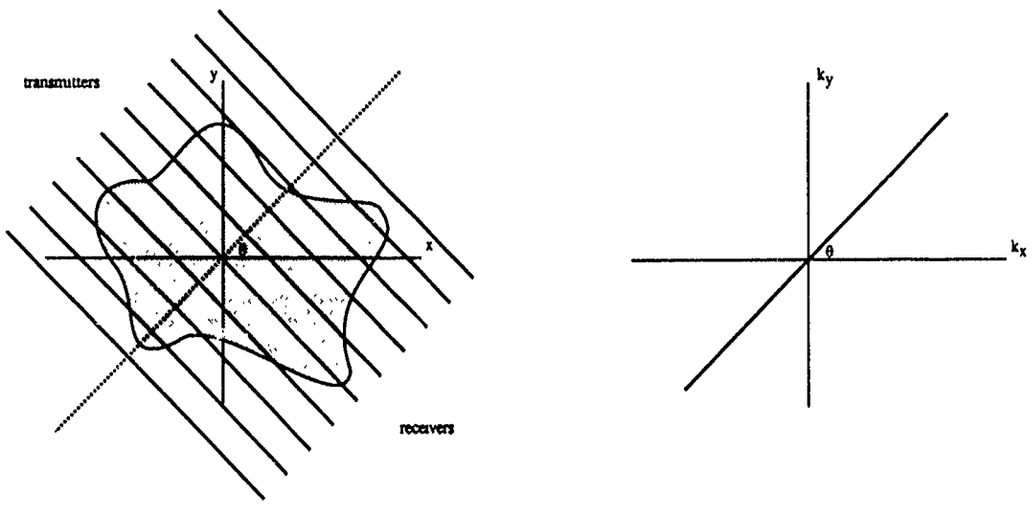


Figure 2. (a) An example of parallel-projection sampling at angle θ; the object is sampled for all values of s. (b) The Fourier space sampled in this case.

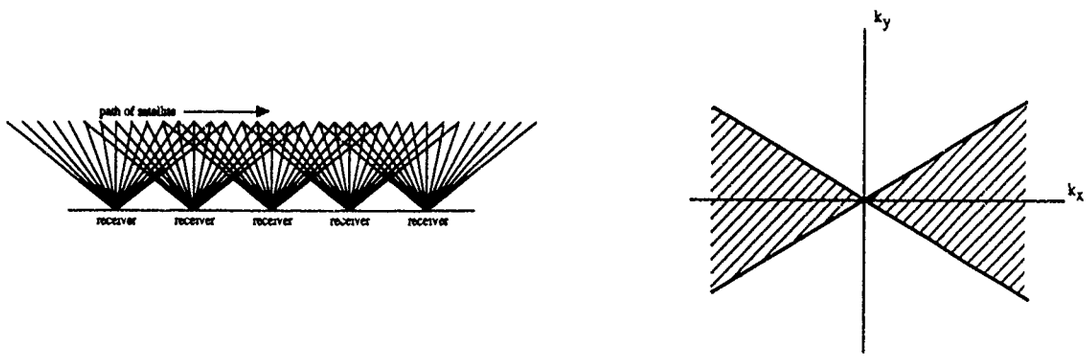


Figure 3. (a) Idealized CIT geometry assuming flat earth and uniformly spaced stations.
(b) The Fourier space sampled in this case.

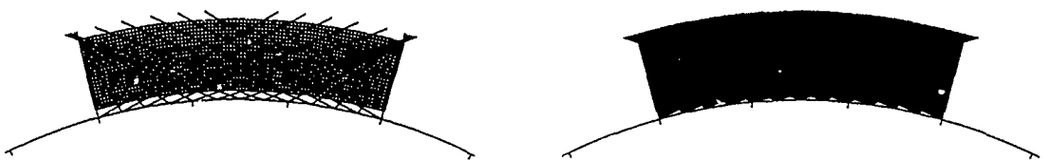


Figure 4. (a) Actual geometry for CIT simulation using 13 receivers. Only the lowest elevation rays for each receiver are shown. (b) The same geometry with all the rays shown; there are 1511 rays.

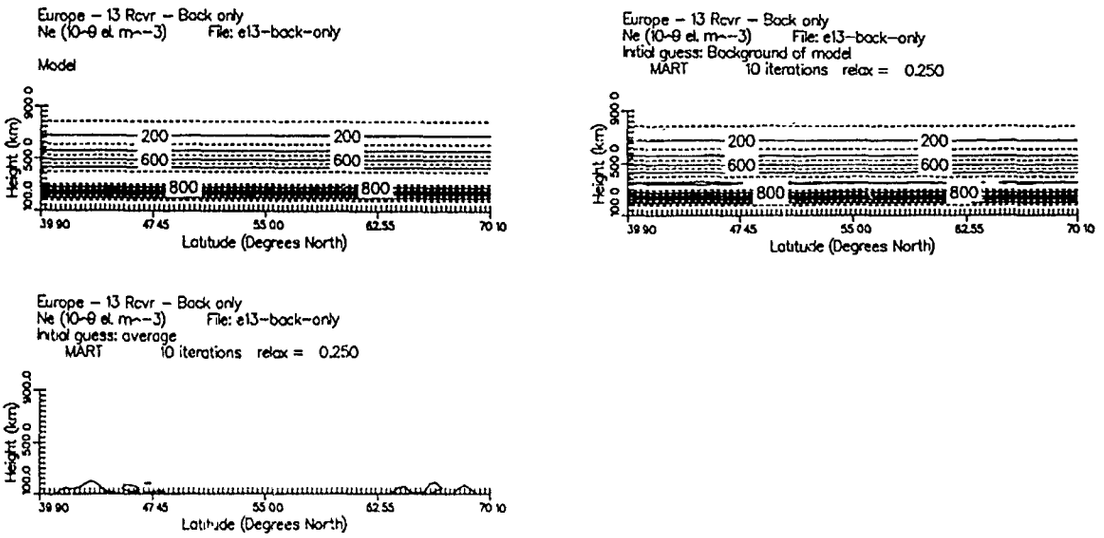


Figure 5. (a) Chapman profile model. (b) Reconstruction with perfect initial guess. (c) Reconstruction with initial guess of uniform density.

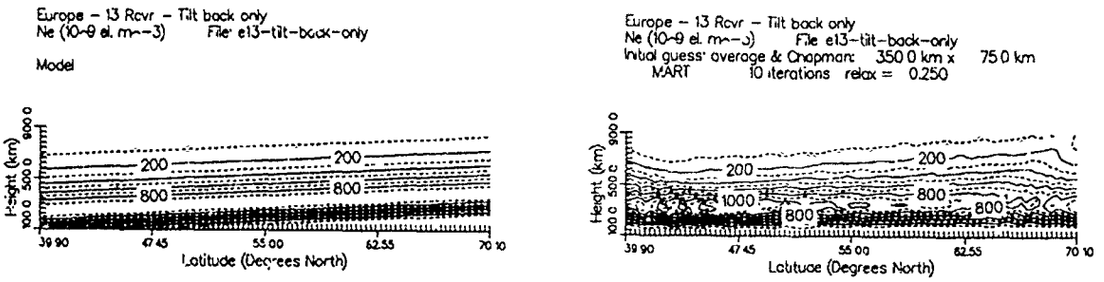


Figure 6. (a) Tilted Chapman profile model. (b) Reconstruction with initial guess equal to a non-tilted Chapman layer.

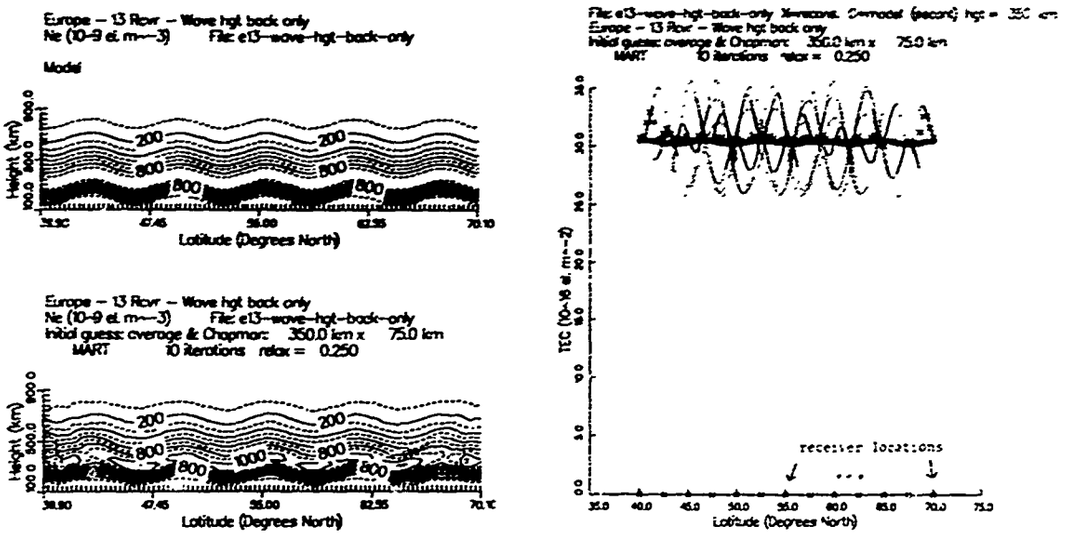


Figure 7. (a) Chapman profile model with wave perturbations. (b) Reconstruction with initial guess equal to a non-perturbed Chapman layer. (c) TEC received at each station, "O" denotes true TEC, "X" denotes reconstructed TEC, solid lines denote TEC computed nearly overhead at each station, and dotted lines denote TEC for all angles at each station.

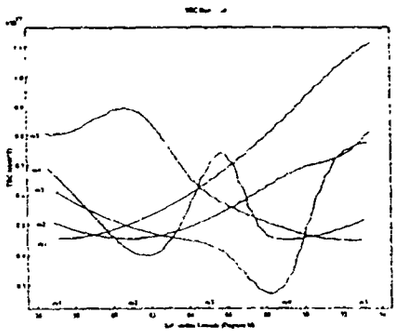


Figure 8. TEC data used as input to the reconstruction algorithm.

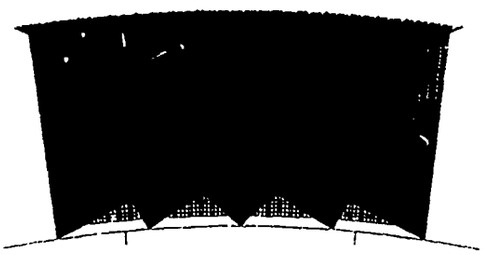


Figure 9. Geometry for the trough reconstruction.

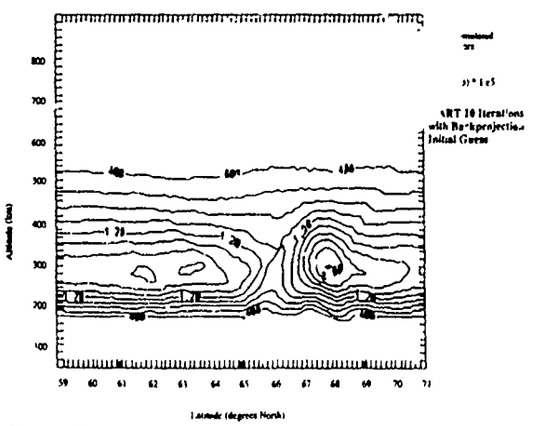


Figure 10. Reconstructed image of the trough.

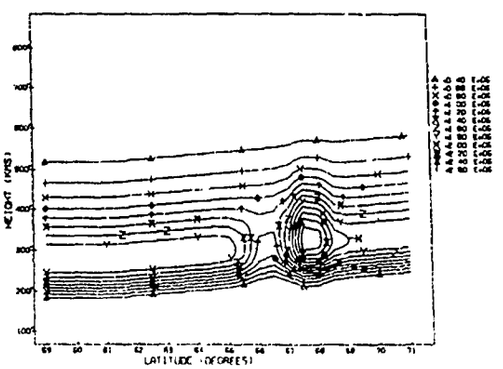


Figure 11. Model image of the trough; the TEC data shown in Figure 8 was generated from this model.

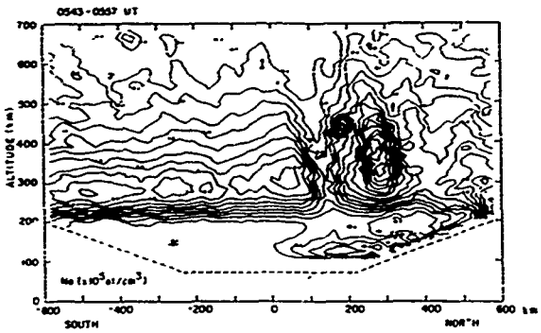


Figure 12. Actual radar map of the trough; the model (Figure 11) was based on this data (from Weber, et al. [1985]).

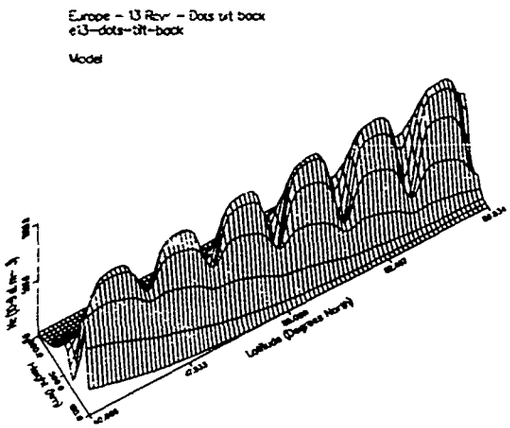


Figure 13. Model image showing a tilted Chapman profile with Gaussian depletions.

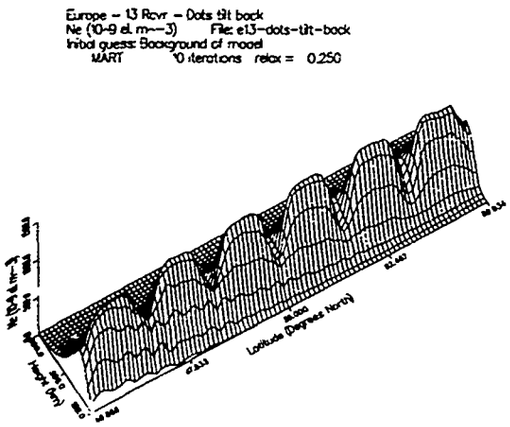


Figure 14. Reconstruction with no noise.

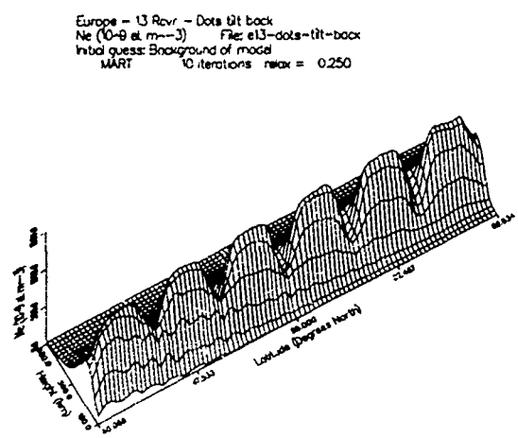


Figure 15. Reconstruction with normal amount of noise.

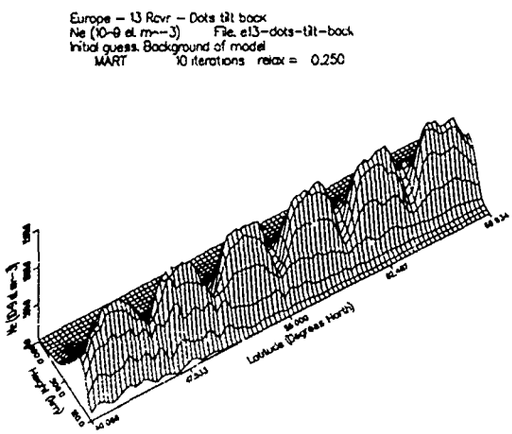


Figure 16. Reconstruction with ten times the normal amount of noise.

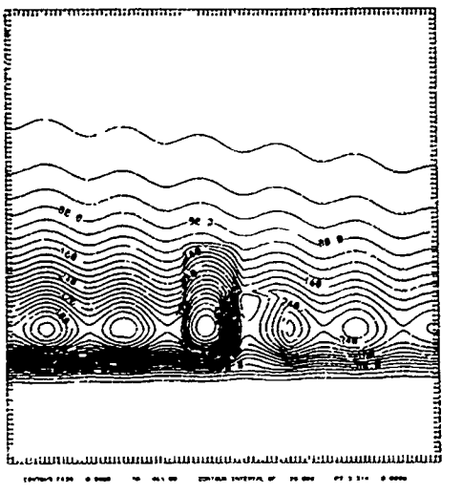


Figure 17. Model ionosphere showing irregularity.

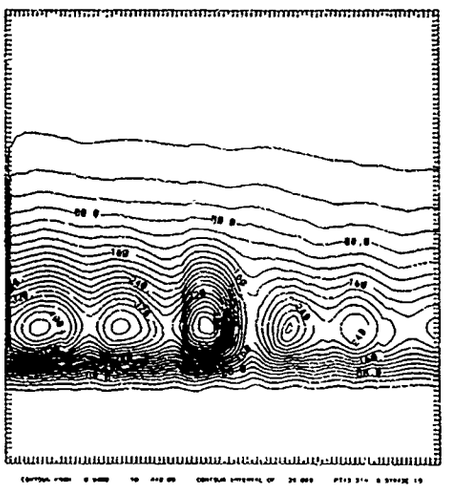


Figure 18. Reconstructed ionosphere.

Figure 19. Geometry used for radar example; the tops of the rays denote the path of the target and the point where the rays converge denote the location of the radar.

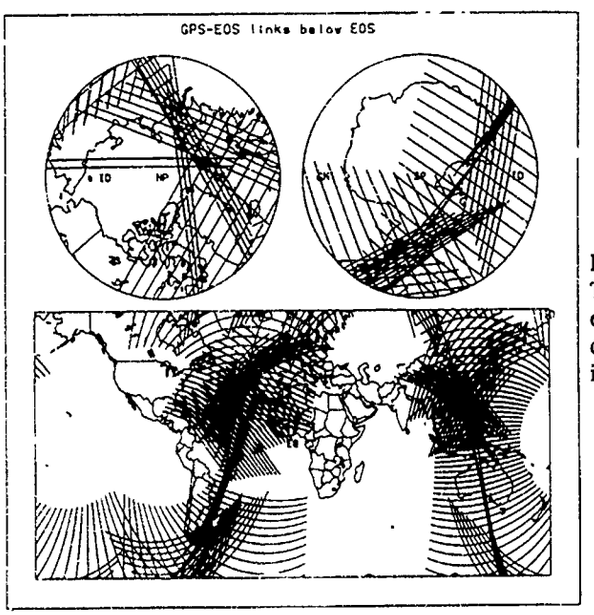


Figure 21. EOS/GPS geometry. The lines show the ionospheric intersection points of the rays with the ionosphere. Top left is view of North polar region; top right is view of South polar region.

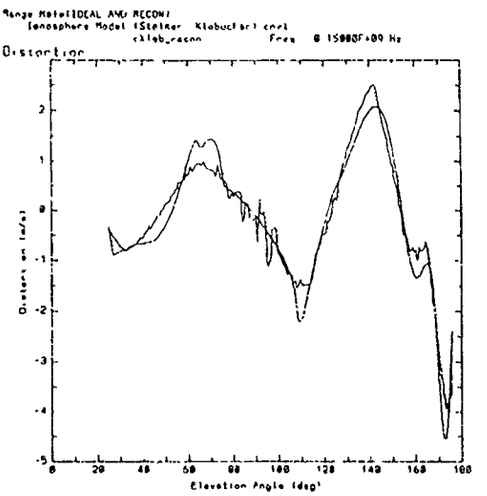


Figure 20. Range rate errors due to ionospheric effects. The smooth curve is the true error, calculated from the electron density model; the rough curve is the error calculated from the CIT reconstructed electron density image.

AD-P006 267

IONOSPHERIC MODIFICATION BY CHEMICAL RELEASES AND HIGH POWER RADIO WAVES

P. A. BERNHARDT, W. A. SCALES, M. J. KESKINEN, H.L. ROWLAND
SPACE PLASMA BRANCH, PLASMA PHYSICS DIVISION
NAVAL RESEARCH LABORATORY
WASHINGTON, DC 20375-5000

L. M. DUNCAN
DEPARTMENT OF PHYSICS AND ASTRONOMY
CLEMSON UNIVERSITY
CLEMSON, SC 29634

ABSTRACT

Ionospheric plasma density irregularities can be produced locally by chemical releases from space vehicles or remotely by a beam of high power radio waves transmitted from the ground. F-region plasma modification occurs by (1) chemically enhancing the electron number density, (2) chemically reducing the electron population, or (3) physically convecting the plasma from one region to another. The three processes (production, loss, and transport) can determine the effectiveness of ionospheric chemical releases and high frequency electromagnetic wave transmissions in subtle and surprising ways. Initially, a chemical release produces a localized change in plasma density and high power radio waves heat the electrons to yield enhanced transport from pressure gradients in the heated region. Subsequent processes, however, can lead to enhanced transport in chemically modified regions and modified reaction rates in the regions affected by high power radio waves. Both methods of ionospheric modification excite artificial enhancements in airglow intensities either by exothermic chemical reactions or by energetic electron fluxes which collisionally interact with neutral species. Numerical models have been developed to describe the creation and evolution of large scale (>1 km) density irregularities and airglow clouds generated by manmade sources. Experimental data compares favorably with these models. In general, we find that chemical releases produce transient, large amplitude perturbations in electron density while continuous wave heating by powerful electromagnetic waves yields weaker, longer-lived irregularities.

INTRODUCTION

The ionosphere is the primary medium for radio wave propagation. Natural irregularities can degrade a communication channel by inducing amplitude or phase fluctuations as a result of multipath mixing or scattering. Artificial modification of the F-region can be used to generate artificial density structures, control existing irregularities or to study processes which govern their evolution.

The two main classes of ionospheric modification are (1) the release of netural chemically reactive substances and (2) the transmission of high power radio waves into the F-region. Any type of ionospheric modification can be understood in terms of the electron and ion continuity equations

$$\frac{\partial n_j}{\partial t} + \nabla\cdot(n_j \underline{v}_j) = P_j - L_j \qquad (1)$$

where the subscript "j" refers to electrons to ions, n_j is the concentration, $\underline{v}_j$ is the velocity, P_j is the production rate and L_j is the loss rate for the plasma species. Chemical releases can directly affect the right side of the continuity equation by enhancing the production by photo or collisional ionization of neutral vapors. Or chemical releases can accelerate losses through electron attachment or dissociative recombination processes. Powerful radio waves can affect the electron production by accelerating electrons to high enough energies to cause neutral breakdown but this does not occur in the ionosphere for facilities currently in existence. Temperature dependent electron-ion recombination rates are modified in the artificially heated environment.

91-09660

High power radio waves most directly couple to the electron densities in the ionosphere by modifying the transport term $[\nabla\cdot(n_e\underline{v}_e)]$ in the continuity equation for electrons. The velocity of each plasma species is determined by the equation of motion:

$$\frac{\partial \underline{v}_j}{\partial t} + (\underline{v}_j\cdot\nabla)\underline{v}_j = -\frac{\nabla P_j}{\rho_j} + \frac{q_j}{m_j}(\underline{E}+\underline{v}_j\times\underline{B}) + \underline{g} - \nu_{jn}(\underline{v}_j-\underline{U}) - \nu_{ji}(\underline{v}_j-\underline{v}_i) \qquad (2)$$

where $P_j = n_jkT_j$ is pressure, $\rho_j = n_jm_j$ is mass density, T_j is electron or ion temperature, m_j is mass, q_j is charge, $\underline{E}$ is electric field, $\underline{B}$ is ambient magnetic flux density, $\underline{g}$ is gravitional acceleration, ν_{jn} is the electron- (or ion-) neutral collision frequency, $\underline{U}$ is neutral velocity, ν_{ji} is the electron- (or ion-) ion collision frequency, and $\underline{v}_i$ is the velocity of the other plasma species. Most of the terms on the right side of (2) can be artificially changed. The flow of plasma out of a region of high power radio waves is a result of enhanced plasma pressure in the heated region. Suprathermal electron fluxes are produced by large amplitude electric fields near the regions where the electromagnetic waves resonate with the electron plasma. For chemical releases, enhanced neutral velocities and polarization electric fields at the edge of expanding clouds accelerate the ambient electrons.

This paper is an overview of the relative effects of production, loss and transport on large scale ionospheric modification by chemical releases and high power radio waves. The next three section discuss chemical mechanisms, parallel and perpendicular transport, and the effects of injection velocities on the interactions. The following section describe ionospheric modification by high power radio waves with emphasis of thermal self focussing in a convecting plasma. The results from computer simulations are compared with experimental observations when possible. The last section summarizes the relative effectiveness of the two methods for production of ionospheric irregularities.

CHEMICAL RELEASES

The chemistry of ionospheric depletion chemicals has been discussed in detail by Bernhardt [1987]. Here a simple comparison of the chemical processes is made to illustrate the relative strengths and weaknesses of the different classes of chemical modification in the ionosphere.

There are four classes of initial reactions in the upper atmosphere following a neutral gas release. The first class is oxidation. Atomic oxygen is the most abundant species between about 250 and 600 km altitude. It can react rapidly with the released material to yield substances which may be less effective as ionospheric modifiers. This reaction has the form

$$XY + O \rightarrow XO^* + Y \qquad (3)$$

where XY represents the injected substance and XO^* is the product of the reaction which may be left in an excited state. Airglow emissions from the excited states may interfere with other light emissions which are related to the ion-electron chemistry. From the point of view of ionospheric modification, the Class I neutral reactions are unwanted because they reduce the efficiency for ionospheric modification and they produce airglow emissions which contaminate emissions used to diagnose plasma chemistry.

The second class of reaction in the upper atmosphere is photoionization. This reaction is generically represented as

$$XY + h\nu \rightarrow XY^+ + e^- \qquad (4)$$

When easily ionizable materials such as barium, cesium, sodium, europium, etc. are released in sunlight, photoionization yields enhanced ion and electron concentrations. Optical diagnosis of these releases is possible if the parent neutral and or the product ion fluoresce in sunlight. The reaction for barium is

$$Ba + h\nu \rightarrow Ba^+ + e^- \qquad \nu_i = 0.036\ s^{-1} \qquad (5)$$

The primary disadvantage for the Class II reactions is that sunlight is required to produce the ionospheric modification.

The third class of ionospheric modification reaction is the positive-ion, molecule charge transfer. This reaction takes the form

$$XY + O^+ \rightarrow XO^+ + X \qquad (6)$$

where XO^+ is a positive molecular ion which can rapidly recombine with electrons. The dissociative recombination reaction is

$$XO^+ + e^- \rightarrow X^* + O^* \qquad (7)$$

where X^* and O^* are excited states. The sequence of reactions from a CO_2 release is

$$CO_2 + O^+ \rightarrow O_2^+ + CO \qquad k = 10^{-9}\ cm^{-3} \qquad (8)$$

$$O_2^+ + e^- \rightarrow O(^1D) + O(^3P) \qquad k = 10^{-7}\ cm^{-3} \qquad (9)$$

$$O(^1D) \rightarrow O(^3P) + h\nu(630\ nm) \qquad \nu_R = 6.7 \times 10^{-3}\ s^{-1} \qquad (10)$$

where $O^1D)$ is the lowest excited state of atomic oxygen which yields red-line emissions upon transition to the ground $O(^3P)$ state. Class III reactions eventually lead to electron density reductions after the molecular product ions have recombined with electrons. The primary limitation of this class of reaction is that at two step process (reaction (6) followed by (7)) are required to produce electron density reductions.

The fourth class of reaction directly depletes the electron concentration by dissociative electron attachment. The form of this reaction is

$$XY + e^- \rightarrow X^- + Y \qquad (11)$$

where X^- is the negative ion dissociation product. The reaction with tri-fluoromethylbromide is

$$CF_3Br + e^- \rightarrow Br^- + CF_3 \qquad k = 10^{-7}\ cm^{-3} \qquad (12)$$

Comparing reactions (8) and (12), we see that the initial electron attachment reactions can be one hundred times more rapid than the initial ion-molecule reactions. Complications arise from the temperature dependence of the electron attachment reactions. Below the activation energy of reaction, the reaction rates decrease rapidly with temperature. The activation energy for (12) is 0.086 eV. The negative ion created by the Class IV reactions either reacts with the atomic oxygen ion

$$X^- + O^+ \rightarrow X^* + O^* \qquad (13)$$

or becomes photo dissociated

$$X^- + h\nu \rightarrow X + e^- \qquad (14)$$

The photodetachment reaction from solar illumination for the bromine anion from (12) is

$$Br^- + h\nu \rightarrow Br + e^- \qquad \nu_D = 0.27\ s^{-1} \qquad (15)$$

The primary limitation for ionospheric modification from electron attachment reactions is that they must be conducted in darkness for long lived effects. Enhanced airglow from the excited state in (13) can provide a diagnostic of the chemical and physical processes in the modified plasma.

PARALLEL TRANSPORT

The motion of the plasma following a chemical release can be considered in terms of parallel and perpendicular transport relative to $\underline{B}$. The two directions tend to be uncoupled owing to the relatively large parallel conductivity which makes the magnitude of the ambipolar electric fields oriented parappel to $\underline{B}$ much less than the magnitude of the polarization fields across the magnetic field lines.

The details of the numerical model of parallel transport is described by Bernhardt [1990]. The distribution of ionized species along magnetic field lines is computed from the equations of continuity (1) and momentum (2). Only variation along the magnetic field direction is considered. The inertial terms on the right side of (2) are dropped. The spatial distributions of the ambient species (e^- and O^+) and the chemically produced species (XY^+ or X^-) are computed self consistently with chemical reactions and transport. The ambipolar electric field ($\underline{E}$) is calculated assuming constant current density along $\underline{B}$.

Simulations have been carried out for the ionospheric modification by Class II, III, and IV chemicals deposited at 300 km altitude into an F-region plasma with an undisturbed density of 3 x 10^5 cm^{-3}. To illustrate a Class II reaction, the release of 10^{26} atoms (23 kg) of barium vapor was simulated (Figure 1). The barium ion concentration reaches a maximum of 5.8 x 10^6 cm^{-3} at the center of the cloud. The diffusion along the magnetic field line causes the Ba^+ ions to extend 70 kilometers from the release point 9 minutes after the release. The O^+ ions are excluded from the center of the barium cloud and are compressed at the sides. The electron concentration is the sum of the positive ion concentrations.

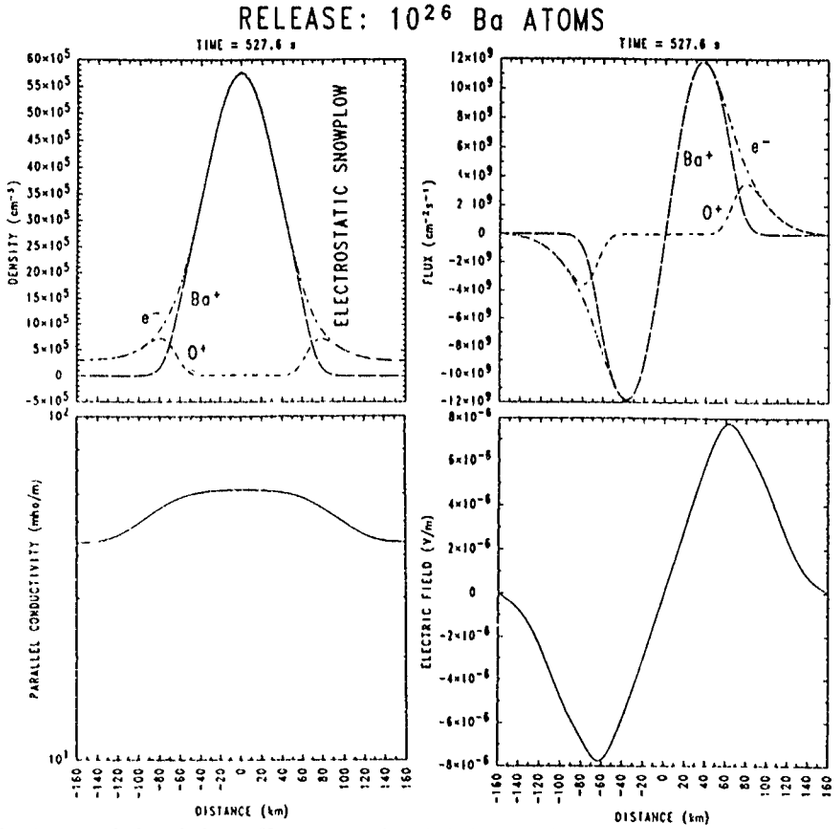


Figure 1. Computed densities, fluxes, conductivities, and ambipolar electric fields for a solar-illuminated barium release in the ionosphere.

The O^+ depression at the center of the barium cloud is a result of the ambipolar electric fields at the edges of the cloud. These fields reach a maximum of 8 μV/m and are oriented to keep the mobile electrons from escaping from the region of high density barium. These same fields will exclude the positively charged O^+ ions from entering the barium region. This process is called the electrostatic snowplow [Schunk and Szuszczewicz, 1988]. The fluxes of all species are directed outward from the center of the cloud because the O^+ ions are being carried along with the expanding barium plasma. Figure 1 also shows a 40% increase in the parallel conductivity at the center of the cloud which is much smaller than the 20-fold increase in electron density. Electron-ion collisions limit the magnitude of this conductivity.

The electrostatic snow-plow of ambient NO^+ ions during the Places barium release is shown in Figure 2 [Narcisi and Szuszczewicz, 1983]. The enhancement of NO^+ of about 50% at 149 km altitude (designated by the arrow in Figure 2) can be attributed to ambipolar electric fields at the edge of the barium cloud.

The Class III reactions initially reduce the O^+ density. The ionospheric modification 12.8 seconds after the release of 10^{26} CO_2 molecules at 300 km altitude is illustrated in Figure 3. At the center of the cloud, most of the O^+ ions have be converted into O_2^+ ions. Recombination of the diatomic ion with electrons produces a 15% reduction in the electron density. The hole is being replenished by field aligned fluxes of O^+ and e^- toward the center of the depletion. The molecular ion concentration is reduced by both chemistry and by outward fluxes. The parallel conductivity perturbation from this release is negligible and the ambipolar electric fields are relatively weak (less than 4 μV/m).

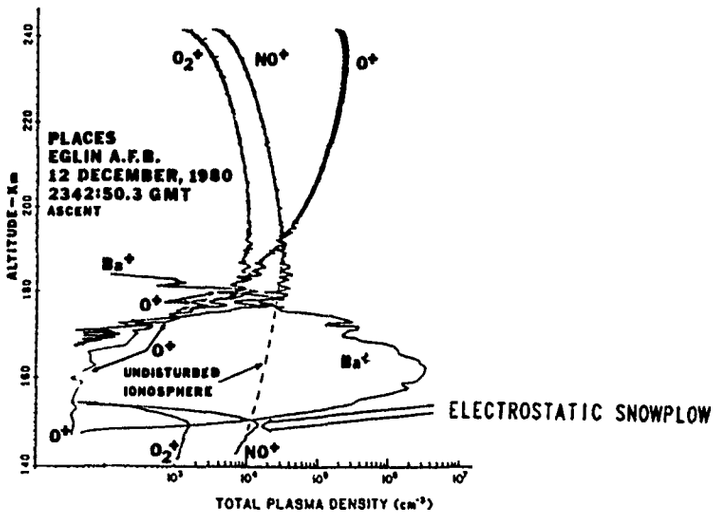


Figure 2. Ion mass spectrometer measurement of NO^+ density showing the effects of ambipolar electric fields.

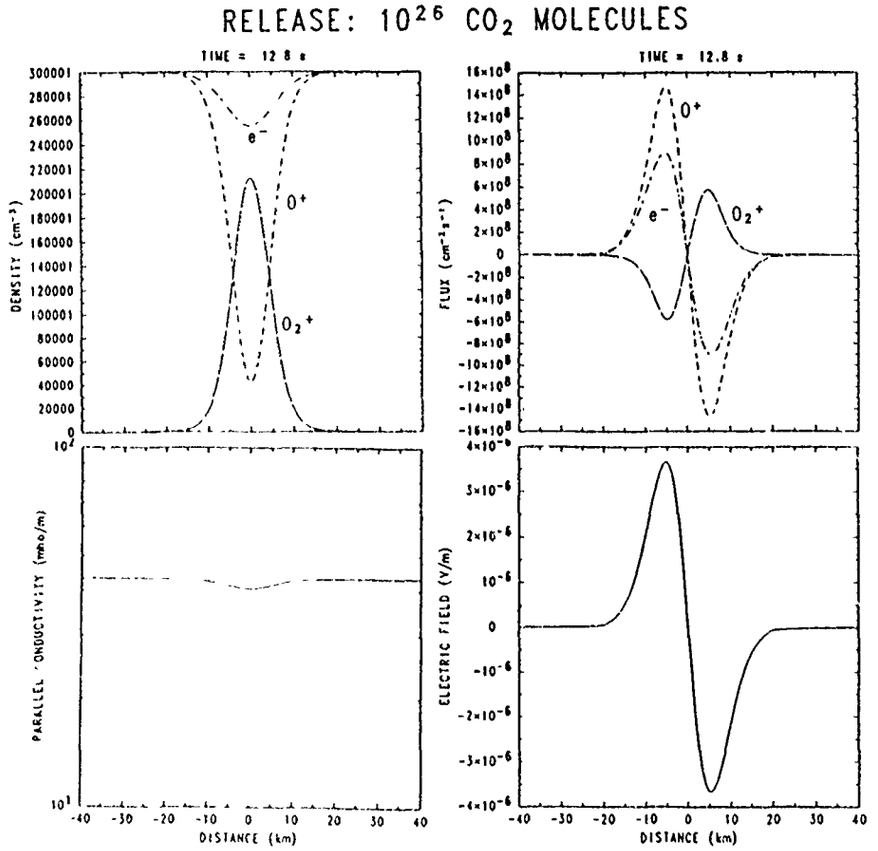


Figure 3. Computed densities, fluxes, conductivities and ambipolar electric fields for a CO_2 release into a uniform ionosphere.

During the NASA sponsored Spacelab II experiment, the Space Shuttle Orbital Maneuvering Subsystem (OMS) engines were used to deposit 280 kg of CO_2, H_2O, H_2, CO, and N_2 molecules into the F-layer over Arecibo, Puerto Rico. Figure 4 shows the evolution of the localized depletion near the peak of the layer as measured with the Arecibo incoherent scatter radar. The observed density depression is consistent with the numerical models involving Class III reactions and ambipolar diffusion. The apparent sinking of the hole into the bottomside ionosphere is attributed to cross-field plasma drift of a depleted flux tube across the radar line-of-site [Bernhardt et al., 1988 a,b].

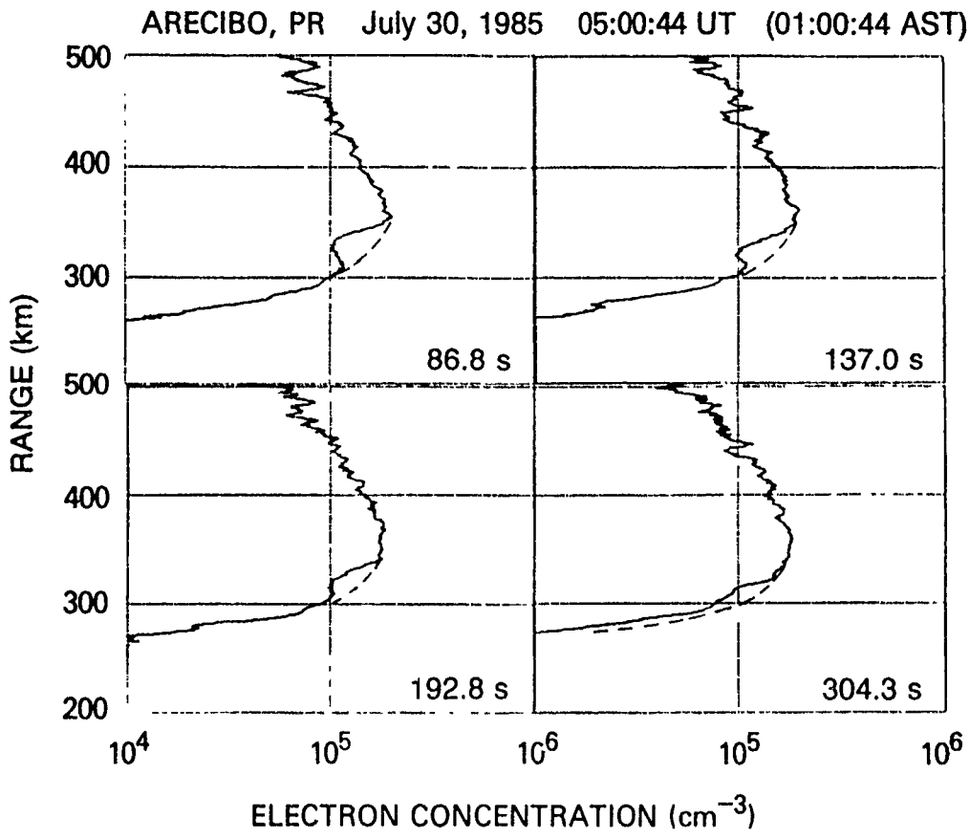


Figure 4. Ionospheric profiles showing the evolution of the hole produced by the Space Shuttle OMS engine burn over Arecibo, Puerto Rico.

The class IV reactions produce much larger perturbations in electron density, plasma fluxes, parallel conductivity, and electric fields (Figure 5). Nine seconds after the release of 10^{26} molecules of CF_3Br, nearly all electrons within a seven kilometer radius are attached to form Br^- ions. The ambipolar electric field strength is 130 μV/m with a polarity that inhibits electrons from flowing into the center. These fields act to compress the ambient O^+ ions toward the center as shown in Figure 5. The steep gradients at the edge of the modified region produce electron fluxes that are a factor ten greater than in the previous examples.

Near the release point, the electron concentration is reduced to the point that ions are only carriers of electric current. This limits the parallel conductivity reduction to two orders of magnitude. The class IV reactions are preferred for experiments which need the maximum changes in longitudinal conductivity.

As illustrated by reaction (13), mutual neutralization of the negative and positive ions leaves the products in excited states. This can produce airglow the can be recorded from the ground. Figure 6 is an amage of the airlgow cloud produced by the SF_6 release which occurred during the NASA/Boston University SPINEX-2 experiment [Fulford et al., 1987]. The 777.4 nm emissions are a result of the mutual neutralization of SF_6^- and O^+ ions yielding excited $O(^5P)$ states [Bernhardt et al., 1986]. The airglow cloud fills the region initially depleted by the electron attachment reactions.

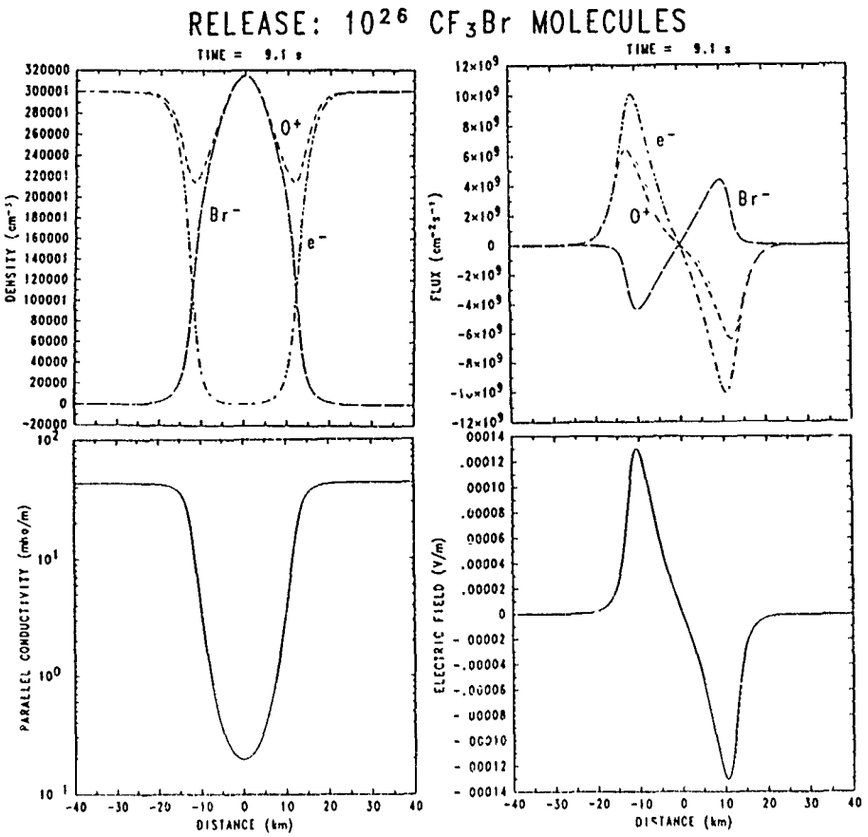


Figure 5. Simulation of the release of an electron attachment material in a uniform plasma.

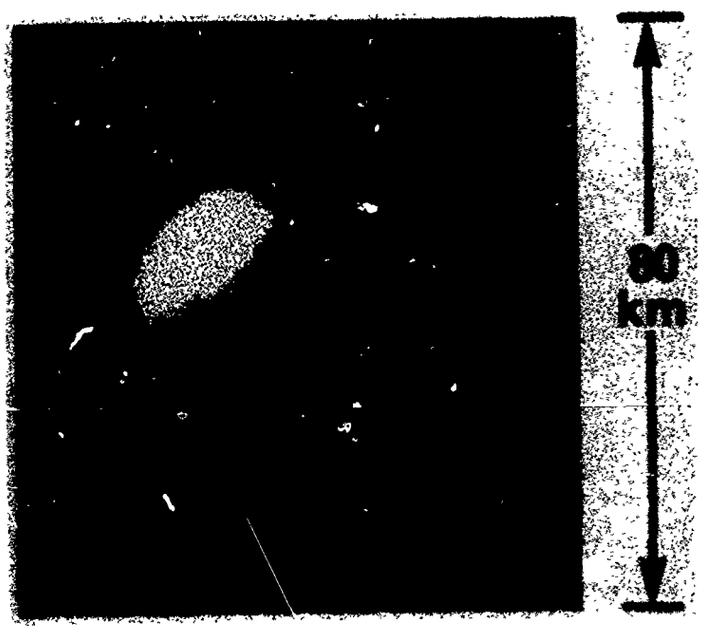


Figure 6. 777.4 nm airglow emissions for a nighttime, 40 kg release of SF_6 at 252 km altitude during the 28 April 1986 SPINEX-2 experiment. The maximum intensity is 160 Rayleighs.

PERPENDICULAR TRANSPORT

Chemically produced density irregularities become polarized by collisions with neutrals that are moving across magnetic field lines. The resulting electric fields cause ExB drifts of the ions and electrons and produce a distortion of the cross-section of the modified region. The cross-field motion of ionization enhancements from Class II reactions has been extensively studied [eg., Scannapieco et al., 1974]. Two dimensional simulations of this flow have been studied for chemical modification produced by electron attachment (Class IV reactions) by Bernhardt [1988] using a numerical model. These simulations included the effects of electron attachment and mutual neutralization chemistry, self-consistent electric fields, and three-species plasma transport but ion and electron inertia [ie., the right side of (2)] were neglected.

The results of these previous studies is briefly summarized as follows: Pedersen conductivity enhancements such as barium ion clouds and positive-ion, negative-ion clouds become polarized by neutral winds. The polarization fields cause motion of the clouds in the direction of the neutral wind. The down wind side of the enhancement steepens and eventually splits (bifurcates) into field aligned irregularities. In contrast, the polarization direction is reversed for Pedersen conductivity depressions such as initially produced by class III reactions or by Class IV reactions after the positive and negative ions have mutually neutralized. The plasma cavity drifts across B in the direction opposing the neutral wind. The ExB gradient drift instability produces irregularities on the upwind edge of the cavity.

We will now present several new results from a study of the release of electron attachment chemicals (see class IV reactions) directed at orbital velocities across the ambient magnetic field. A similar study for the injection of materials which photoionize (see class II reactions) has been studied by Mitchell et al. [1985].

The neutral cloud moving perpendicular to B begins to attach electrons to form a negative-ion positive-ion plasma cloud. Because the negative-ions are created with initial cross-field momentum, they become polarized. An electric field is set up across the plasma cloud that causes it to ExB drift in the direction of the neutral cloud. The distance the cloud moves across the magnetic field is an important aspect of the "skidding" process. Collisions with the background neutral particles will cause the neutral cloud as well as the plasma cloud to eventually slow down and stop. The radial expansion of the clouds initially depends on the release temperature and, as collisional interaction with the background becomes important, on the temperature of the atmosphere.

To study these processes, we have developed a two-dimensional plasma simulation. It is an extension of the one described by Bernhardt [1988] which was limited to low speed flow where the neutral wind dominated. In this simulation code the effects of the ion inertia on the right hand side of (2) have been included. Inertial effects become important at sufficiently high altitudes in the rarefied atmosphere where the ion neutral collision frequency becomes small. Generally, when the time between ion and the background neutral collisions becomes on the order of or larger than characteristic time scales of the phenomena under investigation, these effects must be incorporated into the theoretical model. Inertial terms in continuity equation introduce an effective inertial capacitance that causes (1) the initial potential to decay at a slower rate and (2) the region of highest potential to lag further behind the neutral cloud.

Figures 7,8 and 9 show the results of the simulations with and without the effects of ion inertia for the release of an electron attachment material at 7km/s across the magnetic field. In the figures, the magnetic field B is directed out of the page. The background plasma density (composed of O+ and e-) is assumed to be $4x10^5$ cm^{-3} while the background neutral density is $4x10^8$ cm^{-3}. The positive and negative ion-neutral collision frequencies are 0.61 and 0.14 s^{-1}, respectively. The left sides of the figures show (1) contours of the plasma Pedersen conductivity which are proportional to the ion densities and (2) the plasma flow vectors that are determined from the ExB velocity. The right sides contain contours of the electrostatic potential.

Figure 7 illustrates the evolution of the conductivity and potential after 1 second in the simulation. The material is released at the point (0,0) in the simulation box. Initially we see that the neutral cloud attaches electrons and forms negative ions. This produces an enhancement in the conductivity of the plasma and leads to the generation of an electrostatic potential. Figure 8 shows the collision dominated case where ion inertia has been neglected. After 16.7 seconds in the simulation, we see that the plasma cloud has expanded causing a significant enhancement in the plasma conductivity and has moved roughly 30 km. At this time, the plasma velocity is less than 100 m/s so the plasma cloud has essentially stopped moving with respect to the neutrals. The maximum potential has decayed significantly from the initial value

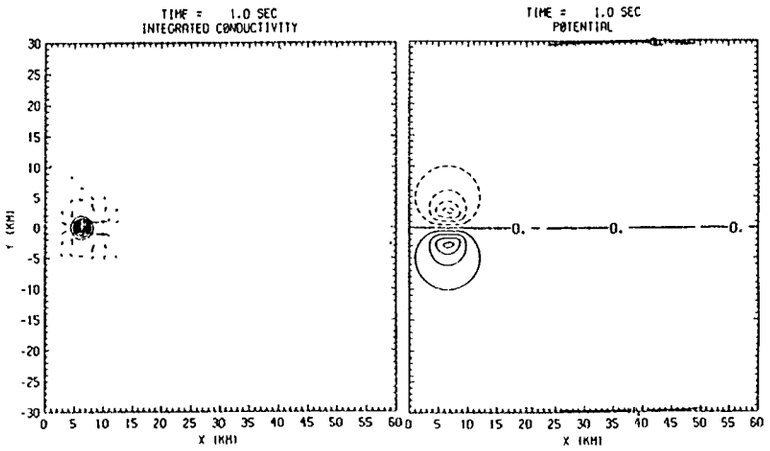


Figure 7. Field-line integrated Pedersen conductivity (mho), velocity field (km/s) and electrostatic potential (V) from the initial injection of an electron attachment vapor at orbital velocity (7 km/s). One second after release, the maximum plasma velocity has a magnitude of 3 km/s. The electric potential reaches 200 V inside the negative-ion cloud.

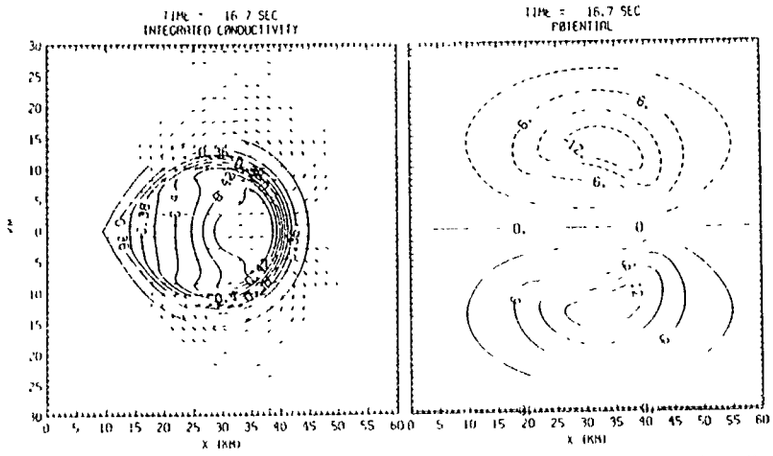


Figure 8. Integrated Pedersen conductivity (mho), plasma drift velocities (km/s) and electric potential (V) for a simulation without ion inertia. Sixteen seconds after release, the maximum flow vector has a magnitude of 0.82 km/s. The peak electric potential is 12 V.

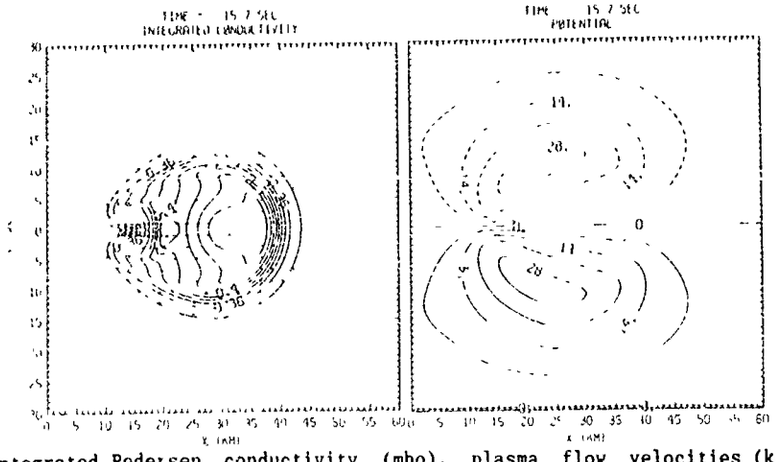


Figure 9. Integrated Pedersen conductivity (mho), plasma flow velocities (km/s) and electic potential (V) for a simulation which includes ion inertia. Sixteen seconds after release, the maximum flow velocity has a magnitude of 3.1 km/s. The peak electric potential is 30 V.

of about 200 Volts. From the neutral cloud contours (not shown) we determine that the region of largest conductivity is essentially collocated with the region of largest neutral cloud density. This is true in the inertial simulation as well.

Figure 9 illustrates the evolution of the plasma cloud including the effects of ion inertia, 15.7 seconds after the release. The main differences between this and Figure 8 are that the potential has decayed with a significantly slower rate and the region of the strongest potential lags behind the region of highest conductivity. Due to the relatively strong electric fields near the back edge, a velocity shear causes the cloud to bifurcate. This behavior should be observable in experiments. In general both the inertial and noninertial simulation runs show that a plasma cloud created by injection of electron attachment materials at orbital velocities may skid several 10's of km across the magnetic field.

HIGH POWER RADIO WAVES

Resistive heating by high power radio waves provides an energy source in the ionosphere to form large scale cavities [Duncan et al., 1988]. The resulting thermal cavities can refract the incident radio wave causing self-focussing. Simulation of these effects require coupling continuity equations (1) and equations of motion (2) to equations which describe the electron and ion temperatures and the propagating electromagnetic wave.

The energy equation for the electron gas is derived in a reference frame which moves with the $\underline{E}\times\underline{B}$ drift of the plasma. Heat flow is restricted to the direction of the ambient magnetic field. The one dimensional heat equation is

$$n_e \frac{3k}{2}\frac{\partial T}{\partial t}e + n_e v_s \frac{3k}{2}\frac{\partial T}{\partial s}e + n_e kT_e \frac{\partial v}{\partial s}s - \frac{\partial}{\partial s}\left[K_e \frac{\partial T}{\partial s}e\right] = Q_e - L_e \qquad (16)$$

where the coordinate "s" is along the magnetic field, v_s is the field aligned component of electron drift, K_e is the electron thermal conductivity, Q_e is the heat input to the electrons, and L_e is the heat lost to ions and to neutrals. A similar equation describes the ion temperatures. The artificial heat source for ohmic heating from high power radio waves is given by

$$Q_{e\omega} = \varepsilon_0 |E_\omega|^2 X \nu_e/2 \qquad (17)$$

where $|E_\omega|$ is the magnitude of the electric field of the electromagnetic wave with frequency ω, $X = \omega_p^2/\omega^2$, ω_p is the plasma frequency, ε_0 is the dielectric constant of free space, and ν_e is the total electron collision frequency. Conversion of electromagnetic waves to slowly propagating plasma waves can be another heat source to the plasma.

For a single frequency time variation $e^{i\omega t}$, the wave equation which describes both electromagnetic and electrostatic wave propagation is written as:

$$-\nabla^2 \underline{E}_\omega + [I - \underline{M}\delta_e]\nabla(\nabla\cdot\underline{E}_\omega) - k_0^2[I-\underline{M}X]\underline{E}_\omega = 0 \qquad (18)$$

where

$$\underline{M} = \begin{bmatrix} U & iY_z & -iY_y \\ -iY_z & U & iY_x \\ iY_y & -iY_x & U \end{bmatrix}^{-1}$$

$U = (1-i\nu_e/\omega)$, $\underline{Y} = e\underline{B}/m_e\omega$, $k_0 = \omega/c$, $\delta_e = v_{Te}^2/c^2$, v_{Te} is the electron thermal velocity, and c is the speed of light. Details about the wave equation and the notation are given by Yeh and Liu [1972]. Equation (18) should be used for a complete electrostatic and magnetoionic treatment of radio wave coupling to the plasma.

Bernhardt and Duncan [1982, 1987] have have developed a model of ionospheric heating which incorporates equations (1), (2), (16), (17), and (18) with the following simplifications:

(1) The propagation is assumed to be isotropic so $\underline{Y}$ is set to zero.
(2) The heating is underdense ($\omega > \omega_p$) so that $X<1$ everywhere.
(3) The conversion to plasma waves is neglected.

The thermal cavity produced by a 40 $\mu W/m^2$ beam at 4.4 MHz is shown in Figure 10. The HF electromagnetic waves heat the plasma continuously for 1600 seconds. The electron temperature

near the center of the beam at the peak of the layer is elevated by 400 K. The maximum density is reduced by ten percent. For this simulation, the ambient plasma is drifting northvard at 50 m/s. Asymmetric thermal cavities are produced by the 50 degree inclination of the ambient magnetic field and by an imposed 50 m/s northward plasma drift. Continuous wave heating of the drifting ionospheric should eventually causes a depression in the plasma down stream from the heated region.

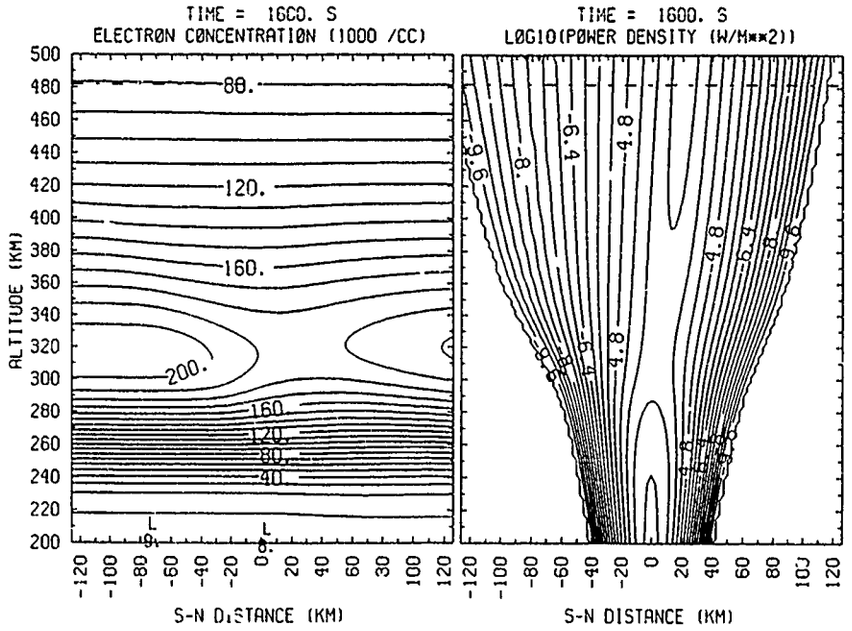


Figure 10. Computed thermal cavity produced by a high power radio wave with a frequency of 4.4 MHz. The critical frequency of the unmodified layer is 4.1 MHz.

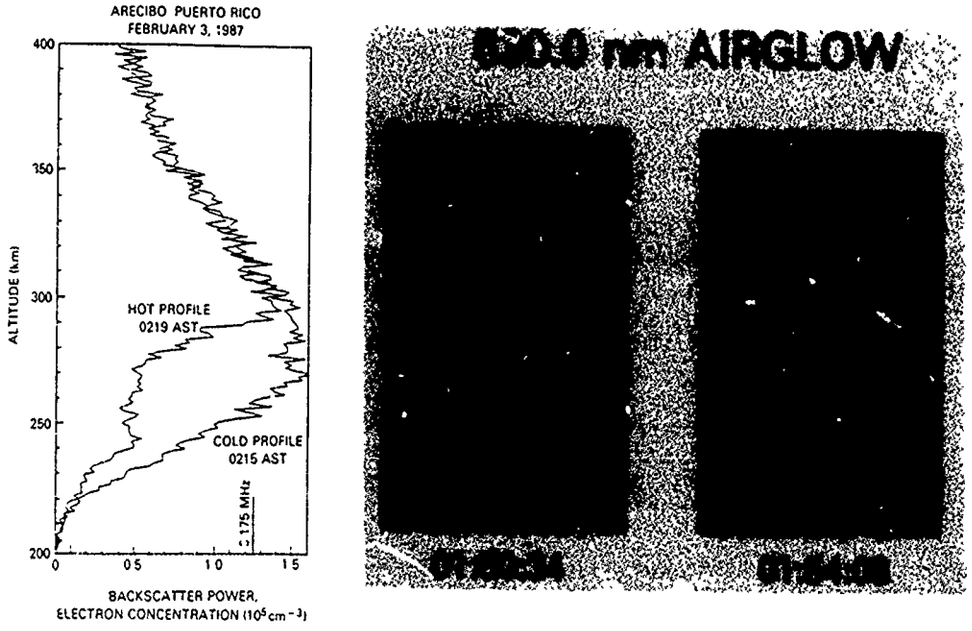


Figure 11 Ionospheric profiles and airglow clouds in a drifting ionosphere. The frequency ot the high power radio wave 3.175 MHz. Comparison of the hot and cold profiles shows the effects of the 50 μW/m2 beam. The airglow clouds outline the elliptical beam which has a major axis in the north-south (top-bottom) direction. The airglow clouds and associated cavities drift to the west (right) at 50 m/s.

Simultaneous radar and optical measurements at Arecibo during ionospheric heating experiments provide the size and location of large scale plasma cavities. The radar profiles in Figure 11 show the modified (hot) and unmodified (cold) ionospheres. A density reduction of 40 percent is estimated from the radar backscatter data. Energetic electrons accelerated out of the heated region produce the airglow clouds seen on the right side of Figure 11. The motion of these clouds indicates convection of the ionospheric cavity which "entrains" the heater beam.

Recent observations by Bernhardt et al. [1988c, 1989a,b] have demonstrated that continuous, overdense ionospheric heating can produce periodic structures in the ionospheric. A schematic of this process is shown in Figure 12. When a transmitter is turned on at time t=0, the ray paths interact with a horizontally stratified plasma. Some of the rays near the center of the vertical beam reflect where the wave frequency ω is equal to the plasma frequency ω_p. This resonance region is where the largest amplitude electric fields are found and where most of the energy is absorbed into the F-region. As a cavity forms and begins to drift, the ray paths become refracted away from the zenith and a resonance region may no longer be found. If this happens, the regeneration rate for the cavity is greatly decreased, the cavity drifts away, and the radio beam "snaps-back" to the zenith.

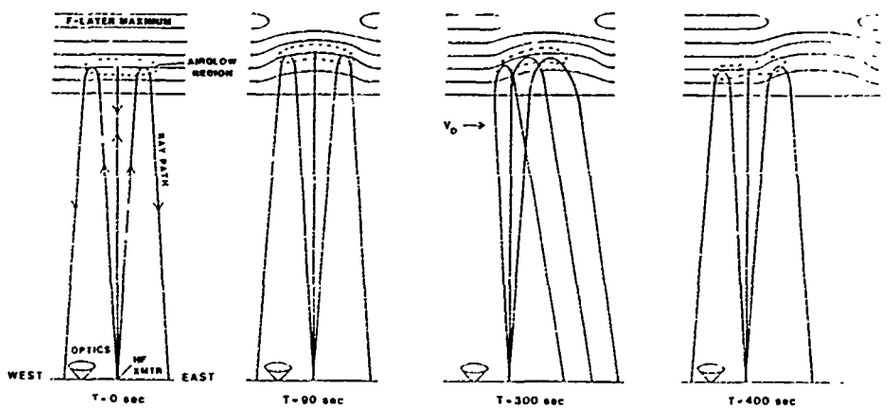


Figure 12. Schematic of the snap-back process during continuous ionospheric heating. This process leads to the generation of periodic irregularities.

Our coupled fluid, electromagnetic model cannot simulate overdense ionospheric heating. We have, instead, produced a different type of model to demonstrate the generation of irregularities in a convecting ionosphere. Within the restrictions of geometric optics one can describe the motion of an electromagnetic wave packet by a Hamiltonian (Landau and Lifshitz,1971). For the case of a wave packet travelling through a plasma, the equations of motion can be written in the same form as those for a single particle moving through a potential (Stix, 1963). Beyond the requirements that the that spatial gradients be long compared to the wavelength of the radiation and that changes in the medium occur on time scales long compared to the radiation's frequency, to be able to use this approach, one must be able to differentiate the local dispersion relation for the mode of interest. This opens up the possibility of using many of the techniques of particle simulations to study the large scale propagation of electromagnetic waves in spatial and time varying plasmas.

The case of a high frequency electromagnetic mode in an unmagnetized plasma is particularly straightforward. This branch obeys a dispersion relation

$$\omega^2 = \omega_p^2(z) + (kc)^2 , \qquad (19)$$

where ω is the frequency of the wave, ω_p is the local plasma frequency, k is the local wavevector, z is the vertical coordinate, and c is the speed of light. The motion of the wavepacket is $\partial z/\partial t = v_g$ where v_g is the group velocity. The group velocity is $\partial\omega/\partial k$. Therefore,

$$\partial z/\partial t = (c^2/\omega)\, k. \qquad (20)$$

We need an equation of motion for k which can be obtained by differentiating the dispersion relation,

$$2c^2 k\, (\partial k/\partial t)\, (\partial t/\partial z) = -\partial\omega_p^2(z)/\partial z. \qquad (21)$$

This reduces to

$$\partial k/\partial t = -(2\omega)^{-1}\ \partial\omega_p^2(z)/\partial z. \qquad (22)$$

Now $\omega_p^2(z) = (q_e/m_e)\ n(z)$. Letting $v' = c^2k/\omega$ and $n'(z) = n(z)/n_{e0}$, the equations of motion for the wave packet is

$$\partial z/\partial t = v' \text{ and } \partial v'/\partial t = -\alpha\ \partial n'(z)/\partial z, \qquad (23)$$

where $\alpha = (1/2)(c\omega_{p0}/\omega)^2$. The density of the plasma acts as a potential field which determines the motion of the packet.

To perform a preliminary study of the effects of resonant heating by the RF beam, a simple model of a drifting plasma was constructed. The density is initialized as an increasing Gaussian in the z-direction and uniform in the horizontal x-direction. The initial normalized temperature is uniform (= 1.0). A third array, C(x,z), is the product of n(x,z) and T(x,z) and does not change in time. These three arrays are shifted to the right in the x direction by one cell each time step. The values of the three arrays on the left hand boundary are not changed so that an unperturbed plasma flows into the simulation box. The wave packets are injected into the plasma at a constant rate of eight packets per time step at z=0 and at x= 15 cells. V_z' is chosen so that the critical layer is approximately half way up the density gradient. V_x' has a magnitude of 20% of v_z' and is multiplied by a uniform random number between -1 and +1 so that the beam has an angular spread. Each wave packet carries a fixed amount of energy (ΔQ). When $v_z' < 0$, this indicates that the particle has reached the critical layer and the energy is deposited in the plasma. This is done by increasing the temperature at the nearest grid point and its six closest neighbors.

$$T(x,z) = T(x,z) + W* \Delta Q \qquad (24)$$

where W is a linear weighting. The particle is then removed from the simulation. The density is changed by $n(x,z) = C(x,z) / T(x,z)$. This new density determines the forces acting on the wave packets.

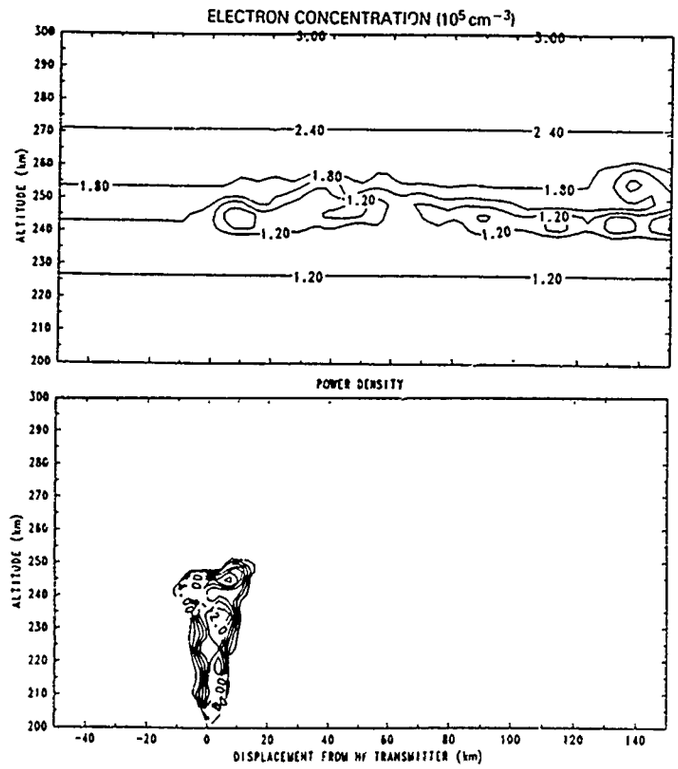


Figure 13. Simulation of a structured plasma layer produced by continuous wave heating. The irregularities tend to form with a spatial period of 40 km. The ambient plasma drift is 100 m/s from left to right.

Figure 13 shows the electron density contours and the skewed radio beam at t=1500 seconds. An unperturbed plasma flows in from the left hand boundary. As it flows over the transmitter energy is absorbed at the critical layer, the temperature increases and a cavity is formed. At the critical layer the width of the beam is approximately 20 km. The plasma is flowing to the right and the structures to the right of the transmitter were formed at earlier times. The modification of the density is not smooth but is bursty in nature forming localized cavities. The density gradients in the x direction act to focus the beam. The plasma drift then acts to bend the beam to the right. By reducing the density at the critical layer it is possible for the beam to reach higher altitudes. This can be seen in the structure at 140 km where a cavity was formed above the ambient critical layer.

The dynamics of this interaction is still being studied. Injecting the same number of wave packets but reducing the amount of energy carried by each wave packet by 2 orders of magnitude leads to a much smoother response. The density at the critical layer is reduced by one percent and one does not see the large oscillations and cavity structures. Also, one does not see the bending of the beam.

The ionospheric cavities produced by high power radio waves are subject to the effects of polarization by neutral winds. This has been discussed in a previous section with reference to ionospheric holes produced by chemical releases. A neutral wind can polarize a thermal cavity to cause it to steepen on the up wind side . We are currently investigating the effects of this steepening on the high power radio waves with respect of beam propagation and beam stability, eg., thermal self-focussing instabilities. In a stationary heated ionosphere, a plasma depression acts like a convergent lens to further intensify the electromagnetic beam. Neutral convection relative to the plasma may induce cross-$\underline{B}$ drifts which further intensify the density gradients on one side of the cavity. As a result, temperature gradients due to ohmic heating will appear across the cavity. Preliminary calculations show that these density and temperature gradients, in conjunction with convection, will have a stabilizing influence on the thermal self-focussing instability.

CONCLUSIONS

Ionospheric modification by chemical releases and by high power radio waves provide complementary means for generation of large scale plasma irregularities. Chemical releases produce localized density gradients either by enhancing or depleting the electron concentrations. Ambient neutral winds and electric fields will transport modified region away from the release location. Nonuniform velocity gradients can steepen the density gradients and may eventually yield small scale structures. The initial velocity from orbital releases can stiumulate velocity-shear (ie., Kelvin-Helmholtz) instabilites which yield irregularities on the back sides of the clouds.

Ground based HF transmitters with equivalent radiated powers of 80 MW or more can produce thermal cavities near where the wave frequency equals the plasma frequency of the ionosphere. At high enough powers, periodic irregularities can be produced in a convecting layer. Even with continuous wave transmission, a relaxation oscillator type process can produce multiple snap-backs of the heater beam. As with chemical releases, once a large scale cavity is formed, $\underline{E}\times\underline{B}$ gradient drift processes may yield shorter scale sizes along their edges.

ACKNOWLEDGEMENTS

This work is supported by the NASA CRRES Projects Office and the Office of Naval Research.

REFERENCES

Bernhardt, P.A. and L.M. Duncan, "The feedback-diffraction theory of ionospheric heating," J. Atmos. and Terr. Phys., 44, No. 12, 1061, 1982.

Bernhardt, P. A., E. J. Weber, J. G. Moore, J. Baumgardner, and M. Mendillo, "Excitation of oxygen permitted line emissions by SF_6 injection into the F region", J. Geophys. Res., 91, 8937-8946, 1986.

Bernhardt, P.A., "A critical comparison of ionospheric depletion chemicals," J. Geophys. Res., 92, 4617, 1987.

Bernhardt, P.A. and L.M. Duncan, "Ionospheric focused heating - a hybrid modification experiment," J. Atmos. Terr. Phys., 49, 1107, 1987.

Bernhardt, P.A., "Cross-B convection of artificially-created negative-ion clouds and plasma depressions: low speed flow," J. Geophys. Res., 93, 8696-8704, 1988.

Bernhardt, P.A., B.A. Kashiwa, C.A. Tepley, and S.T. Noble, "Spacelab 2 upper atmospheric modification experiment over Arecibo, 1, neutral gas dynamics," Astrophys. Lett. and Comm., 7, 169-182, 1988a.

Bernhardt, P.A., W.E. Schwartz, M.C. Kelley, M.P. Sulzer, S.T. Noble, "Spacelab 2 upper atmospheric modification experiment over Arecibo, 2, plasma dynamics," Astrophys. Lett. and Comm., 7, 183-198, 1988b.

Bernhardt, P.A., L.M. Duncan, and C.A. Tepley, "Artificial Airglow Excited by High-Power Radio Waves," Science, Volume 242, pp 1022-1027, 18 November 1988c.

Bernhardt, P. A., L. M. Duncan, and C. A. Tepley, Heater-induced cavities as optical tracers of plasma drifts", J. Geophys. Res., 94, 7003-7010, 1989a.

Bernhardt, P. A., C. A. Tepley, and L. M. Duncan, "Airglow enhancements associated with plasma cavities formed during ionospheric heating experiments", J. Geophys. Res., 94, 9071-9092, 1989b.

Bernhardt, P.A., Parallel-to-B dynamics of chemically induced perturbations to the ionosphere, submitted to J. Geophys. Res., 1990.

Duncan, L.M., J.P. Sheerin, R.A. Behnke, Observations of ionospheric cavities generated by high-power radio waves, Phys. Rev. Lett., 61, 239-242, 1988.

Fulford, J.A., J.W. MacDougall, P.A. Forsyth, M. Mendillo, and P.A. Bernhardt, Ionospheric electron content measurements during the second space-plasma negative-ion experiment (SPINEX-2), Canadian J. Phys., 65, 403, 1987.

Mitchell, H.G., J.A. Fedder, J.D. Huba, and S.T. Zalesak, Transverse motion of high-speed barium clouds in the ionosphere, J. Geophys. Res., 90, 11091-11095, 1985.

Narcisi, R.S., and E.P. Szuszczewicz, Plasma composition and structure characterization of an ionospheric barium cloud, Active Experiments in Space, pp. 299-304, Alpbach, Austria, 24-28 May 1983.

Scannapieco, A.J., S.L. Ossakow, D.L. Book, B.E. McDonald, and S.R. Goldman, Conductivity ratio effects on the drift and deformation of F region barium clouds coupled to the E region ionosphere, J. Geophys. Res., 79, 2913, 1974.

Stix, T.H., The Theory of Plasma Waves, McGraw-Hill Book Company, New York, pp. 55-59, 1962.

Yeh, K.C., and C.H. Liu, Theory of Ionospheric Waves, Academic Press, New York, pp. 130-307, 1972.

Schunk, R. W., and E.P. Szuszczewicz, Early time plasma expansion characteristic of ionized clouds in the ionosphere, J. Geophys. Res., 93, 12901-12915, 1988.

AD-P006 268

FORECASTING MORPHOLOGY AND DYNAMICS OF F-LAYER IRREGULARITIES

Jules Aarons
Center for Space Physics
Boston University
725 Commonweath Ave
Boston, MA 02215

ABSTRACT

Data and analyses are at hand to forecast the morphology and dynamics of global F-layer irregularities. For the planners of systems which are impaired by the scattering characteristics of F-layer irregularities, forecasting the morphology allows them to evaluate the utility of operating systems and to plan means for integrating back-ups. For the active users forecasting the dynamics of changes in intensity and characteristics of the irregularities as a function of geophysical conditions allows for warning operators about impending problems.

For the polar region the morphological parameter used as the principal forcing function for intensification of F-layer irregularities is the solar flux. Intense irregularities appear during high sunspot number years. For the auroral region the parameter used as forcing function is magnetic activity as shown by studies of radio stars, low altitude satellites and synchronous satellites. Increased solar flux plays a role both pushing the F-layer irregularity region equatorwards in latitude and increasing the quiet day intensity. For the sub-auroral region the magnetic activity during the injection phase of the magnetic storm plays a leading role. During the recovery phase of the magnetic storm the effect of the decay of ring current ions on the sub-auroral region is such that irregularities are formed of low intensity without auroral or magnetic activity.

For the equatorial region the effect of solar flux is prominent. With high solar flux the plume regions show greater intensity and greater height. With high solar flux the anomaly regions at the edges of the equatorial zone show high electron densities at sunset hours before and during the appearance of irregularities. When irregularities appear intense microwave scintillations are observed during the pre-midnight time period. During years of low solar flux the anomaly region fails to show such levels. Studies with synchronous satellites and with optical observations have allowed the morphology and nature of the irregularities in this region to be clearly understood. Another "forcing function" to be noted is the importance of the month of the year at various longitudes with a lowering of irregularity occurrence and intensity during December and January in the Pacific sector and a lowering of these levels in June, July and August in the 0-75° West sector.

INTRODUCTION

For forecasting ionospheric propagation parameters we have two distinct needs. The planner is only concerned with the effects of magnetic storms such as were observed March 13th and 14, 1989 as a piece of his statistics, perhaps the 99 percentile bracket. The user of satellite communications at 250 MHz is dismayed when auroral effects on transmissions are noted as far from the auroral oval as Mexico and Puerto Rico. We need morphology and dynamics.

Data and analyses are at hand to forecast the morphology and dynamics of global F-layer irregularities. For the planners of systems which are impaired by the scattering characteristics of F-layer irregularities, forecasting the morphology allows the evaluation of the utility of operating systems and the planning of back-ups. For the active users, forecasting the dynamics of changes in intensity and characteristics of the irregularities as a function of geophysical conditions allows the sytem to warn operators about impending problems.

In addition the needs of users, even real time users, differ so that one aspect to be considered is the relevant parameter. The equatorial ionosphere, the scene of propagation problems and interesting science, can be used at night for trans-equatorial propagation; amateur radio operators try to set records of the length of the path of VHF transmissions across the equator during nights of high sunspot number years. Reflection or scattering of ground to ground transmissions however is affected by the peak of the F-layer irregularities. At the peak of the last sunspot cycle, amateur radio operators recorded transmissions from S. Africa to Athens. Two way distances of 4475 miles and one way of 4930 miles were recorded with scattering off the anomaly region (Cracknell et al, 1981).

Scintillation of satellite radio signals is an integration of the scattering that takes place along the whole path, thus the thickness of the scattering layer is important. On some days a thin layer is formed at F layer heights on the magnetic equator On other nights a plume extending to 1000 km is formed. Observers of GPS in its public mode of 1.6 GHz are dismayed by garbled transmissions over many hours when they observe from sites in the equatorial anomaly region such as Ascension Island or Hong Kong; a model of scintillation activity would have told them of problems However a model of scintillation activity and intensity is not totally relevant for ground to ground transmissions. A model of all seasons is not a model for all purposes

91-09661

SPREAD F AND OLDER STUDIES

Even the gross aspects of morphology may differ from what has on occasion been discussed as the same parameter. In the equatorial region we are confronted by the term Spread F used for describing F layer irregularities. The term is not only a misnomer but is misleading. We know for example that of the two kinds of spread F as determined by ionsondes, only range spread at the equator is correlated with scintillation; frequency spread is not. (This is not necessarily the case at high latitudes). Both types are relevant to HF transmissions in which there is renewed interest since new methods of coding, interleaving, and redundancy are being incorporated into HF systems as well as advances in technology. In this area h'F2 and foF2 are important rather than models of Total Electron Content which have their utility in establishing accuracy in navigation and radar systems.

Some of the earlier studies of spread F established general areas of interest but failed to include some areas of importance. A rather nice study in 1965 (Kumar and Rao, 1965) with results illustrated in Figure 1 failed to show the importance of the anomaly region in equatorial microwave scintillation. However it did show the negative correlation of middle latitude spread F with sunspot number. Another study of high latitude spread F (Figure 2, (Penndorf, 1962) indicated a permanent maximum near Foxe Basin which recent studies have failed to find and more importantly failed to show the importance of the polar region even though the data was taken during a year of high solar flux. However the study did show spread F maxima moving with the auroral oval Most spread F studies have concentrated on the occurrence of irregularities rather than their intensity.

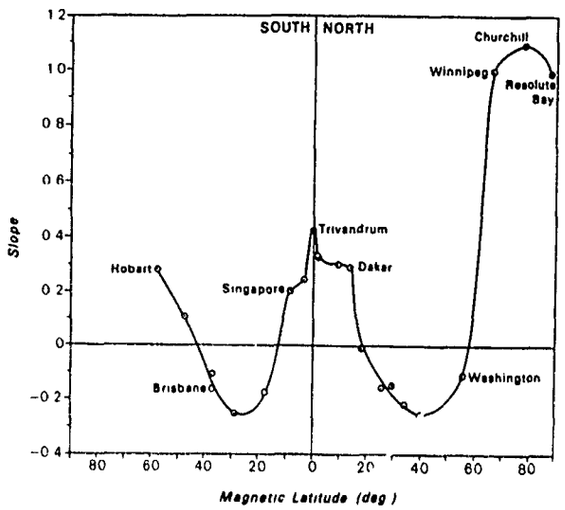


Figure 1. Spread F data of Kumar and Rao (1965) indicate the correlation of sunspot number with the occurrence of spread F. It is of interest to note the negative correlation of spread F occurrence with sunspot number in the middle latitude region.

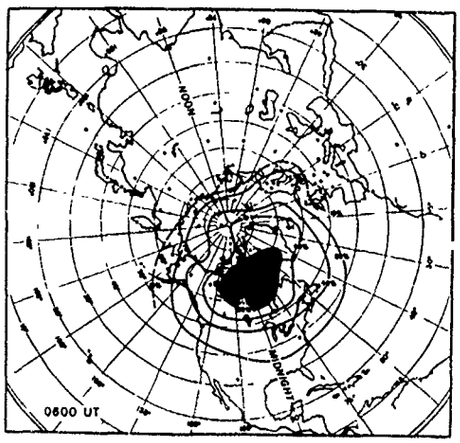


Figure 2. Contours of spread F at 0600 UT in September, 1957 (Penndorf, 1965). The contours indicate a low latitude of effective spread F in the auroral region in the midnight time period but fail to show the polar peak

THE F-REGION IRREGULARITY WORLD ACCORDING TO AARONS AND ACCORDING TO BASU

Now that we have established the need for a plethora of models and statistics, essentially a range of designer models, we should see where we are relative to some of the problems.

Figure 3 is a cartoon of the nighttime F-layer irregularities affecting scintillation during years of high solar flux. The word cartoon is used to reflect the broad brush somewhat inaccurate nature of the picture. It is an updated version of similar global maps presented in the past (Aarons, 1982). It does establish the importance of the auroral, the polar, and the equatorial regions for this parameter.

Figure 4 from Santimay Basu (personal communication) is an accurate picture of the 1.2 GHz intensity of scintillation activity during various periods of solar flux. Figure 3 is used to frighten users; Figure 4 makes users demand to know the complete statistics. These are given in the best references of the occurrence and intensity of both phase and amplitude fluctuations (Basu et al, 1985 and Basu et al, 1988).

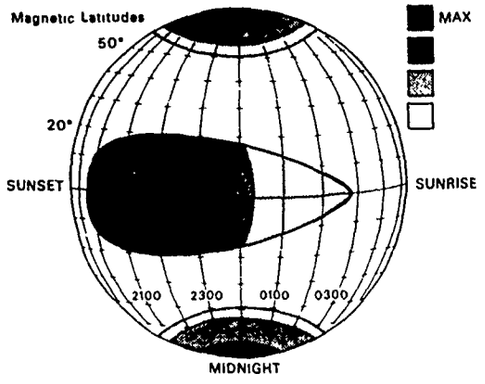


Figure 3. The F-Layer Irregularity World According to Aarons

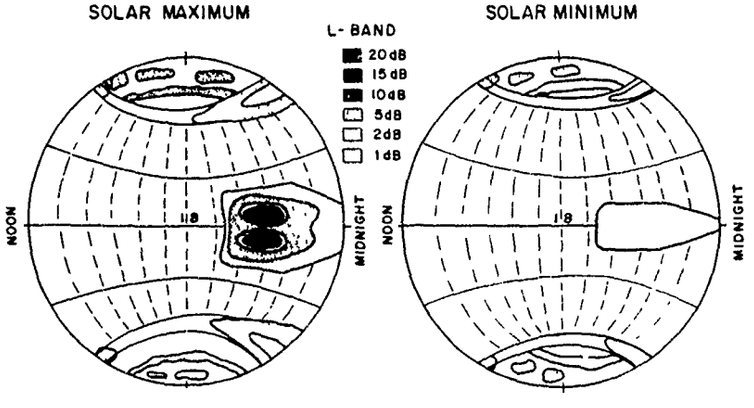


Figure 4. The F-Layer Irregularity World According to S Basu.

Much research has been done on the effect of magnetic activity on F layer irregularities at auroral and at equatorial latitudes. At high latitudes a review of the data indicates a rather complex picture with a seasonal, latitudinal and longitudinal dependence. In most of the studies (e.g. Rino and Matthews, 1980), the data are analyzed by sorting into magnetically disturbed and quiet days depending on whether the sum of Kp or Ap is greater than or less than a particular value. However the effects of magnetic storms on F region irregularities undoubtedly depend on the phases of the storms. For a particular observing point this means a correlation of the effects of storm development as a function of local time. This has been shown to be amply true in the case of total electron content obtained from satellite observations (Mendillo et al, 1972). In the case of equatorial F-layer irregularities several studies have shown correlation with the phase of the storm in the equatorial region (Dasgupta et al., 1985 and at auroral latitudes (Aarons et al, 1988 and Kersley and Chandra, 1988).

THE POLAR REGION

For the polar region the forcing function for intensification of F-layer irregularities is the solar flux. Intense irregularities appear during high sunspot number years (Basu, Su. et al, 1985). It is a complex region with large scale patches moving primarily anti-sunwards as is shown in Figure 5 (E. Weber, private communication and Weber et al, 1984). The patches produce intense activity during years of high solar flux but produce very low level scintillation at 250 MHz during years of low solar flux.

Early studies of spread F in the polar region failed to note the intense activity during years of high solar flux and the dramatic decrease of irregularity intensity during years of low solar flux. In a recent 48 hour test period in 1989, A. Johnson (personal communication) found that 90% of the time at 250 MHz there was intense scintillation activity of 30 dB peak to peak. The polar region as a function of month of the year and of local time clearly shows the effect of solar illumination with the maximum occurrence during the winter months and during hours when the sun is considerably below the horizon.

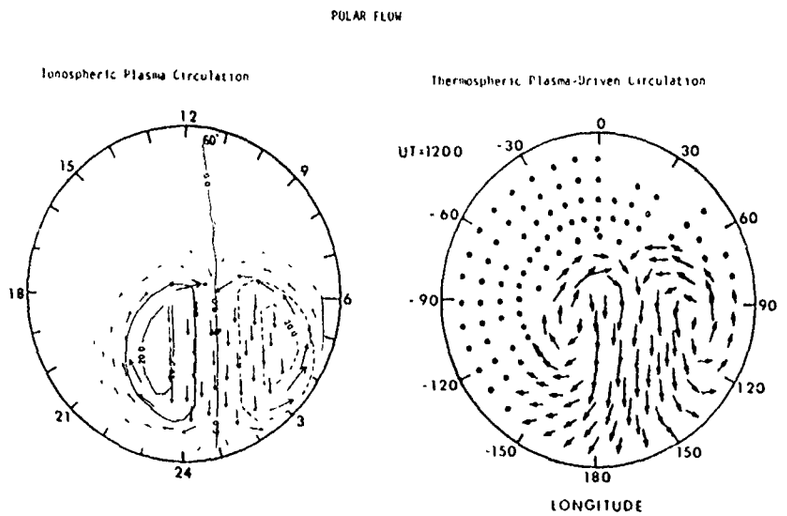


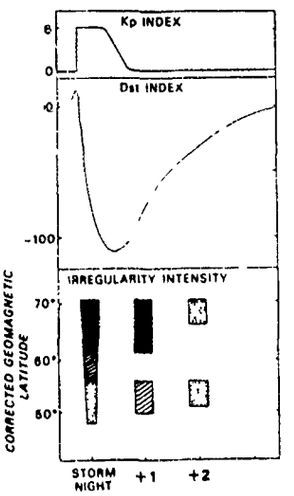
Figure 5 Ionospheric plasma circulation which brings patches from the dayside across the pole (E. Weber, private communication).

THE AURORAL REGION

In the high latitude region, F layer irregularities are found to be generated at auroral latitudes during the injection phase of the magnetic storm, the strength and duration of which vary with the intensity and duration of the magnetic activity The effect during this initial phase of a magnetic storm is proposed to be the prompt electric field penetration to low latitudes due to the rate of motion of the equatorward edge of the auroral zone electric fields, the shielding layer (Fejer et al, 1989 and Foster and Aarons , 1989). When the magnetic storm develops, the conditions which produce F-layer irregularities descend to what had been sub-auroral latitudes. During the recovery phase of the magnetic storm, the effect of ring current ions on the sub-auroral region is such that irregularities are formed of low intensity without auroral or magnetic activity. The appearance of intense F layer irregularities at sub-auroral latitudes over several days in the recovery phase of magnetic storms is in a complex manner correlated with the level and recovery pattern of Dst, the measure of ring current energy density (Aarons and Rodger, 1989).

From the modeling point of view, the irregularity data needs a geophysical parameter around which to organize occurrence and intensity. For the auroral region the forcing function is magnetic activity as shown by studies of low altitude satellites such as WIDEBAND and TRANSIT as well as earlier work with radio stars and synchronous satellites. Increased solar flux plays a role both pushing the F-layer irregularity region equatorwards in latitude and increasing the quiet day intensity. For the sub-auroral region the magnetic activity during the injection phase of the magnetic storm plays a leading role. During the recovery phase of the magnetic storm the effect of ring current ions on the sub-auroral region is such that irregularities are formed of low intensity without auroral or magnetic activity.

A descriptive model of the dynamic effects of great magnetic storms was developed by A.S. Rodger and myself. In Figure 6 we show the apparent motion of the effects of the magnetic storm, in reality the electric field penetration as a function of latitude. It first affects the auroral region with very close correlation between the form of magnetic activity and the timing of scintillation at auroral latitudes. As time proceeds the effects descend to lower latitudes encompassing regions termed in quiet magnetic periods sub-auroral latitudes. During the injection phase the ring current stores the ionospheric and solar ions. After the storm, the ring current decays and this decay is the source for the generation of plasmapause or sub-auroral irregularities - and Stable Auroral Red Arcs.



MODEL OF IRREGULARITY INTENSITY DURING AND AFTER MAGNETIC STORM COMMENCEMENT

Figure 6. High latitude development of F-layer irregularities according to the model of Rodger and Aarons. During the initial injection phase of the magnetic storm, the electric field penetrates to lower latitudes. However with the end of the magnetic storm the ring current decays, producing the conditions for irregularity development at sub-auroral latitudes.

THE EQUATORIAL REGION

For the equatorial region the effect of solar flux is prominent. With high solar flux the plume regions show greater intensity and greater height. With high solar flux the anomaly regions at the edges of the equatorial zone show high electron densities at sunset hours before irregularities appear and these high electron densities continue after sunset. When irregularities appear intense microwave scintillations are observed during the pre-midnight time period. During years of low solar flux the anomaly region fails to show such levels. Studies with synchronous satellites and with optical observations have allowed the morphology and nature of the irregularities in this region to be clearly understood. Another "forcing function" to be noted is the importance of the month of the year at various longitudes with a lowering of irregularity occurrence and intensity during December and January in the Pacific sector and a lowering of these levels in June, July and August in the 0-75° West sector.

The only experimental evidence of the creation of a patch in the equatorial region comes from observations in the post-midnight sector when a change in Dst produces an eastward electric field. If the excursion in Dst appears in the post-midnight time periods, the irregularity generation will be enhanced with the eastward electric field producing a rise in F layer height during a time period when the height normally falls. The layer height after rising then falls and creates irregularities. This is the only observed cause and effect generation noted in equatorial irregularity generation. This was noted in the case of range spread generation (Rastogi and Woodman, 1978), scintillation of radio beacon signals (Rastogi and Aarons 1980), and in a statistical analysis of many cases by Das Gupta et al (1985)

During years of high solar flux and during months of activity, the nightly occurrence of equatorial irregularities on the equator is the norm not the exception. With this in mind the dynamics of the equatorial region are such that the creation of post-sunset patches must be looked at from a different point of view from that discussed i.e. the seeding of irregularity patches. For the months of the year when irregularities are normally created, one should look at spoilers i.e. conditions which prevent the normal occurrence of patches rather than triggers. The equatorial region shows inhibition of patches of irregularities during the injection phase of major magnetic storms On the other hand one can note triggers when patches are normally not observed i.e. during periods of high ring current negative excursions

COUPLING BY ACTION OF THE RING CURRENT

For ionospheric propagation studies, the correlation of Dst with ionospheric parameters at equatorial and sub-auroral latitudes means that one should probe into the ring current which is the source of Dst excursions as well as its constant value. Essentially the ring current, present from 2-7 earth radii, consists of hydrogen, oxygen , nitrogen and other ions trapped in the magnetosphere. The ions come predominantly from the ionosphere with some highly charged ions from the solar wind.

For the equatorial region the ring current depression of the geomagnetic field has the result of decreasing the altitude of the F2 layer in the time period before sunset.

For the higher latitude regions the effects of the ring current are more difficult to assess. In part this is due to the assymetry of the ring current as is shown in Figure 7 from Roelof and Williams (1988). In addition the composition changes from storm to storm. During both great and lesser magnetic storms the proton flux increases. However protons as a percentage of total energy density do not increase as much as the O+ and N+ flux (Lui et al, 1987). Thus the interaction of ring current ions with the ionosphere in the charge exchange process is rather difficult to assess. The assymetry, the ion composition, the energy of the ions make the process of interaction difficult to assess for any particular storm. While the ring current appears to be the source of much of the sub-auroral F layer irregularities, the Dst measure may not reflect in detail the development of those parameters of importance to irregularity generation. The Dst measure in the February 6-9, 1986 storm was well correlated with the inner ring current energy density from storm maximum well into recovery but the correlation was not good during the long developing main phase of this storm (Hamilton et al, 1988).

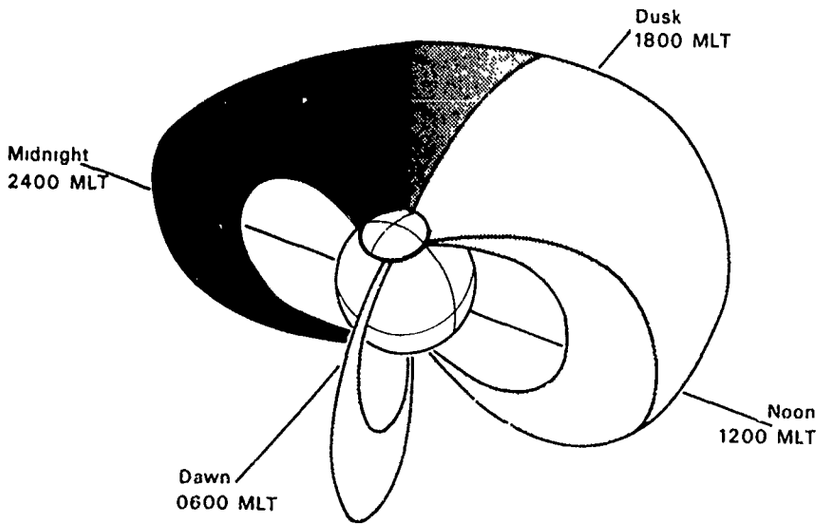


Figure 7. The Ring Current according to Roelof and Williams (1988) Note that the assymetry of the ring current indicates much higher levels in the midnight time sector

FORECASTING

Data and analyses are at hand to forecast the morphology and dynamics of global F-layer irregularities. for the intense irregularities, the forcing parameters are the form, level, and development of the magnetic storm. During any year we may have thousands of users of the transmissions for the satellites in the 250 MHz range. The effect of amplitude fluctuations can be minimized; it is up to the user community to demand attention to this problem.

While we have a handle on forecasting dynamics and morphology we still need to find a way to communicate, not only point to point but scientist to user.

ACKNOWLEDGEMENTS

Support for these studies has been from the Office of Naval Research.

REFERENCES

Aarons, J., Global morphology of ionospheric scintillations, Proc. IEEE 70 360-378, 1982.

Aarons, J., C. Gurgiolo, and A.S. Rodger, The effects of magnetic storm phases on F layer irregularities below the auroral oval, Radio Sci., 23, 309-319, 1988a.

Aarons, J. and A.S. Rodger (1989) Sub-auroral F-layer irregularities and the ring current Ann. Geophysicae 7, (2), 169-176

Basu, Su., S. Basu, E. MacKenzie and H.E. Whitney (1985) Morpholgy of phase and intensity scintillations in the auroral oval and polar cap Radio Science 20, 347-356

Basu,S., E. MacKenzie, and Su. Basu (1988) Ionospheric constraints on VHF/UHF communication links during solar maximum and minimum periods, Radio Science 23, 363-378

Cracknell,R., F. Anderson, and C. Fimerelis (1981) The Euro-Asia to Africa VHF transequatorial circuit during KSolar Cycle 21, QST, Nov. 1981, 31-36

Dasgupta, A., A. Maitra, and S.K, Das, Postmidnight equatorial scintillation activity in relation to geomagnetic disturbances, J. Atmos. Terr. Phys., 47, 911-916, 1985.

Fejer, B.G., E.R.de Paula, I. S. Batista, E. Bonelli, and R.F. Woodman (1989) Equatorial F region vertical plasma drifts during solar maxima, J. Geophys. Res. 94, A9, 12,049-12,054

Foster, J.C., and J. Aarons (1988) Enhanced anti-sunward convection and F-region scintillations at mid-latitudes during storm onset J. Geophy. Res. 93, 11,537-11,542

Hamilton, D.C. , G. Gloeckler, and F.M. Ipavich, W. Studemann, B. Wilken and G. Kremser (1988) Ring current development during the great geomagnetic storm of February 1986 J. Geophys Res 93, 14, 343-14, 355

Kersley, L., and H. Chandra (1988) Scintillation at high latitudes during winter magnetic storms J. Atmos. Terr. Phys., 50, 565-572

Kumar, V.R.J. and Rao, B.R. (1965) Latitude effect on the variation of spread F occurrence with sunspot number, Nature, 207, 1083-1084

Lui, A.T.Y., R.W. McEntire, and S.M. Krimigis (1987) Evolution of the ring current during two geomagnetic storms J. Geophys Res. 92, 7459-7470 (July 1)

Mendillo, M., M. D. Papagiannis, and J A Klobuchar (1972) Average behavior of the mid-latitude F-region parameters NT, N max, and t during geomagnetic storms , J. Geophys. Res. 77,4891, 1972

Penndorf, R. (1962) Diurnal and seasonal variation of spread F in the Arctic, J. Geophys. Res. 67, 2289-2298

Rastogi, R. G. and R. F. Woodman (1978) Spread F in equatorial ionograms associated with reversal of horizontal F region electric field Ann. Geophys. 34,31-36

Rastogi, R.G. and J.Aarons (1980) Nighttime ionospheric radio scintillations and vertical drifts at the magnetic equator J. Atmos. Terr. Phys. 42,583-591

Rino, C. L. and S.J. Matthews (1980) On the morphology of auroral zone radio wave scintillation, J. Geophys. Res. 85, 4139, 1980

Roelof, E.C. and D.J. Williams (1988) The terrestral ring current Johns Hopkins APL Technical Digest 9, 144

Suguira, M. and S. Chapman (1960) The average morphology of geomagnetic storms with sudden commencement Abh.Akad.Wiss. Goettingen Math. Phys Kl.,3

Weber, E.J., J. Buchau, J.G. Moore, J R. Sharber, R.C. Livingston, J. D. Winningham, and B. W. Reinisch, (1984) F layer ionization patches in the polar cap, J. Geophys. Res. 89,1683-1694

AD-P006 269

ATMOSPHERIC GRAVITY WAVES AND TRAVELING IONOSPHERIC DISTURBANCES: THIRTY YEARS OF RESEARCH

Robert D. Hunsucker
Geophysical Institute
University of Alaska
Fairbanks, AK 99775-0800

I. Introduction.

In November 1960 a paper, "Internal Atmospheric Gravity Waves at Ionospheric Heights" was published by Dr. Colin Hines, then working at the Theoretical Studies Group of the Canadian Defense Research Board in Ottawa. This is now one of the most-cited ionospheric publications in the latter half of the 20th century. Several hundred papers have been published describing experimental and theoretical studies of atmospheric gravity waves (AGWs) and their ionospheric manifestations - traveling ionospheric disturbances (TIDs). Earlier attempts to explain TIDs were made by Martyn (1950, 1955), Dungey (1955), Akasofu (1956), and Hines (1955 and 1956), but met with only limited success until Hines (1960) accomplished the full analytical treatment of internal atmospheric gravity waves. In simplest terms, AGWs are freely propagating waves in the neutral atmosphere, and the two dominant forces are gravity (acting downward) and buoyancy (acting upward). AGWs should not be confused with the cosmological gravitational waves as described in a book by P.C.W. Davies (1980). We shall use "atmospheric" in this paper, as first introduced by Hines (1960) instead of "acoustic" as is sometimes used in the literature on AGW phenomena. We will also restrict our discussion to AGWs propagating in the ionosphere of the "large-scale" category (periodicities of ~70 min. to several hours and horizontal speeds of ~300 m/s or higher) and of "medium-scale" (periodicities of ~15 to 70 min. and horizontal speeds of ~90 to 250 m/s). Sources of the large-scale AGWs are thought to be primarily energetic events in the auroral oval, whereas the sources of the medium-scale AGWs are numerous, and probably include ionospheric, tropospheric and earth-surface events (volcanoes, earthquakes, etc.); other sources include large man-made explosions and possibly other phenomena. We will not include in this discussion the small-scale AGWs which seem to have lower and middle atmosphere sources and have also been the subject of extensive study (Fritts, 1984).

In the first decade after publication of Hines seminal work, most of the papers described observations of TIDs which seemed to be explainable as AGW-associated phenomena (Tveten, 1961; Du Castel and Faynot, 1964; Baker and Gledhill, 1965; Liszka and Taylor 1965; Thome, 1966; Hunsucker and Tveten, 1967; Georges, 1968). It was realized that TIDs - which had been detected since the late 1940's by investigators using various radio techniques - were, indeed, manifestations of AGWs. Theoretical studies during this first decade concentrated upon describing AGW propagation characteristics (Hines and Reddy, 1967).

During the second decade "after Hines (AH)" many more observations of AGW/TID phenomena were reported, and extensive computer modeling studies were carried out, as summarized in reports and reviews by Croft, 1972; Testud, 1973; Hines, 1974; Yeh and Liu, 1974; Francis, 1975; Richmond, 1978a & b; 1979a & b; Richmond and Matsushita, 1975; Richmond and Roble, 1979; Roble et. al., 1979). A detailed analytical treatment of the generation of AGWs by the auroral electrojet and their subsequent propagation equatorward was presented by Chimonas and Hines (1970) and detailed experimental examinations of the auroral source functions were presented by Hunsucker (1977), Richmond (1978) and Brekke (1979). AGW phenomenology and theory have also been extensively described in three books published during the decade of the 1970s (Hines et. al., 1974; Gossard and Hooke, 1975 and Beer, 1974) and a conference report describing the effects of AGW/TIDs on EM wave propagation was published (AGARD, 1972).

The third decade "AH" has been characterized by continued reports of observation of AGW/TID phenomena, concentrating on worldwide campaigns utilizing some of the newer radio instrumentation now available, and a review was published on the subject of AGWs generated in the high-latitude

91-09663

ionosphere (Hunsucker, 1982). A series of "Worldwide Atmospheric Gravitywave Studies (WAGS)" (Argo and Hunsucker, 1985) was initiated with the "WAGS1" campaign 14 - 19 October 1985. Results of WAGS1 were reported in a special issue of Radio Science (Argo and Hunsucker, 1988) and in papers by Crowley and Williams, 1987); Hunsucker (1987); Williams et. al., 1988.

Subsequent WAGS campaigns have been carried out during the periods 8-15 April 1988 (WAGS2 - mainly in the European sector), and WAGS3 (28-31 August 1989). Results of these campaigns are currently being analyzed and papers are being prepared for publication (personal communications from Phil Williams and Bill Wright, 1990). Figure 1 is a schematic global overview showing the relation of high latitude ionospheric auroral AGW sources and propagation.

II. Outline of AGW/TID Theory.

Hines (1960) derived the basic mathematical formalism describing AGWs in a rather elegant manner, starting with the governing fluid dynamic equations using the "Eulerian Derivative" operator

$$\frac{D}{Dt} \equiv \left(\frac{\partial}{\partial t} + \bar{v} \cdot \nabla \right)$$

The Continuity Equation can then be written as,

$$\frac{D\rho}{Dt} + \rho \nabla \cdot \bar{V} = 0 \qquad 1$$

Similarly, the Momentum Equation

$$\rho \frac{D\bar{V}}{Dt} = -\bar{\nabla} p + \rho \bar{g} \quad \textit{(neglecting other terms)} \qquad 2$$

and the Adiabatic Form of the Perfect Gas Law

$$\frac{Dp}{Dt} + c^2 \frac{D\rho}{Dt} \qquad 3$$

and resulting in the well-known (by now) dispersion equation for an AGW in a planar, horizontally stratified, isothermal, single species, windless and non-rotating adiabatic atmosphere.

$$\omega^4 - \omega^2 C^2 \left(K_x^2 - K_z^2 \right) + (\gamma - 1) g^2 K_x^2 + i\omega^2 \gamma g K_z = 0 \qquad 4$$

where:

ω = *the AGW angular frequency* = $2\pi f$

K_x = *horizontal propagation vector*

K_z = *vertical propagation vector*

γ = *ratio of specific heats for the atmospheric gas* = $\frac{Cv}{Cp}$

Cv = *specific heat at constant volume*

Cp = *specific heat at constant pressure*

c = *local speed of sound in an ideal gas for adiabatic processes* = $\left| \gamma \frac{p}{\rho} \right|^{\frac{1}{2}}$

g = *acceleration of gravity*

Figure 2 is a diagram from Hines (1960) showing the relation between AGW wave amplitude, phase and energy propagation. Some of the properties of AGWs can be deduced by a closer examination of

equation 4. We can define two "critical frequencies", ω_a & ω_b as the acoustic cutoff frequency and the buoyancy, or Brunt-Vaisala frequency, respectively.

$$\omega_a \equiv \frac{\gamma g}{2C} \qquad 5$$

and

$$\omega_b \equiv \frac{(\gamma-1)^{\frac{1}{2}} g}{C} \qquad 6$$

For $\omega \ll \omega_b$ the waves are "generic gravity waves," and waves for which $\omega \approx \omega_b$ are called "buoyancy waves" as shown further in Figure 3 and Figure 4.

Equation 4 is a relatively simple dispersion equation, which was derived for the conditions stated and has proven to be reasonably accurate for most experimental work (Williams, 1982). Other more complex dispersion equations have been derived for more realistic atmospheres, including an eighth order, quartic equation (Hickey & Cole, 1987) - which is useful in numeric simulation of AGWs. We will use Hines equation for attempting to understand the "1st order" AGW properties (Rice, 1989).

We can then solve equation 4 for the propagation constant as,

$$K_z^2 = \frac{\omega^2 - \omega_a^2}{C^2} + K_x^2\left(\frac{\omega_b^2}{\omega^2} - 1\right) \qquad 7$$

For large-scale AGW/TIDs, we can use an asymptotic approximation in which $K_z^2 \gg \omega_a^2/C^2$ and $\omega \ll g/C$, which yields

$$K_z^2 = K_x^2\left(\frac{\omega_b^2}{\omega^2} - 1\right) \qquad 8$$

A propagating wave must have real numbers, so $\omega < \omega_b$. The angle Θ associated with the phase velocity V is given by

$$\Theta = \tan^{-1}\frac{K_z}{K_x} = \tan^{-1}\left|\frac{\omega_b^2}{\omega^2} - 1\right|^{\frac{1}{2}} \qquad 9$$

As $\omega \to 0$, $\Theta \to 90°$

The angle associated with the packet velocity is

$$\Phi = \tan^{-1}\left(-\frac{K_z}{K_x}\right) \qquad 10$$

We see that Φ and Θ differ by 90°, so the energy in an AGW propagates perpendicular to its apparent phase as shown in Figure 2. The various AGW velocities are worth some discussion. The phase velocity, V is

$$v = \frac{\omega K}{K} \qquad 11$$

Physically, phase velocity describes the motion of a point of constant phase (crest, trough, etc.) over time. Since this motion is confined within the wave packet, this phase velocity describes the apparent motion at a particular observation site. The group velocity, G, is given by

$$G = \frac{K d\omega}{dK} \qquad 12$$

This motion is perpendicular to the group plane, and indicates the motion of a particular time-limited portion of the packet. The "horizontal trace speed" is defined as

$$G_h = \left| \left(\frac{dK_x}{d\omega} \right)^2 + \left(\frac{dK_y}{d\omega} \right)^2 \right|^{-\frac{1}{2}} \qquad 13$$

This speed will be observed as the speed of the TID by a single observer who cannot view the entire AGW packet. Since large-scale TID packets typically extend for 1000's of kilometers, G_h is the generally reported TID speed.

The packet velocity, P is defined in terms of its components, (Hines, 1960)

$$P_x = \frac{\partial \omega}{\partial K_x} \quad and \quad P_z = \frac{\partial \omega}{\partial K_z} \qquad 14$$

For instance, the packet velocity in the x-direction is,

$$P_x = \frac{\omega K_x (\omega^2 - \omega_b^2)}{(\omega^4/C^2 - K_x^2 \omega_b^2)} \qquad 15$$

As Hines pointed out, an experimental propagation speed determined from observations is rather poorly defined because it is based on the times of arrival of different components of the packet--Hines labels this experimentally-determined speed the "disturbance trace speed". This disturbance trace speed will usually be close to the horizontal group trace speed for a given mode. If observed over very large distances, the disturbance trace speed becomes a rough approximation of the packet speed, P_x.

Another topic that merits some comment is the relationship between AGWs and TIDs. The basic linkage between the two phenomena is "collisional coupling" between neutrals and ions--but the actual process is somewhat more complex as has been shown by Hooke (1968) and Testud (1972), as a result of some complicated chemistry changes described by Hooke (1968). If the object of the investigation is to accurately portray AGW characteristics from measured characteristics of a TID, one can relate the TID electron density parameters (Ne) to the AGW parameters in the neutral atmosphere [U(Zo)] by the equation,

$$N_e(x,y,z,t) = N_{eo}(z)\, U_b(z_o) \sin I \frac{\exp[K_{zi}(z - z_o)]}{\omega} \left| \phi_x^2 + \phi_z^2 \right|^{\frac{1}{2}}$$

$$\exp\left\{ i \left| \omega t - k_x x + k_y y - k_z z + \frac{\pi}{2} - \tan^{-1} \frac{\phi x}{\phi z} \right| \right\} \qquad 16$$

Where:

$$\phi x = \frac{k_{br}}{\sin I}$$

$$\phi_z = \frac{1}{N_{eo}} \frac{\partial N_{eo}}{\partial z} + K_{zi}$$

$U_b(z_o)$ = *amplitude of the AGW projected along the field line*

Z_o = *arbitrary reference height*

N_{eo} = *unperturbed electron density*

$k_z = k_{zr} + i k_{zi}$

$k_{br} = Re\{k \quad B\}$

Testud (1972) showed that in addition to the electron density fluctuation, there are other effects complicating the AGW/TID coupling mechanism His model predicts large electron temperature (Te) variations (as observed by incoherent scatter radars), which have been reported [see Hunsucker, 1982-Figures 20 through 24].

If one is only concerned with the effects of TIDs on radio propagation systems, the somewhat complex relationship between AGWs and TIDs is of lesser importance than to an ionospheric scientist studying the physics of AGWs.

An idealized representation of the rather complex relationship between auroral ionospheric sources, AGWs and TIDs is shown in Figure 5. A Joule heating source is assumed, resulting in a modified E-region electron density profile, the launching of an AGW, increasing in "amplitude" as it reaches the F-region, where the resulting TID propagates equatorward, and is observed by an ionosonde.

III. Experimental techniques being utilized to investigate TIDs.

Several different radio techniques have been and are being utilized to observe, detect and characterize TIDs. Currently they include: "analog and digital" ionosondes; incoherent scatter radars (ISRs); HF backscatter sounders; HF Doppler sounders; satellite radio beacons; riometers and even astronomical radio telescopes. Additional valuable information on AGW sources has been obtained by imaging satellites (DE, DMSP, etc.) and by other multipurpose satellites (HILAT, Polar Bear, etc.).

Basically, radio techniques are capable of measuring the frequency, amplitude, time delay, phase, and spectrum of the received signal, then derive other quantities from them. Each specific technique is sensitive to its particular physical interaction theory and sometimes to particular ionospheric models. It is important then, to keep in mind that no one particular technique "tells the complete truth" about AGW/TID phenomena! Thus, the most complete description of AGW/TID sources, propagation and interaction with EM signals (or for that matter, any physical phenomena) is given by simultaneous observation by several independent techniques (see Hunsucker, 1990).

IV. Experimental Results in the 1980's.

Experimental investigation of AGW/TID phenomena from approximately mid-1970's until 1981 have been reported in a review paper by Hunsucker (1982). There was a relative scarcity of published reports on AGW/TID phenomena from 1981 until the WAGS1 campaign (14-19 October 1985) - which seemed to stir up considerable interest and resulted in about 10 papers published in the scientific literature from 1987 - 1989. An interesting series of studies was initiated independent of WAGS activities by Hajkowicz (1983 a,b) using latitudinal "chains" of riometers and ionosondes to study AGW/TID phenomena with sources in the auroral ionosphere. It was found that rapid onsets of riometer absorption were associated with substorm-related energetic particle precipitation - resulting in the launching of large-scale TIDs equatorward from conjugate sources (Hajkowicz and Hunsucker, 1985 and 1987. Figure 6 shows the velocities and latitudinal positions of large-scale TIDs observed during this study. TIDs observed during this study originated between L = 4 - 5 and propagated equatorward for distances in excess of 7000 km and it was postulated that since they might constructively interfere in the equatorial regions as suggested by Gross et. al (1984).

Further evidence that the region around L = 4 is also a source of medium-scale TIDs was presented by Hunsucker and Hargreaves (1986 and 1988). It was found that these radio beacon studies of the variation of total electron content (TEC) showed that these medium-scale TIDs were present 98% of the daytime hours and no unambiguous sources were identified. The results of another "non-WAGS" investigation made in 1974 using Faraday rotation measurements of the INTELSAT IIF3 beacon and they conclude that a delineation between medium-scale and large-scale TIDs occurs for a midlatitude observing station for wave periods of 50-70 min. and a speed of 200-250 m/s (Kersley and Hughes, 1989). This is in general agreement with the theoretical values obtained by Richmond (1978a & b). Kersley and Hughes conclude that only for periods greater than about 70 min. and speeds in excess of 270 m/s is auroral generation of AGW/TIDs possible over horizontal distances greater than 1500 km under the conditions assumed in theory. Sampson et. al. (1989) have used the high-resolution HF radar located at Goose Bay, Labrador to study gravity waves and have concluded that TIDs observed by this technique originate near the ionospheric convection reversal and propagate equatorward as earth reflected waves.

The concept of a Worldwide Atmospheric Gravitywave Study (WAGS) was discussed during the fall American Geophysical Union (AGU) meeting in San Francisco in 1984. The first campaign was planned for October 15-18, 1985 and *co-facilitators* Paul Argo and Robert Hunsucker were appointed to coordinate the WAGS 1 campaign. An announcement was published in EOS (Argo and Hunsucker, 1985) and announcements were mailed to ~60 investigators who had expressed interest.

Some 20 different groups from ~12 countries participated in WAGS 1 (15-18 October 1985) and results have been published by Williams et al. (1988), Crowley et al. (1988), Hunsucker (1987), Rice et al. (1988), Sheen and Liu (1988a&b), Crowley et al. (1988), Walker et al. (1988), Kersley and Hughes (1989). A Data

Workshop was held in Boulder, Colorado, in January 1987 to facilitate analysis of WAGS 1 data. Early WAGS 1 results and other AGW/TID reports were presented in volume 1 of the WITS Handbook (1988).

Selected results obtained from the WAGS 1 campaign are:

1. Propagating TID/AGWs were observed over large areas of East Asia using a network of 10 ionosondes operating at 5 minute intervals. The TIDs were observed at nighttime and most of them propagated equatorward with periods form 40-210 minutes. There was no convincing evidence of increasing period with distance traveled equatorward by the TID. Sources of the AGWs were not defined (Walker et al., 1988).

2. A theoretical model of ion velocity and electron density variations caused by TID/AGWs was developed and compared with TID measurements made by the Sondrestrøm and Millstone Hill ISRs. Very good agreement between model and experimental data was reported in 3 cases (Sheen et al., 1988 - Paper 1).

3. The *background wave spectra* observed during several magnetically quiet days at the Sondrestrøm ISR were compared with a theoretical model and it was concluded that, indeed, these wave spectra confirm that gravity waves produce these variations in ion velocity and electron density (Sheen et al., 1988 - Paper 2).

4. Data from a tristatic HF Doppler sounder system in the UK revealed both large-scale and medium-scale TIDs during the WAGS period. During the highest magnetic activity, the TIDs were consistent with large-scale aurorally-generated AGWs. Large-scale TIDs of low amplitude existed even during quiet magnetic conditions. There was no observed *wave filtering* by neutral winds except for the quietest day (Crowley et al., 1988).

5. A rather convincing *cause-effect* observation of an AGW generated by a Joule heating event in the ionosphere north of central USSR was documented. This heating event generated 2 AGWs which produced 2 TIDs observed at Sondrestrøm, Leicester, Millstone Hill, Wallops Island and possibly at Arecibo. In this case, a large, impulsive joule heating event produced 2 TIDs which were *tracked* from the auroral oval source region equatorward to a distance of 9000 km. The propagating TID appeared to increase in period as it propagated southward (Rice et al., 1988).

6. Other investigators used the EISCAT ISR, the HF Doppler system in the UK, and chains of riometers and magnetometers in another *cause-effect* study of AGW/TID phenomena during WAGS 1. They reported that auroral joule heating events were stronger than particle heating and that the Lorentz forcing term is always dominant below 100 km in generating AGW. They also found evidence for a periodic source function producing TIDs of the same period, and concluded that longer wave length TIDs propagated further equatorward than shorter wave length TIDs (Williams et al., 1988, JATP and Crowley and Williams, 1987, Nature).

7. Hunsucker (1987, Nature) reviewed the early results of the WAGS 1 campaign and emphasized that there seems to be at least 2 distinct types of source function for generation of AGW's in the auroral ionosphere; single events and multiple peaked events. In the latter case, the periodicity of the *source function* determines both the magnitude and periodicity of the TIDs.

V. Effects of TIDs on radio systems.

Early reports of amplitude fading on the order of tens of minutes of HF signals can, at least qualitatively, be explained by TIDs passing through the ionospheric reflection points on the great circle propagation paths. In general, TIDs do not produce profound effects on the typical HF ionospheric propagation path probably because the rather wide antenna beams and low "signaling rates" utilized. It was only in the late 1960's when high resolution HF backscatter sounders (see Croft, 1972) became capable of detecting the "fine-structure" of the ionosphere that we became aware of the extent of the problem. Also during the 1960's early OTH, accurate HF/DF and higher signaling rate HF systems started to "come on line" and anomalies with "AGW periodicities" started to show up. Figure 7 shows typical variations of plasma frequency-virtual height due to passage of TIDs - which could produce effects in "high-resolution systems". [See also the investigations of TID effects on HF bearings reported by Morgan (1972)].

During 1967 and 1968 a coordinated series of high-resolution backscatter soundings using the ESSA HF backscatter sounder, 5 ionosondes and two transponders was carried out. The backscatter sounder (Tveten & Hunsucker, 1969) was located in Boulder, Colorado, the transponders were located in Meridian, Mississippi and Orlando, Florida and the boresite azimuth of the backscatter sounder was 114°. Figure 8 is an example of DOA signatures in azimuth and elevation during a series of TIDs, showing the close association of the two signatures (Tveten and Hunsucker, 1968).

There is some evidence that under certain conditions, TIDs can "trigger" spread-F (Whitehead, 1971; Kelley et. al., 1981; Yeh, 1984; Hunsucker and Hargreaves, 1986) which could also produce some degradation of performance of a "high-resolution system" (see Figure 7).

VI. Future Directions.

It is clear that much effort has been expended and quite a bit has been learned about the morphology and physics of AGW/TID phenomena in the last three decades. Some remaining problems in this field include:

A. Obtaining better occurrence statistics on TIDs as a function of time of day, season, sunspot cycle and solar-terrestrial disturbance level to enable forecasters and predictors to do a better job.

B. Finding the best solar-geophysical parameters (Kp, Kn, auroral oval status, solar flare occurrence, etc.) for predicting the occurrence of large-scale TIDs.

C. Documenting and quantifying specific effects on "high-resolution" systems (OTH radars, Direction-finding systems, etc.)

D. Obtaining more knowledge of the physics of AGW sources (for large and medium-scale TIDs), i.e. - Joule and particle heating, Lorentz forcing, mechanisms for single and "periodic" sources.

E. Global propagation of TIDs (do conjugate simultaneous AGWs produce TIDs which sometimes form "standing-wave" structures in the equatorial ionosphere?)

F. Why are medium-scale TIDs ubiquitous at certain regions?

A large-scale global campaign called "SuperWAGS" (Liu and Hunsucker, 1990) is scheduled for 10-24 November 1990 in a continuing effort to find answers to such questions as those posed above. Figure 9 is a "time-line" for the various WAGS worldwide campaigns We hope that the global network of experimenters will include the community of users of high-resolution radio systems. Under the auspices of the STEP program it is hoped that future "WAGS-type" campaigns will be scheduled to continue investigations of AGW/TID phenomena.

VII. Acknowledgements.

National Science Foundation Grant ATM 8720541 supported the preparation of this paper and Ms. Sheila Finch and Carol Gering prepared the camera-ready-copy.

VIII. References.

AGARD, (1972). Effects of atmospheric acoustic gravity waves on electromagnetic wave propagation. Conf Proc No. 115

Akasofu, S.-I., (1956). Dispersion relation of magneto-hydrodynamic waves in the ionosphere and its application to the shock wave, Sci Repts Tohoku Univ Geophys 8, 24.

Argo, P.E., and R.D. Hunsucker, (1985). Gravity wave study EOS Trans AGU, 66, 121.

Argo, P.E., and R.D, Hunsucker, (1988). Foreword: The Worldwide Atmospheric Gravity wave Study (WAGS). Radio Sci 23, 865.

Baker, D.C., and J.A. Gledhill, (1965). An unusual traveling disturbance in the F-region of the ionosphere. J. Atmos Terr. Phys , 27, 1223.

Banks, P.M., and J.V. Evans, (1979) An initial feasibility study to establish a very high latitude incoherent scatter radar. NSF Grant ATM-7826533 to Utah State Univ

Beer, T , (1974). Atmospheric waves. John Wiley, New York, N.Y.

Brekke, A., (1979). On the relative importance of Joule heating and Lorentz force in generating atmospheric gravity waves and infrasound in the auroral electrojets. J. Atmos Terr. Phys., 41, 475.

Chimonas, G., and C.O. Hines, (1970). Atmospheric gravity waves launched by auroral currents. Planet. Space. Sci. 18, 565.

Croft, T.A., (1972). Sky wave backscatter: A means for observing our environment at great distance. Rev. Geophys. Space Phys., 10, 73.

Crowley, G., and P.J.S. Williams, (1987). Observations of source and propagation of atmospheric gravity waves. Nature 328, 231.

Crowley, G. and I.W. McCrea, (1988). A synoptic study of TIDs observed in the United Kingdom during the first WAGS campaign. Radio Sci. 23, 905.

Davies, P.C.W., (1980). The Search for Gravity Waves. Cambridge Univ. Press, Cambridge, UK.

DuCastel, F., and J.M. Faynot, (1964). Some irregularities observed simultaneously in the upper and lower ionosphere at middle latitudes. Nature 204, 984.

Dungey, J.W., (1955). Electrodynamics of the outer atmosphere. The physics of the ionosphere (The Physical Society, London), 229-236.

Francis, S.H., (1975) Global propagation of atmospheric gravity waves: A review. J. Atmos. Terr. Phys., 37, 1011.

Fritts, D., (1984). Gravity wave saturation in the middle atmosphere: A review of theory and observation. Rev. Geophys. Space. Phys., 22, 275.

Georges, T.M., (1968). Ionospheric effects of atmospheric waves. ESSA Tech Rept. 57-ITSA 54, U.S. Gov. Printing Office; Washington, D.C.

Gossard, E.E., and W.H. Hooke, (1975). Waves in the atmosphere. Elsevier, New York.

Gross, S.H., C.A. Reber and F Huang, (1984). Large-scale waves in the thermosphere observed by the AE-C satellite. IEEE Trans on Geoscience. Remote Sensing GE-22, 340-352.

Hajkowicz, L., (1983a). Auroral riometer absorptions and the the F-region disturbances observed over a wide range of latitudes. J. Atmos. Terr. Phys., 45, 175.

Hajkowicz, L., (1983b). Conjugate effects in the generation of traveling ionospheric disturbances (TIDs) in the F-region. Planet. Space. Sci. 31, 1409.

Hajkowicz, L., and R.D. Hunsucker, (1985). The conjugate source location of large-scale TIDs. Ionospheric Research Section Rept. No. 12, Dept. of Physics, University of Queensland, Brisbane, Australia.

Hajkowicz, L., and R.D. Hunsucker, (1987). A simultaneous observation of large-scale periodic TIDs in both hemispheres following an onset of auroral disturbances. Planet. Space. Sci., 35, 785.

Hickey, M.P. and K.D. Cole, (1987). A quartic dispersion equation for internal gravity waves in the thermosphere. J. Atmos Terr. Phys., 49, 889.

Hines, C.O., (1955). Hydromagnetic resonance in ionospheric waves. J. Atmos Terr. Phys., 7, 14-30.

Hines, C O., (1956). Electron resonance in ionospheric waves. J. Atmos. Terr. Phys., 9, 56-70.

Hines, C.O., (1960). Internal atmospheric gravity waves at ionospheric heights Can. J. Phys., 38, 1441.

Hines, C.O., and C.A. Reddy, (1967). On the propagation of atmospheric gravity waves through regions of wind shear. J. Geophys. Res., 72, 1015

Hines, C.O., et. al. (1974). The upper atmosphere in motion. Geophys Monogr. Ser , Vol. 18, American Geophysical Union, Washington, D.C.

Hooke, W.H., (1968). Ionospheric irregularities produced by internal atmospheric gravity waves. J Atmos. Terr. Phys., 30, 795.

Hunsucker, R.D. and L.H. Tveten, (1967). Large traveling-ionospheric disturbances observed at midlatitudes utilizing the high resolution HF backscatter technique. Proc. IEEE 57, 487

Hunsucker, R.D., (1977). Estimate of the relative importance of Joule heating and the Lorentz force in generating atmospheric gravity waves from the auroral electrojet. J. Geophys. Res., 82, 4826.

Hunsucker, R.D., (1982). Atmospheric gravity waves generated in the high-latitude ionosphere: A review. Rev. Geophys. Space Phys., 20, 293-315.

Hunsucker, R.D., and J.K. Hargreaves, (1986). A study of gravity waves in ionospheric electron content at L = 4. Proceedings of the International Beacon Satellite Symposium June 9-14, 1986; Oulu, Finland, P. 243.

Hunsucker, R.D., (1987). The sources of gravity waves. Nature, 328 (6127), 204-205.

Hunsucker, R.D., and J.K. Hargreaves, (1988). A study of gravity waves in ionospheric electron content at L = 4 J. Atmos. Terr. Phys., 50, 167-172.

Hunsucker, R. D., (1990). Radio techniques for probing the terrestrial ionosphere. Springer-Verlag: Heidelberg, Berlin, New York.

Kelley, M.C., M.F. Larsen, C. LaHoz and J.P. McClure, (1981). Gravity wave initiation of equatorial spread-F: A case study. J. Geophys. Res., 86, 9087.

Kersley, L., and A. Hughes, (1989). On the distinction between large-scale and medium-scale atmospheric gravity waves. Annales Geophysicae, 7 (5), 459-462.

Liszka L., and G.M. Taylor, (1965). A synoptic study of large-scale ionospheric irregularities using observations of the Faraday rotation of satellite signals. J. Atmos. Terr. Phys., 27, 843.

Liu, C.-H., and R.D. Hunsucker, (1990). The SuperWAGS Campaign, EOS Trans AGU, April 10, p.2.

Martyn, D.F., (1950). Tidal phenomena in the ionosphere. Proc. Roy. Soc. A., 201, 216.

Martyn D.F., (1955). Interpretation of observed F2 "winds" as ionization drifts associated with magnetic variations. The physics of the ionosphere. (The Physical Society,London),163-165.

Morgan, M.G., (1972). The effect of ionospheric disturbances on the bearing of incoming sky waves. AGARD Conf. Proc. No. 115; 28-1/28-17.

Rice, D.D., (1989). Observations of gravity wave generation and global propagation in the ionosphere. M.S. Thesis, University of Alaska-Fairbanks.

Rice, D.D., R.D. Hunsucker, L.J. Lanzerotti, G.Crowley, P.J.S Williams, J.D. Craven and L. Frank, (1988). An observation of atmospheric gravity wave cause and effect during the October 1985 WAGS campaign. Radio Sci., 23, 919.

Richmond, A.D., (1978a). The nature of gravity wave ducting in the thermosphere. J. Geophys. Res., 83, 1385.

Richmond, A.D., (1978b). Gravity wave generation, propagation and dissipation in the thermosphere. J. Geophys. Res., 83, 4131.

Richmond, A.D., (1979a). Large-amplitude gravity wave energy, production and dissipation in the thermosphere. J. Geophys. Res., 84, 1880.

Richmond, A.D., (1979b). Thermospheric heating in a magnetic storm: Dynamic transport of energy from high to low latitudes. J. Geophys. Res., 84, 5259.

Richmond, A.D. and S. Matsushita, (1975). Thermospheric response to a magnetic substorm. J. Geophys. Res., 80, 2839.

Richmond, A.D., and R.G. Roble, (1979). Dynamic effects of aurora-generated gravity waves on the mid-latitude ionosphere. J. Atmos. Terr. Phys., 41, 841.

Roble, R.G., R.E. Dickinson and E.C. Ridley, (1979). Thermospheric response to the November 8-9 1969 magnetic disturbances. J. Geophys Res., 84, 4207.

Sampson, J.C., R.A. Greenwald, J.M. Ruohoniemi and K.B. Baker, (1989). High-frequency radar observations of atmospheric gravity waves in the nihg-latitude ionosphere. Geophys. Res. Lett. 16, 875.

Sheen, D.R. and C.-H. Liu, (1988a). Modeling F-region gravity waves observed during the WAGS 1 special event. Radio Sci. 23, 879.

Sheen, D.R. and C.-H. Liu, (1988). Modeling of F-region gravity waves during the WAGS campaign 2. Background wave spectra. Radio Sci. 23, 894.

Testud, J., (1972). Interaction between gravity waves and ionization in the ionospheric F region. COSPAR Space Res XII, Vol. 2, edited by S.A. Bowhill et. al. pp. 1163-1178. Akademie-Verlag, Berlin.

Testud, J., (1973). Ondes atmospheriques de grande echelle et sous-orages magnetiques. These de doctorat d'etat. Univ of Paris.

Thome, G., (1968). Long-period waves generated in the polar ionosphere during the onset of magnetic storms. J. Geophys. Res. 73, 6317-6336.

Tveten, L.H., (1961). Ionospheric motions observed with high frequency backscatter sounder. J. Res. NBS., Nat'l Bur. of Standards, Sect. D65, 115.

Tveten, L.H., and R.D. Hunsucker, (1968). Some observations on backscatter and direction of arrival relationships. 1968 ARPA/OHD Proceedings - Monterrey, CA. October 1968.

Tveten, L.H., and R.D. Hunsucker, (1969). Remote sensing of the terrestrial environment with an HF radio high-resolution azimuth and elevation scan system. Proc. IEEE 57, 487-492.

Walker, G.O., Y.W. Wong, T.H.K. Ma, T. Kikuchi, K. Nozaki, Y.N. Huang and V. Badillo, (1988). Propagating ionospheric waves observed throughout east ASia during the WAGS October 1985 campaign. Radio Sci. 23, 867-878.

Whitehead, J.D., (1971). Ionization disturbances caused by gravity waves in the presence of an electrostatic field and background wind. J. Geophys. Res., 76, 238.

Williams, P.J.S., A.P. van Eyken and F. Bertin, (1982). A test of the Hines dispersion equation for atmospheric gravity waves. J. Atmos. Terr. Phys. 44, 573.

Williams, P.J.S., G. Crowley, K. Schlegel, T.S. Virdi, I. McCrea, G. Watkins, N. Wade, J.K. Hargreaves, T. Lachlan-Cope, H. Muller, J.E. Baldwin, P. Warner,A.P. van Eyken, M.A. Hapgood and A.S. Rodger, (1988). The generation and propagation of atmospheric gravity waves observed during the Worldwide Atmospheric-gravity wave study (WAGS). J. Atmos. Terr. Phys., 50, 323.

WITS Handbook No. 1 (Edited by C.-H. Liu and Belva Edwards, SCOSTEP Secretariat, Univ. of Illinois, Urbana, IL)

Yeh, K.-C., and C.-H. Liu, (1974). Acoustic-gravity waves in the upper atmosphere. Rev. Geophys. Space Phys., 12 (2),193.

Yeh, K.-C., (1984). Ionopsheric irregularities and gravity waves. Proc. of International Beacon Satellite Symposium; New Delhi, India pp. 479-487.

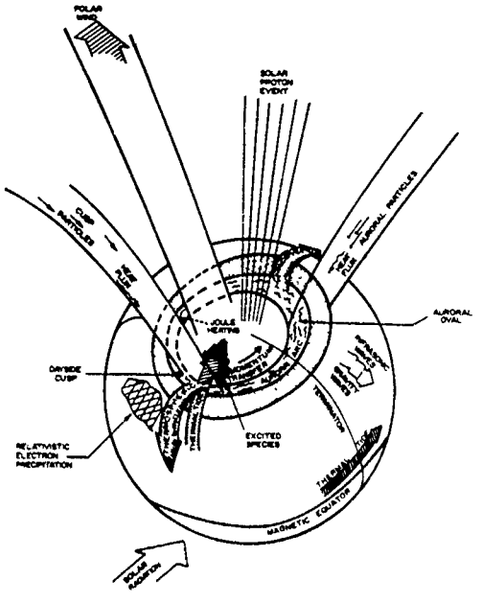


Figure 1. Relation between AGW auroral ionospheric sources and propagation on a global scale (after Banks & Evans, 1979)

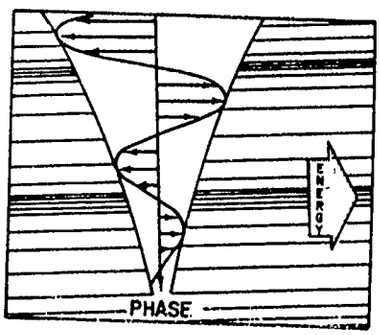


Figure 2. Hines original schematic diagram of relation of the propagation velocities of an AGW (Hines, 1960).

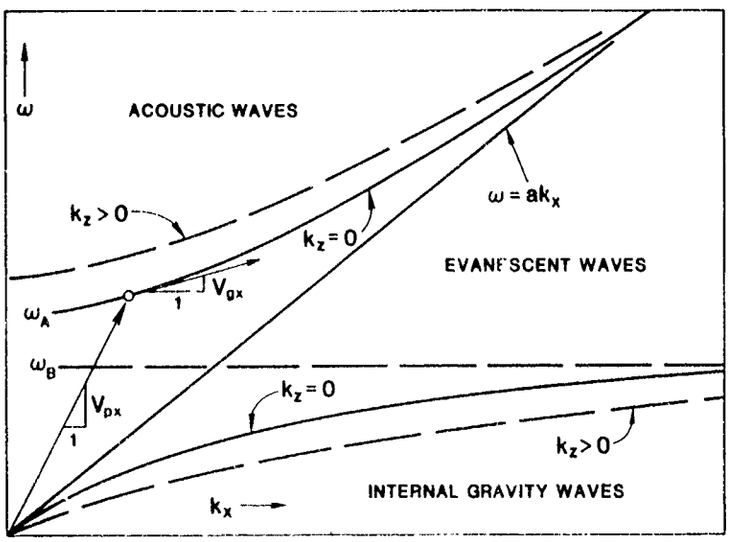


Figure 3. "Diagnostic Diagram" for AGWs acoustic waves & evanescent waves

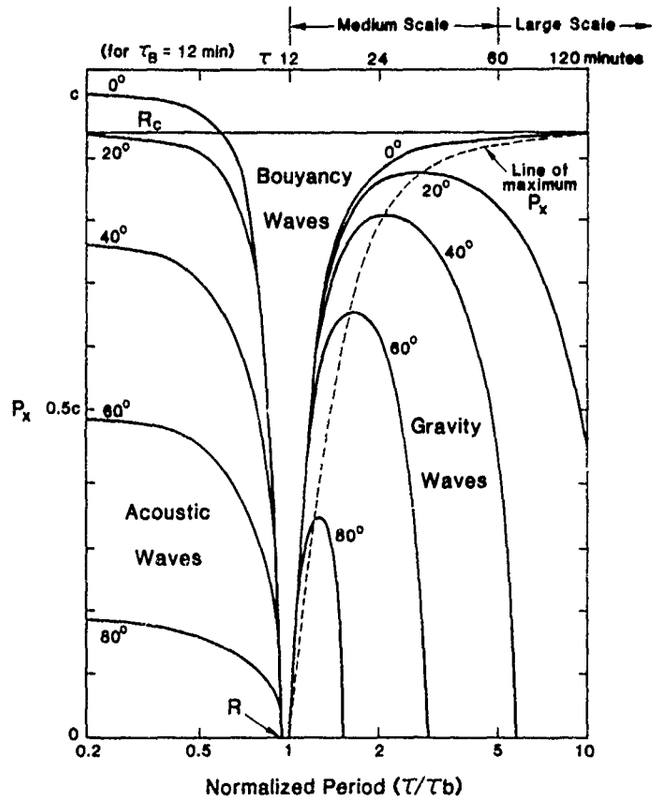


Figure 4. Acoustic, AGW and buoyancy wave spectra.

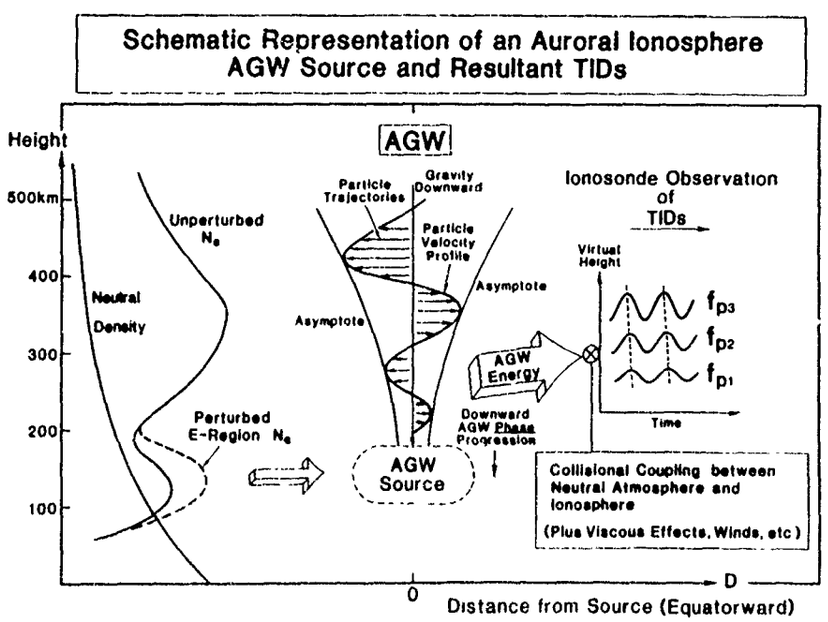


Figure 5. Schematic diagram of Auroral Ionospheric Source, AGW and TID relationship.

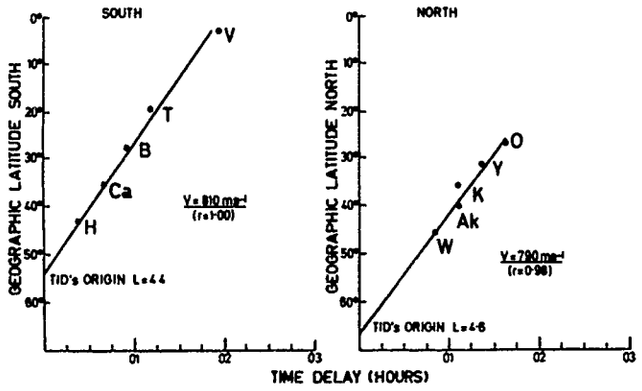


Figure 6. Velocities and latitudinal position of large-scale TIDs on 9 November 1979 (after Hajkowicz and Hunsucker, 1987).

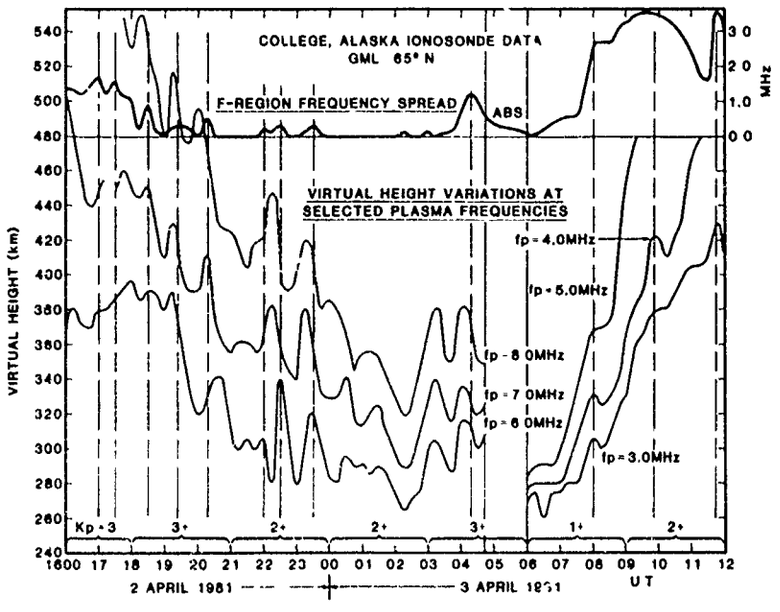


Figure 7. Virtual height and "frequency-spread-F" variations vs time at selected plasma frequencies at College, Alaska. Frequency spread of the spread-F echo is shown in the upper plot and virtual height variations are shown in the bottom plot for a disturbed period in 1981. (from Hunsucker and Hargreaves, 1988).

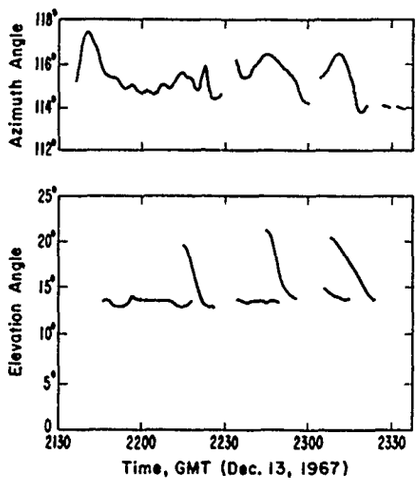
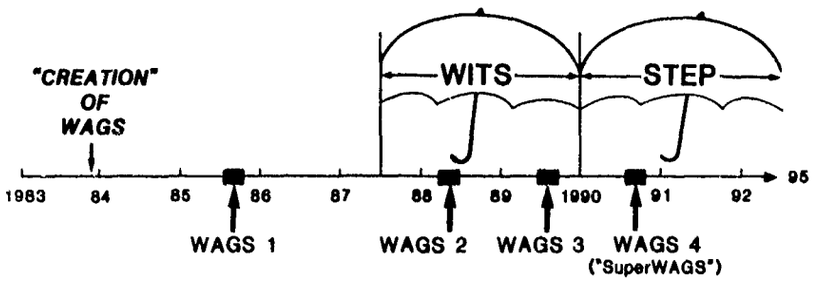


Figure 8. Simultaneous observation of TIDs (from HF backscatter sounder and DOA variation (after Tveten and Hunsucker, 1968).

WAGS CAMPAIGN "TIME-LINE"



(Plus regular world days specified on IUWDS Calendar

also 5 special intervals in 1988, and 2 in 1989)

AD-P006 270

THE TROUGH IN THE DAYTIME F LAYER: A MACROSCOPIC EFFECT OF IONOSPHERIC-MAGNETOSPHERIC CONVECTION

J. A. Whalen
Ionospheric Effects Branch (LIS)
Geophysics Laboratory, Hanscom AFB, MA 01731

ABSTRACT: The daytime F layer trough is a major result of ionospheric-magnetospheric convection, appearing in the winter high latitude ionosphere as a continuous band thousands of kilometers in extent in which the daytime F layer electron density is depleted, often by an order of magnitude. As observed by a global array of ionospheric sounders during solar maximum, the trough occurs in regions of sunward convection, in the morning corresponding to the dawn cell and in the afternoon corresponding to the dusk cell. The formation of the trough is consistent with the transport of low density nighttime plasma into the day sector where it displaces high density daytime plasma, although other mechanisms such as ion chemical effects may also play a role. The trough is observed to be present on every day, contracting to latitudes above ~70° MLAT during quiet conditions but expanding so as to extend from polar cap to midlatitudes during disturbed conditions. As activity increases the afternoon trough occurs at an increasingly earlier MLT as observed at a fixed MLAT, and at an increasingly lower MLAT as observed at a fixed MLT. It is typically stationary in MLAT/MLT coordinates for durations of many hours. Since the convection pattern resides in the magnetospheric frame of reference, the daytime trough, which is its ionospheric result, resides in solar geomaghetic coordinates (Mlat/MLT). On the other hand solar ion production and the undisturbed daytime F layer in general reside in solar terrestrial coordinates (geographic latitude/LT). The 2 coordinate systems are skewed with respect to one another principally because of the offset of the geographic and geomagnetic poles. Since the relation between the coordinates changes with location on the earth, the relation between the trough and the daytime F layer changes with location on the earth. In particular in the northern hemisphere, MLT leads LT in eastern MLONG but lags LT in western MLONG so that the trough pattern occurs at an earlier LT in the east than in the west. The trough is so prominent that its effect is to truncate the monthly median LT distribution of foF2 and displace the midday maximum relative to the undisturbed distribution to earlier LT in the east and to later LT in the west, effects which increase with increasing MLAT. In the southern hemisphere the longitudinal relation between the local time coordinates is reversed from that of the north, and as a result, the sense of the LT displacement with longitude of the midday F layer maxima is reversed. In addition because the offset of the poles is greater in the south, the magnitude of the LT displacements of the midday maxima are observed to be greater.

1. INTRODUCTION

The existence of the trough in the winter daytime F layer at high latitudes has been known for many years (Muldrew, 1965), but its importance as a permanent, macroscopic feature has only recently been recognized (Whalen 1987, 1989a, 1989b; Sojka et al 1990). This trough is so prominent that it can be said to control the daytime F layer, increasingly with increasing MLAT and with increasing activity. Its relation to the daytime F layer, the auroral oval and its existence as a separate phenomenon from the nighttime trough can be shown in the ionospheric soundings from a chain of 9 stations situated principally in Scandanavia between 50° and 66.3° MLAT. foF2, the plasma frequency at the F layer maximum, is recorded at each station at hourly intervals continuously during the Dec 1958 IGY period. A gridded represention of the measurements in a simulated 3 dimensional orthographic presentation, compares a quiet day on which the daytime trough is absent (Figure 1a) with moderate day when it is present (Figure 1b). The vertical axis is electron density at the F layer maximum, NmF2 (in electrons/cm^3), the lower and upper limits of which correspond approximately to 2 and 20 MHz respectively. This parameter is plotted as a function of magnetic latitude (MLAT) in the corrected geomagnetic system and local (solar) time

91-09662

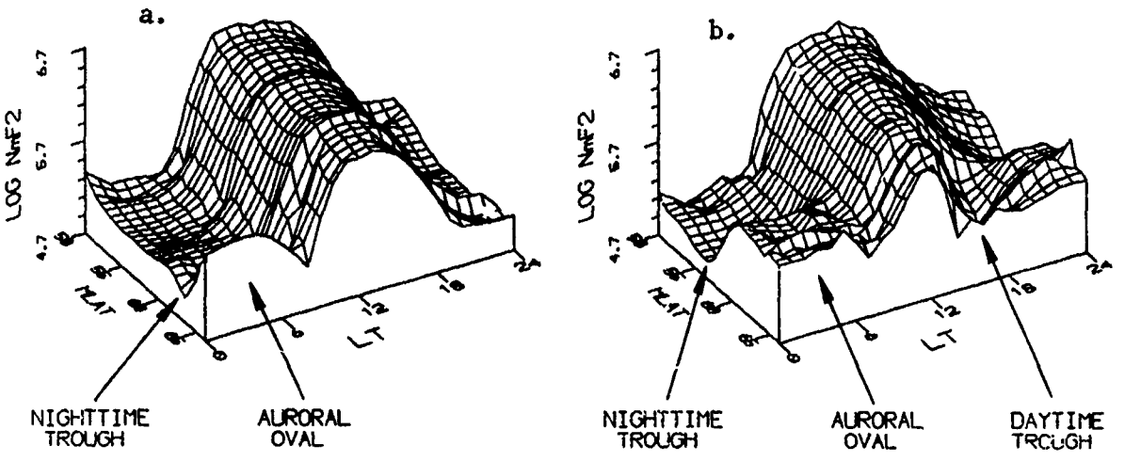


Fig. 1. The daytime trough in relation to the other major F layer phenomena. A 3-dimensional plot of electron density at the F layer maximum, NmF2 (electrons/cm^3), versus MLAT (deg) and LT (h) measured by a meridional chain of 9 ionospheric sounder stations. The undisturbed daytime F layer, the auroral oval and the nighttime trough are indicated for two 24 hour periods: (a) A quiet day (average Kp of 1o) when the daytime trough is not observed. (b) A moderate day (average Kp of 2o) when the daytime trough is observed in the afternoon as a depletion in NmF2 of a factor of 8 as compared to the undisturbed level of Figure 1a.

(LT).

On the quiet day (10 Dec., average Kp of 1o) in Figure 1a, the nighttime trough is seen on the 00 LT meridian as a region of low electron density which decreases with increasing latitude until about 63° MLAT where it undergoes a rapid increase at the poleward trough wall. This feature is the equatorward boundary of what is generally termed the auroral oval, produced in large part by ionization by particle precipitation. However it is not necessarily coincident in location with the instaneous position of the visible auroral oval (Rodger et al 1986), which is an E region phenomenon. In addition this region in the nighttime is the locus of convection which can relocate the particle produced plasma and also transport higher density plasma across the polar cap from the dayside. However the southern hemisphere data suggests that particle production may be the primary source of this plasma as will be discussed below.

The nighttime trough remains at a nearly constant level until about 0600 LT when the increase resulting from sunrise is seen. The oval apparently retreats poleward during this interval.

The daytime F layer is very regular and is taken to be the baseline undisturbed case. Electron density decays following sunset returning by 2200 LT to the level of the nighttime trough seen earlier. The auroral oval is not observed in the evening to midnight sector having apparently retreated poleward of the chain. The daytime trough is not visible by the chain on this day but is observed at other locations and to be located poleward of the chain.

On the second day (21Dec, average Kp of 2o) in Figure 1b, at 00LT the oval is considerably equatorward of its location in the previous case, so that the nighttime poleward trough wall begins at about 57°MLAT. As before the nighttime trough remains at a low level until sunrise.

The daytime trough appears near noon at the highest latitude as a rapid decrease, NmF2 falling by approximately a factor of 8 within 2 hours. The trough moves to lower latitude with increasing time and is followed by an increase produced by the auroral oval.

The 2 troughs are seen in Figure 1 to be different phenomena: the daytime trough existing at midday near maximum solar ion production but the nighttime trough existing only in the absence of ion production, whether solar or particle. The daytime trough is the result of dynamic ion removal processes related to convection. The nighttime trough on the other hand is primarily the result of the decay of the daytime ionosphere bounded in space and time by ion production or transport.

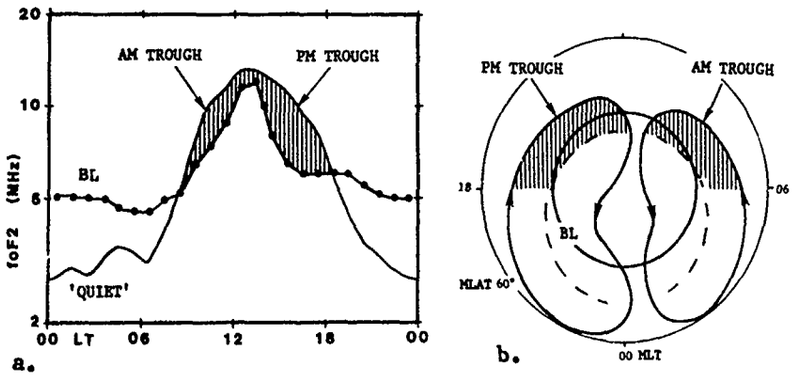


Fig. 2. Relation of morning and afternoon troughs to convection. (a) Median diurnal distribution of foF2 from Baker Lake (BL) showing morning and afternoon trough components as depletions in comparison to an inferred undisturbed 'quiet' distribution. (b) Schematic 2-cell convection pattern showing the morning and afternoon troughs observed in Figure 2a to occur during the passage of BL through regions of the dawn and dusk cells in which high density daytime plasma can be displaced by low density nighttime plasma by convection.

The daytime trough is observed to exist by the world-wide chain of sounders on every day of this period and to extend at least to 75° MLAT where it produces similarly catastrophic depletions of the daytime F layer even under moderate conditions such as seen here.

The fact that the daytime trough moves equatorward with time means that its onset in LT corresponds to its equatorward boundary. As a result any station in the proper latitude range will detect this boundary.

This work uses the measurements from the greatest number of ionospheric sounders, the world-wide network in place in 1958 during the International Geophysical Year, and reviews some of the results reported earlier on the relation to convection, the morphology and the activity dependence of the daytime trough (Whalen 1989b). It expands the study of the longitude dependence in the northern hemisphere to studies in the southern hemisphere and describes general hemispheric properties, north-south and east-west. Finally it attempts sort out those regions of plasma produced by solar radiation from those produced by particle precipitation, a task complicated by the dominance of convective displacement of plasma in the high latitude F layer.

2. CONVECTION-TROUGH RELATIONS

The relationship of the daytime trough to convection has been observed by means of simultaneous measurement of electron density and plasma drift by incoherent scatter radar (e.g. Holt et al., 1984). It can be illustrated in the data from this period in the observations from Baker Lake (BL) NWT, Canada, during Dec.1958. This station is at a sufficiently high MLAT (75.1°) that it observes the trough on all 31 days of this study. In addition the station is at an appropriate MLONG to detect both morning and afternoon components corresponding to the dawn and dusk convection cells.

The median foF2 from BL for the month of December is shown in Figure 2a (points) together with a quiet, undisturbed distribution. This quiet distribution is inferred from a station at the same geographic latitude as BL, but it is at a much lower MLAT (60°) so that it does not observe the trough during the daytime. The shaded areas thus correspond to the AM and PM trough components seen by BL.

Figure 2b is a schematic plot of a 2-cell convection pattern (Heelis et al, 1982), the 2 shaded regions corresponding approximately to the areas in which BL, carried by the rotation of the earth beneath the pattern along the circle near 75°, intersects the 2 trough sectors, one in the dawn cell and the other in the dusk cell. Apparently these regions of lower density consist of plasma convected from the nightside where it replaces high density daytime plasma convected away into the polar cap. In addition other mechanisms such as ion chemical effects may play a role in the plasma depletion.

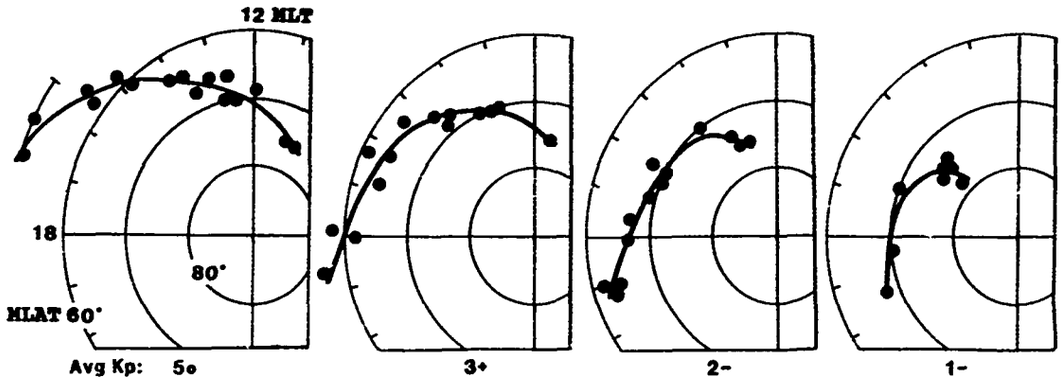


Fig. 3. Activity dependence of the daytime trough. The equatorward boundaries defined by the trough onsets observed in LT by the world-wide array of ionospheric sounders. These boundaries, shown for 4 individual days, are stationary in MLAT/MLT for periods of up to 14 hours UT. This stability contrasts with the extreme variation seen from day to day as designated by Kp averaged over the UT day. The variation with activity results from the expansion and rotation of the dusk convection cell with increasing activity.

Near midday BL observes a maximum in foF2 that approaches the level that would exist in the presence of local solar ion production alone (Figure 2a). In addition this is also near the region of soft particle precipitation in the daytime cusp or cleft. However in Figure 2b this corresponds to the passage of the station through the throat between the 2 convection cells where strong anti-sunward convection can displace locally produced plasma with high density plasma from lower latitudes on the dayside. This effect, reported by Kelly and Vickrey (1984), is also consistent with observations in the southern hemisphere discussed below.

3. ACTIVITY DEPENDENCE

The afternoon trough as characterized by its equatorward boundary, is detected on each of the 31 days of December. It is found to be dependent principally on magnetic activity (Kp) averaged over the period of a day. It is generally stable in MLAT/MLT for many hours and shows little dependence on changes in Kp which take place during such periods. On the most disturbed day, the trough is its most extensive; the trough boundary extends in MLAT/MLT from 75°/1030 to 51°/1630 and is nearly stationary for a duration of 14.5 hours UT as observed by 18 stations spanning 220° MLONG. Figure 3 shows this most disturbed case on the left together with three other days, the average Kp of which decreases from left to right. Note that as the average Kp increases from quiet to disturbed the trough rotates to earlier MLT and expands to lower MLAT.

The fact that the trough contracts poleward during quiet times to very high latitudes where there are very few stations makes it difficult to observe. This behavior has apparently contributed to the mistaken belief that it does not exist during quiet times.

4. HEMISPHERIC ASYMMETRIES

a. LT-MLT RELATIONS

The daytime trough resides in solar geomagnetic coordinates (MLAT/MLT) and is often stationary there, whereas the undisturbed daytime F layer can be regarded as residing in solar terrestrial coordinates (geographic latitude/LT). This section will show that hemispheric differences in these coordinates result in fundamental hemispheric differences in the daytime F layer.

The relation between the latitude and longitude coordinates in the 2 systems is illustrated in a plot of the earth in rectangular coordinates of geographic latitude/geographic longitude in which the Mlat/Mlong coordinates appear as the curved lines in Figure 4 (after Evans et al., 1969). The magnetic meridians through 2 stations are shown, Narssuarssuaq (NQ) in Greenland and Barrow (BW) in Alaska, both stations being at nearly the same MLAT (70°).

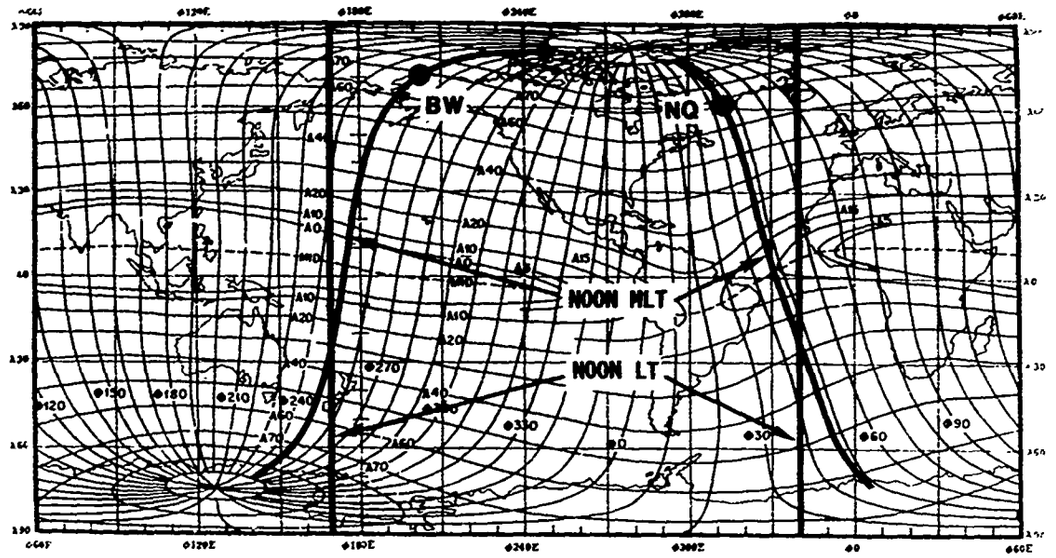


Fig. 4. East-west differences in the relation between LT and MLT. The noon meridian in LT simultaneous with the noon meridian in MLT for 2 stations: Narssuarssuaq (NQ) and Barrow (BW). At NQ noon MLT occurs approximately 2h before noon LT, a situation that is reversed at BW. This is characteristic of the differences between the 2 coordinate systems and results in the difference in the occurrence of the daytime trough at the 2 stations. (Magnetic coordinate map after Evans et al., 1969).

Noon MLT occurs at NQ when its magnetic meridian intersects the subsolar point which is at 23.5° S geographic latitude near winter solstice. The vertical line thus locates the noon LT meridian when NQ is at noon MLT. Accordingly noon MLT occurs at NQ nearly 2 hours earlier than noon LT. At BW on the other hand, the magnetic meridian curves west with decreasing latitude so that noon MLT occurs nearly 2 hours later than noon LT.

As viewed in LT, phenomena that are situated in the MLT reference frame, as is the convection pattern and resulting daytime trough, occur nearly 4 hours apart in LT at the 2 stations. Note that the relation of MLT leading LT in the east and lagging LT in the west is reversed in the southern hemisphere.

The spatial relation between these 2 local time coordinate systems in the high latitude northern hemisphere is shown in a map in MLAT/MLONG of the contours of the LT at which noon MLT occurs (Figure 5a). The 2 time coordinates are the same near 0° and 180° MLONG. In eastern MLONG, noon MLT occurs earlier than noon LT, and in western MLONG noon MLT occurs later that noon LT, increasingly so at increasingly higher latitudes. Thus MLT can be said to lead LT in eastern MLONG, to lag LT in western MLONG and to be in phase with LT near 0° and 180° MLONG.

The relation of these local time coordinates for the southern hemisphere (Figure 5b) shows the east-west effect to be the opposite of that in the north as will be discussed below.

b. NORTHERN HEMISPHERE IONOSPHERIC EFFECTS

The ionospheric effects of these coordinate relations appear in the observations by 5 ionospheric sounders which span the North American continent from NQ in the east to BW in the west near 70° MLAT, the locations of which are shown in Figure 5a.

foF2 as a function of LT for each of the stations of this chain is shown for all days of Dec 1958 in Figure 6, the order top to bottom corresponding to the order of the stations east to west. At NQ the overplot of the distributions shows the increase in foF2 following sunrise to be very similar from day to day, but the afternoon decrease to be extremely variable from day to day. On the other hand the pattern at BW is nearly the opposite of that at NQ, the morning increases variable and the afternoon decreases similar. Accordingly both of these stations see a daytime trough which is variable from day to day but at NQ it is in the afternoon and at BW it is in the morning.

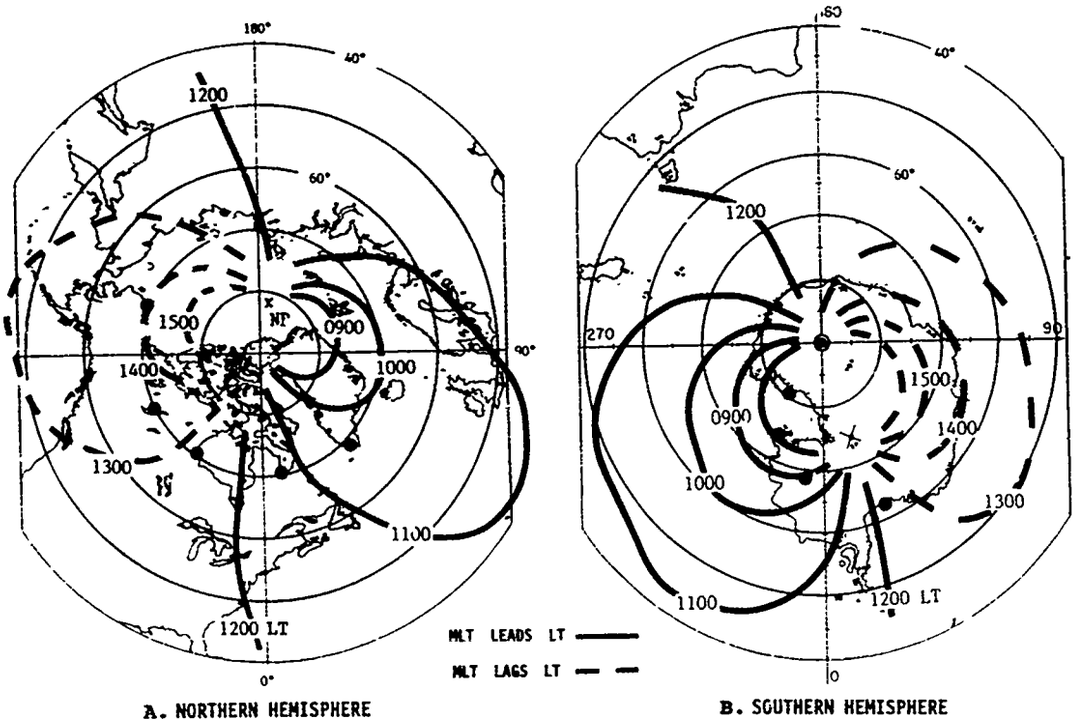


Fig. 5. Spatial relation between LT and MLT for north and south hemispheres. Contours of the LT at which noon MLT occurs mapped in MLAT/MLONG for the 2 hemispheres: (a) Northern hemisphere showing MLT to lead LT in eastern MLONG but to lag LT in western MLONG. The 5 ionospheric sounding stations at ~70°, the locations of which are shown, observe the effects of this difference in the occurrence of the daytime trough (Figure 6). (b) Southern hemisphere (in northern hemisphere projection) showing the LT/MLT relations to be reversed from those of the northern hemisphere. 4 ionospheric sounding stations, the locations of which are shown, span nealy 20° of MLAT at nearly constant geographic latitude. These stations observe the effects of the trough and other convective processes on the F layer (Figures 7 and 8).

Considering all 5 stations, the extent of the afternoon trough decreases and that of the morning trough increases progressively going from east to west. This change accompanies the change in the location of noon MLT plotted at the top of each graph from pre-noon LT to post-noon LT indicating the variation of the 2 coordinate systems. Thus the trough, because it resides in the solar geomagnetic frame of reference, varies systematically with longitude with respect to the undisturbed daytime F layer which resides in the solar terrestrial frame of reference.

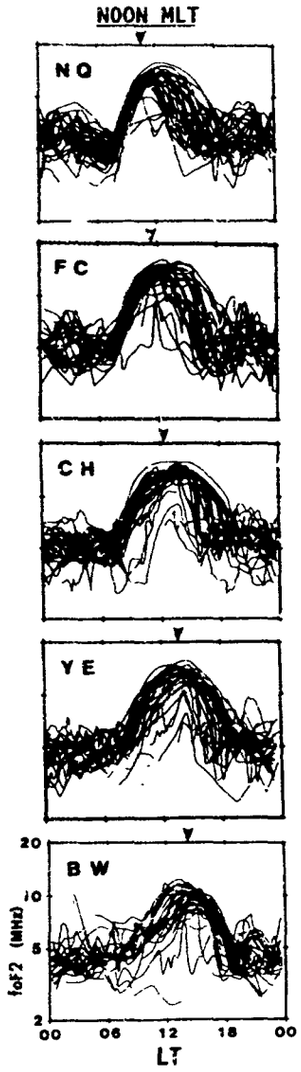
c. SOUTHERN HEMISPHERE IONOSPHERIC EFFECTS

Figure 5b maps in MLAT/MLONG the southern hemisphere polar region in north polar projection (as if viewed through the earth from a point above the north pole) showing the contours of noon MLT as they occur in LT. Here MLT lags LT in the east but leads LT in the west, the opposite of the situation in the north.

Also shown is an array of 4 ionospheric sounding stations all at nearly the same geographic latitude (~78°) but spanning nearly 20° MLAT.

Median values of foF2 for the month of April for the 3 lower latitude stations are shown in Figure 7a, noon MLT indicated at the top of the graph. At Ellsworth (EL) at 62.6° MLAT the distribution is quite regular, the effects of convection on the daytime ionosphere being essentially nil. Note that noon LT and noon MLT are nearly the same.

At Byrd (BD) at 68.7° MLAT the daytime distribution is truncated in the early PM by the trough and the entire distribution is shifted to earlier LT in the direction of noon MLT as compared to EL.



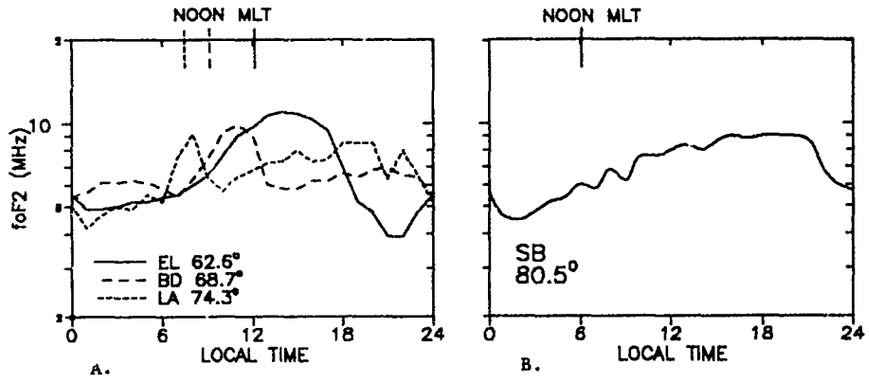


Fig. 7. Daytime trough and other convective effects in the southern hemisphere. Median diurnal distributions of foF2 for April 1958 for the chain of 4 ionospheric sounding stations (Figure 5b) which are at nearly constant geographic latitude (~78° S). (a) The 3 lowest latitude stations. The daytime F layer is displaced in the direction of noon MLT shown at the top of the graph in accordance with the time zones shown in Figure 5b. (b) The highest latitude station at SB. The F layer shows no relation either to noon LT or noon MLT. Instead the minimum is near 0300 UT and the maximum near 1500 UT, the times of minimum and maximum solar illumination of the geomagnetic polar cap. Since SB remains within the polar cap, its ionosphere is dominated by anti-sunward convection and so varies as the illumination of the convection pattern as a whole.

Fig. 6. The longitudinal differences in the daytime F layer produced by the daytime trough in the northern hemisphere. An overplot of diurnal distributions of foF2 for all 31 days of December 1958 for each station in a longitudinal chain at ~70° MLAT shown in Figure 5a. The depletions in foF2 caused by the trough, although highly variable from day to day, occur principally in the afternoon at NQ but in the morning at BW, a change in LT occurrence that takes place progressively with longitude. Noon MLT, shown as the arrow at the top of each graph, shifts from prenoon to post noon over this longitude interval, indicating the variation of the coordinate systems responsible for the variation of the trough.

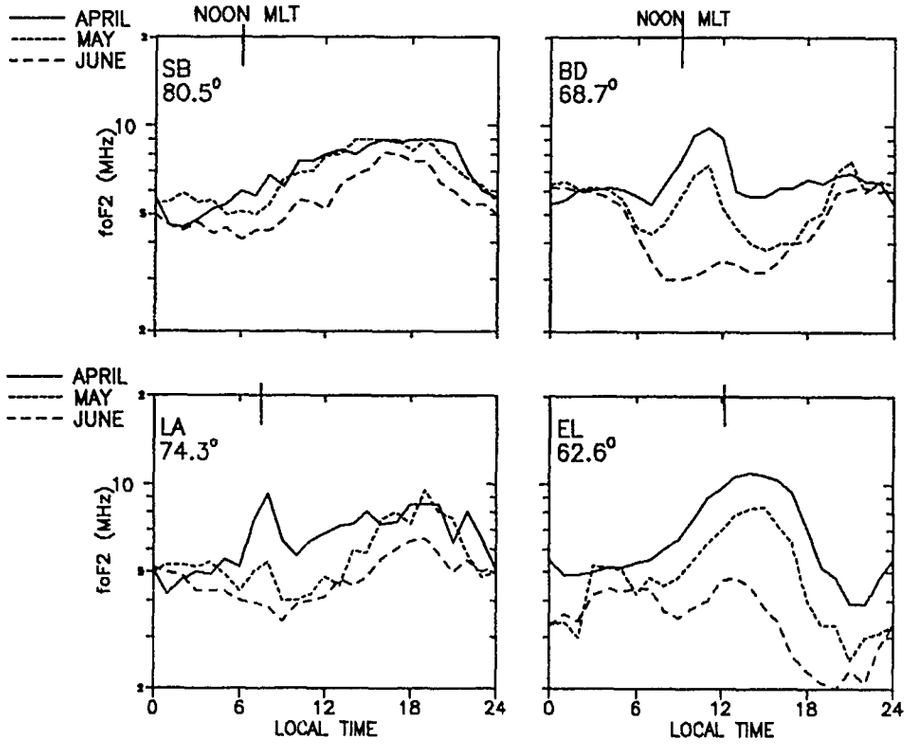


Fig. 8. Variation of the F layer with changing solar ionization for the stations in the southern hemisphere chain. Median diurnal distributions of foF2 for the months of April, May and June identifying as solar-produced, those sectors of the ionosphere which decrease in accordance with the decrease of solar illumination. In particular the reduction of the peak near noon MLT at LA is consistent with solar rather than with particle production. By contrast the near constancy of the levels at BD from evening through morning is consistent with particle rather than with solar production.

At Little America (LA) at 74.3° MLAT there is a maximum at about 08 LT well before significant local solar production. This peak is near the location of noon MLT and so corresponds to the throat of the convection pattern into which plasma of higher density can be transported from a region which is at a later local time and farther equatorward and therefore more sunlit. This interpretation is supported by the seasonal variation described below.

At Scott Base (SB) at 30.5° MLAT (Figure 7b) there is a minimum post midnight and a maximum in the evening. These features are not related to local solar production but are UT dependent as will be discussed below.

(1). ORIGINS OF OBSERVED PLASMAS

Because convection can transport plasma great distances at high latitudes, the source of the plasma is difficult to identify. However it is possible to observe the dependence of the distributions seen in Figure 7 on the variation of solar ion production for a sequence of months from austral autumn to winter, inferring plasma which decreases systematically with season to be of solar origin, and that which remains constant to be of particle origin. The median diurnal distributions of foF2 for 3 months, April, May and June, are shown for each of the 4 stations in Figure 8.

At EL the midday maximum decreases with the decrease in solar production in going from autumn to winter, and at BD a similar decrease is seen in the daytime peak. At LA there is a decrease of similar magnitude that is also consistent with solar production although it is not consistent with local ion production. Apparently daytime plasma is convected from later LT and lower geographic latitude on the dayside into the throat of the convection pattern forming an increase, the

opposite of a trough. Finally the distribution at SB undergoes a decrease but one of much smaller magnitude than that observed by the other stations.

On the other hand, the levels of foF2 at BD, which tend to be temporally constant from about 20 LT through midnight to about 04 LT, are nearly seasonally constant, and so consistent with resulting from production by precipitating particles rather than by solar radiation. This result supports the interpretation of the region called the auroral oval in Figure 1 as resulting primarily from particle precipitation.

(2). UT DEPENDENCE

At SB the broad maximum seen in Figures 7 and 8 occurs near 03 UT and the minimum near 15 UT, times corresponding to the states of maximum and minimum solar illumination of the geomagnetic polar cap. Since SB never escapes from the polar cap, its F layer is dominated by antisunward convection so that its variation is that of the variation of the illumination of the convection pattern. Note that the magnitude of the decreases are not as great as those attributed to solar production alone seen at the other stations in going from autumn to winter.

5. SUMMARY

1. The daytime trough is a major feature of the winter F layer at high latitudes producing order of magnitude depletions in the daytime F layer that are evident in monthly medians.
2. It occurs in the presence of solar ion production, and so is a different phenomenon from the nighttime trough which disappears in the presence of ion production whether solar or particle produced.
3. It is a product of convection, being consistent with low density plasma convected from nighttime so as to displace high density daytime plasma.
4. Because this trough resides in the solar geomagnetic frame of reference, it produces macroscopic hemispheric asymmetries in the daytime F layer in which the properties are reversed east to west and north to south.
5. In the throat region of the convection pattern the opposite of a trough is formed in an enhancement in electron density resulting from high density plasma being convected into a low density plasma region.
6. Control of the F layer by convection increases with increasing MLAT and becomes total within the polar cap.
7. F layer behavior as solar illumination varies with season can be used to infer whether plasma displaced by convection is produced primarily by solar radiation or by precipitating particles.

REFERENCES

Evans, J.E., L.L. Newkirk, and B.M. McCormac, North polar, south polar, world maps and tables of invariant magnetic coordinates for six altitudes: 0, 100, 300, 600, 1000, and 3000 km, DASA Report 2347, Lockheed Palo Alto Research Laboratory, Palo Alto, CA, 1969.

Heelis, R.A., J.K. Lowell, and R.W. Spiro, A model of the high-latitude ionosphere convection pattern, J. Geophys. Res., 87, 6339, 1982.

Holt, J.M., R.H. Wand, and J.V. Evans, Millstone Hill measurements on February 26, 1979, during the solar eclipse and formation of a midday F-region trough, J. Atmos. Terr. Phys., 46, 251, 1984.

Kelly, J.D., and J.F. Vickrey, F region ionospheric structure associated with antisunward flow near the dayside polar cusp, Geophys. Res. Lett., 11, 907, 1984.

Muldrew, D.B., F-layer ionization troughs deduced from Alouette data, J. Geophys. Res. 70, 2635, 1965.

Rodger, A.S., L.H. Brace, W.R. Hoegy, and J.D. Winningham, The poleward edge of the mid-latitude trough - its formation, orientation and dynamics, Jour. Atmos. Terr. Phys., 48, 715, 1986.

Sojka, J.J., R.W. Schunk, and J.A. Whalen, The longitude dependence of the dayside F-region trough: a detailed model-observation comparison, submitted to J. Geophys. Res., 1990.

Whalen, J.A., The daytime F layer trough observed on a macroscopic scale, J. Geophys. Res., 92, 2571, 1987.

Whalen, J.A., The daytime F layer trough, in Physics of Space Plasmas (1988), edited by T. Chang, G.B. Crew and J.R. Jasperse, SPI Conference Proceedings and Reprint Series, No. 8, Scientific Publishers, Cambridge, Mass., 1989a.

Whalen, J.A., The daytime F layer trough and its relation to ionospheric-magnetospheric convection, J. Geophys. Res., 94, 17,169, 1989b.

AD-P006 271

E-LAYER DUE TO PURE PROTON AURORA

B. Basu and J.R. Jasperse
Geophysics Laboratory
Hanscom Air Force Base, MA 01731

ABSTRACT

It is now well-known that in the evening sector of the auroral oval and at latitudes equatorward of the electron precipitation boundary, the auroral E-layer is produced solely by the precipitating energetic protons, the H atoms that are created by charge-exchange collisions of protons, and by the secondary electrons that are generated. A numerical code has been developed to solve the coupled proton-H atom linear transport equations for a pure proton aurora. The solutions give proton and H atom fluxes as a function of altitude, energy, and pitch angle, for a given incident proton flux. From the particle fluxes, energy depostion rates, ionization rates, and electron density profiles (EDP) are calculated. Results of these calculations are presented. Comparisons between calculated EDP and radar measured EDP are shown using nearly coincident satellite data for the incident proton spectra.

INTRODUCTION

The continuous (diffuse) aurora and the associated auroral E-layer are produced by precipitating energetic electrons and protons of magnetospheric origin. The relative contribution from each component to the total energy flux reaching the atmosphere varies as a function of latitude and local time. In the evening sector of the auroral oval, in particular, the ratio of proton to electron energy flux increases with decreasing latitude, and during quiet times it is not unusual for the energy flux carried by protons to actually exceed that carried by electrons near the equatorward edge of the particle precipitation region. This was first shown by Sharber [1981] using ISIS 2 data. More recently, Gussenhoven et al. [1987] have done an extensive study of the latitudinal distributions of electron and proton precipitation as a function of local time using DMSP data. Their study also showed that in the evening sector, the proton precipitation boundary on the average is displaced equatorward relative to the electron precipitation boundary. The magnitude of such displacement may become as large as 4° in invariant latitude as is exemplified in the NOAA 6 data presented in Figure 1.

The total energy flux carried by electrons and protons, obtained from the on-board integration of the detector measurements, is shown in Figure 1 along with electron density data measured simultaneously by Chatanika radar. It is apparent from the data that within the latitudinal range between 64° and 68° invariant, protons carried all of the energy flux. Thus, the electron density measured by the radar, within this latitudinal range, is produced entirely by the precipitating protons, the H atoms that are created by charge-exchange collisions of protons, and by the secondary electrons that are generated.

A theoretical treatment of the proton aurora requires the solution of two transport equations - one for the proton flux and the other for the H atom flux - coupled to each other due to the charge-changing (charge exchange and stripping) collisions. Once the fluxes are known, all the other related quantities can be easily calculated. Previously, these coupled equations were solved analytically [Jasperse and Basu, 1982] and a code was developed to evaluate the analytic expressions for the particle fluxes from which energy deposition rates, ionization rates, and electron density profiles were calculated using simple formulae.

Recently, we have developed a fully numerical code that solves the coupled transport equations to obtain particle fluxes as a function of altitude, energy and pitch angle, when the incident proton flux is specified. The transport equations are simplified by making use of the plane-parallel geometry and the forward scattering approximations only. Otherwise, the equations and their numerical solutions are exact. The purpose of this paper is to present the results from this new code. Since the numerical code avoids all the other approximations that were necessary for obtaining the earlier closed-form analytic solutions, the results from this code are more accurate. A practical motivation for this work is that in our current studies of the optical emissions produced by precipitating protons and H atoms, we found that certain emission features are very sensitive to the value of the E- and lower F- region proton

91-09664

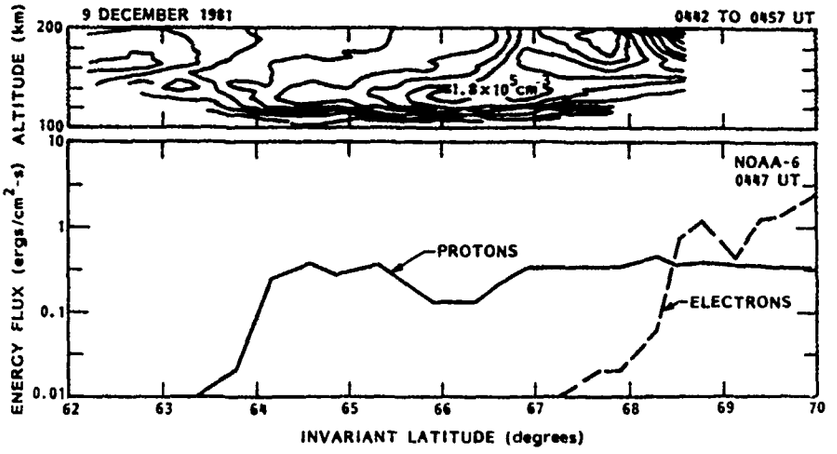


Fig. 1. NOAA satellite data showing auroral E layer produced entirely by precipitating protons.

and H atom fluxes. For these emissions, we concluded that it was necessary to obtain a more accurate solution for the particle fluxes than those provided by the earlier analytical results.

The new code solves a set of two coupled integral equations which are obtained from the differential forms of the linear transport equations. These are presented in the next section. In the subsequent sections, a brief description of the code and the results of numerical calculations are presented.

BASIC EQUATIONS

In plane-parallel geometry, when forward scattering approximation to the angular dependence of the differential cross sections for various processes is used, the steady-state linear transport equations for protons and H atoms are

$$\left[\mu\frac{\partial}{\partial z}+\sum_{\alpha} n_{\alpha}(z)\,\sigma_{\alpha,P}(E)\right]\Phi_P(z,E,\mu)$$

$$=\sum_{\alpha} n_{\alpha}(z)\left\{\sum_{k}\sigma_{\alpha,P}^{k}\left[E+W_{\alpha,P}^{k}(E)\right]\Phi_P\left(z,E+W_{\alpha,P}^{k}(E),\mu\right)\right.$$

$$\left.+\;\sigma_{\alpha}^{01}\left[E+W_{\alpha}^{01}(E)\right]\Phi_H\left(z,E+W_{\alpha}^{01}(E),\mu\right)\right\} \qquad (1)$$

and

$$\left[\mu\frac{\partial}{\partial z}+\sum_{\alpha} n_{\alpha}(z)\,\sigma_{\alpha,H}(E)\right]\Phi_H(z,E,\mu)$$

$$=\sum_{\alpha} n_{\alpha}(z)\left\{\sum_{k}\sigma_{\alpha,H}^{k}\left[E+W_{\alpha,H}^{k}(E)\right]\Phi_H\left(z,E+W_{\alpha,H}^{k}(E),\mu\right)\right.$$

$$\left.+\sigma_{\alpha}^{10}\left[E+W_{\alpha}^{10}(E)\right]\Phi_P\left(z,E+W_{\alpha}^{10}(E),\mu\right)\right\} \qquad (2)$$

where coupling between the proton and the H atom fluxes, denoted by Φ_P and Φ_H respectively, is evident. Here n_α is the density of the neutral species α, μ is the cosine of the angle between particle velocity and the z axis, which is positive upward and E is the particle energy. The z axis is taken to be along the geomagnetic field line. In the northern hemisphere, $\mu \equiv \cos(\pi - \theta)$, where θ is the pitch angle, while in the southern hemisphere, $\mu \equiv \cos\theta$. The summation index k labels excitation and ionization collisions, $\sigma^k_{\alpha,\beta}(E)$ where $\beta = P, H$, is the total cross section for the k-type collision between β and α, and, $W^k_{\alpha,\beta}(E)$ is the energy-loss associated with that collision. The total cross sections for charge-exchange and stripping collisions and the energy losses associated with them have been denoted by $\sigma_\alpha^{10}(E)$, $\sigma_\alpha^{01}(E)$, $W_\alpha^{10}(E)$, and $W_\alpha^{01}(E)$, respectively. Furthermore,

$$\sigma_{\alpha,P}(E) \equiv \sum_k \sigma^k_{\alpha,P}(E) + \sigma_\alpha^{10}(E)$$

$$\sigma_{\alpha,H}(E) \equiv \sum_k \sigma^k_{\alpha,H}(E) + \sigma_\alpha^{01}(E)$$

We next define the scattering depth, $\tau_\beta(z, E)$, as

$$\tau_\beta(z, E) \equiv \sum_\alpha \sigma_{\alpha,\beta}(E) \int_z^{z_o} n_\alpha(z')\, dz'$$

where $\beta = P, H$ and z_O denotes the upper boundary of the atmosphere. Then, Eqs. (1) and (2) can be converted to the integral equation forms given by

$$\Phi_P(z, E, \mu) = \varepsilon\Phi_0(E, \mu) \exp[\tau_P(z, E)/\mu]$$

$$-\frac{1}{\mu}\sum_\alpha \int_z^{z_o} dz' n_\alpha(z') \exp\{[\tau_P(z, E) - \tau_P(z', E)]/\mu\}$$

$$\cdot \left\{ \sum_k \sigma^k_{\alpha,P}\left[E + W^k_{\alpha,P}(E)\right] \Phi_P\left(z', E + W^k_{\alpha,P}(E), \mu\right) \right.$$

$$\left. + \sigma_\alpha^{01}\left[E + W_\alpha^{01}(E)\right] \Phi_H\left(z', E + W_\alpha^{01}(E), \mu\right) \right\} \qquad (3)$$

and

$$\Phi_H(z, E, \mu) = -\frac{1}{\mu}\sum_\alpha \int_z^{z_o} dz' n_\alpha(z') \exp\{[\tau_H(z, E) - \tau_H(z', E)]/\mu\}$$

$$\cdot \left\{ \sum_k \sigma^k_{\alpha,H}\left[E + W^k_{\alpha,H}(E)\right] \Phi_H\left(z', E + W^k_{\alpha,H}(E), \mu\right) \right.$$

$$\left. + \sigma_\alpha^{10}\left[E + W_\alpha^{10}(E)\right] \Phi_P\left(z', E + W_\alpha^{10}(E), \mu\right) \right\} \qquad (4)$$

for $-1 \le \mu < 0$, and zero otherwise. The boundary conditions used in the derivation of (3) and (4) are: $\Phi_P (z = z_o, E, \mu) = \Phi_o(E, \mu)$ and $\Phi_H (z = z_o, E, \mu) = 0$, for $-1 \le \mu < 0$ which defines the downward hemisphere. The use of plane-parallel geometry implies the neglect of the atmospheric lateral spreading of the incident proton flux due to charge-changing collisions. The correction factor ε in Eq. (3) takes into account this spreading effect [see, Jasperse and Basu, 1982]. For an arc width of 200 km, which is typical of many naturally occurring aurorae, $\varepsilon \simeq 0.75$.

Equations (3) and (4) are solved numerically to obtain $\Phi_P (z, E, \mu)$ and $\Phi_H (z, E, \mu)$ for a given $\Phi_O (E, \mu)$. A brief description of the numerical code follows.

DESCRIPTION OF THE CODE

Equations (3) and (4) are a system of two coupled equations which are analytic-type equations on μ, difference-type equations on E, and Volterra-type integral equations on z. In order to solve these equations numerically, a finite, two-dimensional grid of evenly spaced E-points from E_{min} to E_{max} and unevenly spaced z-points from z_{min} to z_o is chosen for each value of μ. E_{max} is chosen sufficiently large so that the particle fluxes are negligibly small for energies greater than E_{max} and are set to zero there. When boundary conditions are imposed, linear interpolation is used to find Φ_P and Φ_H at each of the E x z grid points for each value of μ. This is possible without inverting a matrix because of the specific form of (3) and (4). Care had to be taken so that the quadrature approximation used to calculate $\tau_P (z, E)$ and $\tau_H (z, E)$ is consistent with that used in the z integrations.

After calculating the fluxes $\Phi_P (z, E, \mu)$ and $\Phi_H (z, E, \mu)$, the code then uses standard quadrature methods to calculate the hemispherically averaged fluxes, defined as

$$\Phi_\beta (z, E) \equiv \int_{-1}^{0} d\mu \, \Phi_\beta (z, E, \mu), \qquad \beta = P, H, \tag{5}$$

and the energy deposition rate, given by

$$\eta_E (z) \equiv -2\pi \int_{E_{min}}^{E_{max}} dE \int_{-1}^{0} d\mu \, \mu E \frac{\partial}{\partial z} \left[\Phi_P (z, E, \mu) + \Phi_H (z, E, \mu) \right]. \tag{6}$$

The code also calculates the following two quantities:

$$\Gamma \equiv -2\pi \int_{z_{min}}^{z_o} dz \int_{E_{min}}^{E_{max}} dE \int_{-1}^{0} d\mu \, \mu \frac{\partial}{\partial z} \left[\Phi_P (z, E, \mu) + \Phi_H (z, E, \mu) \right], \tag{7}$$

and

$$Q \equiv \int_{z_{min}}^{z_o} dz \, \eta_E (z), \tag{8}$$

in order to check particle and energy conservations. For example, if the incident flux is a Maxwellian given by

$$\Phi_o (E) = \left(\frac{Q_o}{2\pi E_o^3} \right) E \exp (-E / E_o)$$

where E_O is the characteristic energy and Q_O is the downward energy flux, then the particle conservation requires

$$\Gamma = -2\pi\varepsilon \int_0^\infty dE \int_{-1}^{0} d\mu \, \mu \, \Phi_o (E) \equiv \varepsilon Q_o / (2E_o),$$

and the energy conservation requires $Q = \varepsilon Q_0$. The factor ε takes into account the correction due to atmospheric spreading effect.

The stability of the numerical calculations was checked by increasing the number of E x z grid points until no significant change in the numerical results occurred. For a typical calculation, we found that so far as the energy deposition rate is concerned, 10 points for μ ($-1 \leq \mu < 0$), 70 points for z ($90 \text{ km} \leq z \leq z_0 = 600 \text{ km}$), and 500 points for E ($E_{min} = 1 \text{ keV} \leq E \leq E_{max}$) were satisfactory. To obtain the particle fluxes more accurately, more E points are necessary.

Three neutral constituents (N_2, O_2, and O) are considered and their densities are given by the MSIS-86 model [Hedin, 1987]. The total cross sections for the inelastic processes given in Figures 1-3 of Basu et al. [1987] are used. The values of $W^k_{\alpha,\beta}(E)$ for various inelastic processes are given in Basu et al. [1990].

RESULTS AND DISCUSSIONS

The neutral atmosphere used in the calculations is typical of the eveningtime auroral region. In all the calculations, energy conservation is satisfied to within ± 0.05%, and the particle conservation is satisfied to within ± 2%.

In Figure 2 we have presented the calculated energy deposition rates for three Maxwellian incident proton fluxes. The characteristic energies E_0 of these Maxwellians are 4, 8, and 16 keV, respectively, and the downward energy flux Q_0 for each Maxwellian is 0.5 erg cm^{-2} s^{-1}. The penetration heights, defined to be the altitudes where energy deposition rates have maximum values, for the three Maxwellians are 129, 123, and 118 km, respectively. In comparing the numerical results with our earlier analytical results [Jasperse and Basu, 1982] we find that although the particle fluxes differ appreciably, the energy deposition rates can be made to agree to within less than ± 10% provided that the W, which appears in the analytical solutions, is considered to be a function of E_0 with the values W = 23, 32, and 44 eV for E_0 = 4, 8, and 16 keV, respectively. In fact, since the energy losses $W^k_{\alpha,\beta}(E)$ increase sharply as the particle energy E increases, different values of W, which is the average energy loss, should be used for Maxwellians with different characteristic energies.

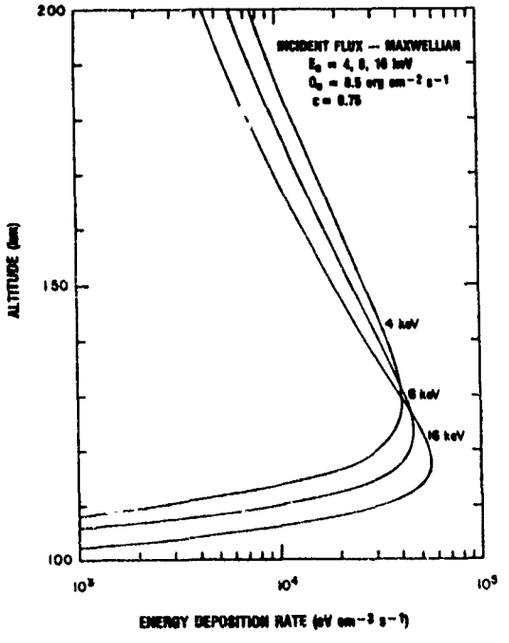


Fig. 2. Energy deposition rate versus altitude as obtained from the numerical code

In the following, we present comparisons between calculated electron density profile (EDP) and radar measured EDP for a pure proton aurora, as shown in the NOAA 6 data presented in Figure 1. For the incident proton spectra needed for our calculations, we use those measured by the NOAA 6 detectors.

In order to calculate the ionization rate rigorously, we need to solve an electron transport equation with the secondary electron production by the protons and the H atoms as a source term. This work has been done and the results will be given in Strickland et al. [1990]. For the purpose of this paper, we calculate the ionization rate by using the formula:

$$\eta_I(z) = -2\pi \int_{E_{min}}^{E_{max}} dE\left[E/W_S(E)\right] f_I(E) \int_{-1}^{0} d\mu\, \mu \frac{\partial}{\partial z}\left[\Phi_P(z, E, \mu) + \Phi_H(z, E, \mu)\right] \tag{9}$$

Here $W_S(E)$ is the estimated average energy expended in creating an electron-ion pair, the values of which are given in Jasperse and Basu [1982], and $f_I(E)$ is the fraction of the total deposited energy that is available for the creation of electron-ion pairs. Estimated values of $f_I(E)$ are given in Basu et al. [1987]. We use the incident proton spectra measured by NOAA detectors to calculate $\Phi_P(z, E, \mu)$ and $\Phi_H(z, E, \mu)$ and then calculate $\eta_I(z)$ from Eq. (9). Next, we calculate electron density profile (EDP) from $\eta_I(z)$ by using

$$n_e(z) = [\eta_I(z)/\alpha(z)]^{1/2} \tag{10}$$

where for the effective recombination coefficient $\alpha(z)$ we use

$$\alpha(z) = (2.5 \times 10^{-6}) \exp[-z(km)/51.2] \tag{11}$$

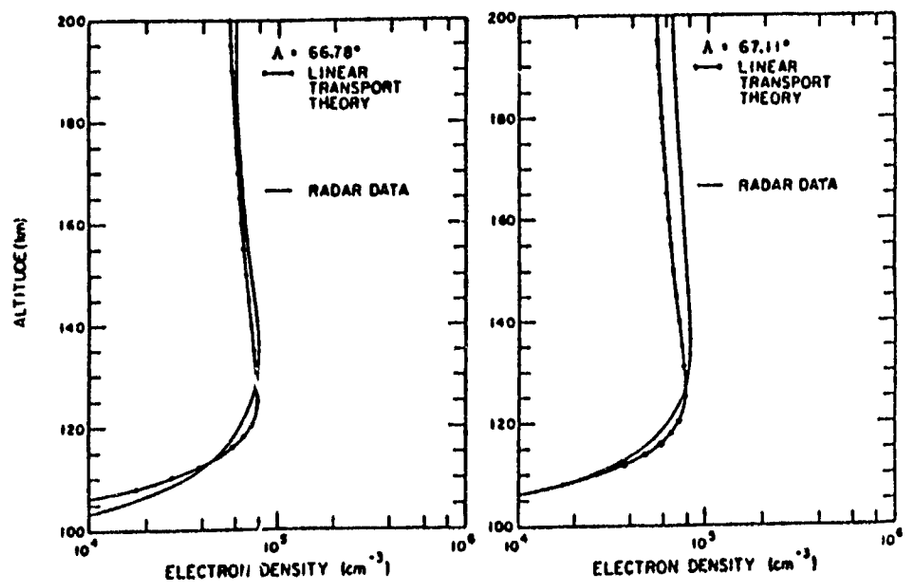


Fig. 3. Comparisons between calculated electron density profile (EDP) and radar measured EDP. Incident proton spectra are given by NOAA satellite data.

Results of the calculations are shown in Figure 3 for two incident proton spectra measured by NOAA at two selected invariant latitudes. Also shown in Figure 3 in solid curves are the radar measured EDP. It is evident that the calculated profiles agree quite well with the observed profiles.

ACKNOWLEDGEMENTS

This work was supported in part by the Air Force Office of Scientific Research. Numerical analysis was done by Neil Grossbard at the Institute for Space Research, Boston College.

REFERENCES

Basu, B., J.R. Jasperse, and N.J. Grossbard, A numerical solution of the coupled proton-H atom transport equations for the proton aurora, submitted to J. Geophys. Res., 1990.

Basu, B., J.R. Jasperse, R.M. Robinson, R.R. Vondrak, and D.S. Evans, Linear transport theory of auroral proton precipitation: a comparison with observations, J. Geophys. Res., 92, 5920, 1987.

Gussenhoven, M.S., D.A. Hardy, and N. Heinemann, The equatorward boundary of auroral ion precipitation, J. Geophys. Res., 92, 3273, 1987.

Hedin, A.E., MSIS-86 thermospheric model, J. Geophys. Res., 92, 4649, 1987.

Jasperse, J.R., and B. Basu, Transport theoretic solutions for auroral proton and H atom fluxes and related quantities, J. Geophys. Res., 87, 811, 1982.

Sharber, J.R., The continuous (diffuse) aurora and auroral-E ionization, in Physics of Space Plasmas, vol. 7 edited by T.S. Chang, B. Coppi, and J.R. Jasperse, p. 115, Scientific Publishers, Cambridge, MA, 1981.

Strickland, D.J., B. Basu, J.R. Jasperse, and R.E. Daniell, Jr., Theory of the coupled electron-proton-hydrogen atom aurora, to be submitted to the J. Geophys. Res., 1990.

AD-P006 272

91-09665

SPACE-TIME STRUCTURE OF AURORAL RADIO ABSORPTION EVENTS OBSERVED WITH THE IMAGING RIOMETER AT SOUTH POLE

J. K. Hargreaves,* D. L. Detrick, and T. J. Rosenberg

Institute for Physical Science and Technology
University of Maryland
College Park, MD 20742

*Also, Environmental Science Division, University
of Lancaster, Lancaster, LA1 4YQ, England

ABSTRACT

An imaging riometer system comprising 49 independent beams has been operating at South Pole station since January, 1988. A study of intense, short-duration events from the pre-midnight sector has defined their typical shape as elliptical, with axial ratio 2.3 oriented along the local L-shell. The space-time evolution shows rapid intensifications of the moving absorption patches.

INTRODUCTION

For more than twenty-five years, the riometer has been a standard instrument for monitoring the incidence of abnormal radio absorption in the D region (*Hargreaves*, 1969), and it has been well established that in auroral regions most of the sporadic absorption events are due to energetic (> 20 keV) electrons penetrating below 100 km. For practical reasons most riometer installations use a relatively small antenna with the consequence that the spatial resolution is about 100 km and smaller structures in the absorption are not resolved. In order to achieve finer resolution, some narrow-beam systems, having larger antennas, have been constructed (*Ansari*, 1964; *Berkey*, 1968; *Theander*, 1972; *Nielsen and Axford*, 1977; *Nielsen*, 1980; *Kikuchi et al.*, 1988). Whereas the earlier systems formed a single beam, Nielsen and collaborators employed the Butler matrix (*Butler*, 1966) to form several beams from a single array.

At South Pole the "Imaging Riometer for Ionospheric Studies," IRIS, has taken the concept of the multiple beam riometer to a further level of sophistication (*Detrick and Rosenberg*, 1990). From an array of 64 crossed dipoles, 49 independent beams are formed simultaneously by means of a network built up from 15 Butler matrices. The beam width depends on the elevation, and is 12° between 3 dB points for the overhead beams. Adjacent beams touch between the 3 dB and 6 dB levels. Projected to the lower ionosphere at 90 km altitude (Figure 1), the inner beams are about 20 km across and the outer are wider and more elongated. The whole array covers 200 to 300 km; for comparison, the dotted circle in Figure 1 shows the area viewed by a typical wide-beam riometer.

The signals received from the 49 beams are recorded by rapid switching into seven riometers. The dwell time for each step is about 1/8 second, and the whole array is scanned every second. The data are recorded digitally, and are averaged to ten seconds for the present analysis.

The instrument operates at a radio frequency of 38.2 MHz. Following standard riometer technique it is necessary to derive 49 quiet-day curves instead of only one, but there are several outstanding advantages. Compared with most standard riometers, the IRIS covers a wider area, has better spatial resolution, achieves excellent temporal resolution, can measure the spatial size of absorption patches up to 200 km across, and can observe their motion. With only a single beam one cannot say whether a change of absorption is due to a movement of the absorbing region or to a true change of intensity. The IRIS enables us to examine this question for different kinds of events

A STUDY OF INTENSE, SHORT-DURATION EVENTS OCCURRING BEFORE MIDNIGHT

Quicklook

Statistically, the absorption events observed at South Pole fall into two groups, one before midnight and the other before noon (*Hargreaves et al.*, 1964). Quicklook plots show that the pre-midnight group contains some short but intense events, often with spiky appearance, and three have been selected for study (Table 1).

Table 1: Selected events

Event	Date	Time (UT)	Peak absorption wide-beam (dB)	Peak absorption IRIS (dB)	Ratio
1	1988 July 22	2039-2046	1.75	8.9	5.1
2	1988 May 18	2329-2335	1.5	4.2	2.8
3	1988 Feb 24	0232-0237	1.2	3.4	2.8

Two points immediately stand out. First, in each case the greatest absorption seen in any of the narrow beams is considerably larger than that in the wide beam, and this means that the precipitation was not uniform over the field of view of the wide-beam riometer. Second, the first event is remarkable for its intensity: 8.9 dB at 38.2 MHz is equivalent to 14.4 dB at 30 MHz, making this possibly the most intense event of the auroral kind that has yet been reported. Figure 2 compares the event as seen by the wide-beam riometer and by one of the narrow beams, and this comparison emphasizes that the event contains fine structure that is not visible to a wide-beam riometer.

Properties at the maxima

A contour plot of the absorption at the peak of the event of July 22 (Figure 3) shows a roughly elliptical absorption patch measuring about 150 km x 60 km between absorption levels half that at the peak. The patch is clearly not circular, but neither is it a highly elongated strip. It has sometimes been supposed that events of this type (high intensity, short duration, pre-midnight) might be extended strips or bands that move rapidly across the observing beam. That is not so in this case, and the events of May 18 and February 24 present similar results. Table 2 gives the median values from 23 measurements taken during the growth and decay of these three events.

Table 2: Median results

Major dimension	205 km
Minor dimension	85 km
Axial ratio	2.3
Orientation	65°

To consider the possible significance of the axial ratio, we suppose that the energetic particle precipitation producing the absorption is due to an active region in the magnetosphere situated near the nose of the field-line passing through the antenna. A circular region of diameter D, at the nose of the field-line, maps in along dipole field-lines to an elliptical patch at the Earth's surface, having dimension $y = D/L^{3/2}$ and $x = D/(L^{3/2}(4 - 3/L)^{1/2})$ in the zonal and the meridional directions, respectively, where L is the usual L-parameter. The axial ratio $x/y = (4 - 3/L)^{1/2}$. At South Pole, $L \sim 13$, giving $x/y = 1.94$. This compares tolerably with the observed value of 2.3, though the comparison could obviously be refined by more realistic field-line modeling. The circular active region in the magnetosphere would be 8000 to 9000 km in diameter.

The orientation of the South Pole IRIS was set by magnetic compass, and the markings SWNE on Figure 1 are notional directions. Computed L-shells show that L = 13.2 (invariant latitude –74°) passes virtually through the center of the beam pattern at an angle of 73° measured clockwise from notional S-N. This is almost the same as the measured orientation of the absorption patches (Table 2). Figure 4 illustrates a contour plot from the event of February 24 for which the orientation was about average. There is some variation in the orientation from event to event, but the data are not sufficient to indicate any systematic variation with season or time of day.

Finally, it appears that for each of the events studied there is a relation between the magnitude of the event and its area (Figure 5), where the area is defined by the contour of absorption half that of the peak. In each example, the event covers the smallest area while at its most intense, suggesting that the additional precipitation is confined to a more localized region.

Dynamics

Each of the events showed movement and, although the movements are not simple, it is clear that they were not in themselves responsible for the growth and decay of the major absorption as observed. In each case, the peak absorption (Table 1) was due to an intensification lasting about one minute, and the absorption patch did not move very far in that time. However, if we look over a longer period then movements can be seen.

In Figure 6, a cross-section of the absorption taken from the seven beams along the S-N axis is plotted against time. This plot, for the July 22 event, covers 193 km in distance and fifteen minutes of time at 10-s resolution. The event moved into the field of view from the equatorward side (north), rapidly intensified and decayed (at point 47), and retreated towards the equator. The events of May 18 and February 24 were similar in so far that intensifications of short duration occurred within a longer period showing movements, though each was different in detail.

Quasi-periodic fluctuations appear in some events. The local peak at point 33 in Figure 6 maximized at 8 dB, though the peak was off the S-N axis. The time between that and the peak at point 47 was 140-s. The question of periodicities in IRIS data remains to be investigated in more detail.

CONCLUSIONS

The IRIS at South Pole is able to resolve spatial detail in ionospheric radio absorption regions down to 20 km and temporal detail to one second. A study of three intense short-duration events has shown that the precipitation region is elliptical, typically 205 x 85 km. The axial ratio of 2.3 and the orientation near the local L-shell (on average) are consistent with a source region about 8000 km across in the magnetosphere near the nose of the local field-line. The area covered by the event tends to shrink as the event intensifies.

Movements of the absorbing patches can be observed but the intense events studied here are due to short-lived intensifications of absorption rather than to movement of steady absorption over the riometer beam. Quasi-periodic fluctuations are observed in some cases.

ACKNOWLEDGMENTS

The IRIS project from design to installation was supported by NSF grant DPP 8419272. Operation and maintenance of the instrument, as well as data processing and analysis, are being supported by NSF grants DPP 8610061 and DPP 8818229.

REFERENCES

Ansari, Z. A., The aurorally associated absorption of cosmic noise at College, Alaska, *J. Geophys. Res. 69*, 4493 (1964).

Berkey, T., Co-ordinated measurements of auroral absorption and luminosity using the narrow beam technique, *J. Geophys. Res. 73*, 319 (1968).

Butler, J. L., Digital, matrix and intermediate frequency scanning, in *Microwave Scanning Antennas. Vol. 3, Array Systems*, R. C. Hansen, Ed., p. 262. Academic Press, New York (1966).

Detrick, D. L., and Rosenberg, T. J., A phased-array radiowave imager for studies of cosmic noise absorption, *Radio Sci.*, in press (1990).

Hargreaves, J. K., Auroral absorption of HF radio waves in the ionosphere - a review of results from the first decade of riometry, *Proc. IEEE 57*, 1348 (1969).

Hargreaves, J. K., Chivers, H. J. A., and Petlock, J. D., A study of auroral absorption events at the South Pole, 1. Characteristics of the events, *J. Geophys. Res. 69*, 5001 (1964).

Kikuchi, T., Yamagishi, H., and Sato, N., Eastward propagation of CNA pulsations of Pc 4-5 periods in the morning sector observed with scanning narrow beam riometer at L = 6.1, *Geophys. Res. Lett. 15*, 168 (1988).

Nielsen, E., Dynamics and spatial scale of auroral absorption spikes associated with the substorm expansion phase, *J. Geophys. Res. 85*, 2092 (1980).

Nielsen, E., and Axford, W. I., Small scale auroral absorption events associated with substorms, *Nature 267*, 502 (1977).

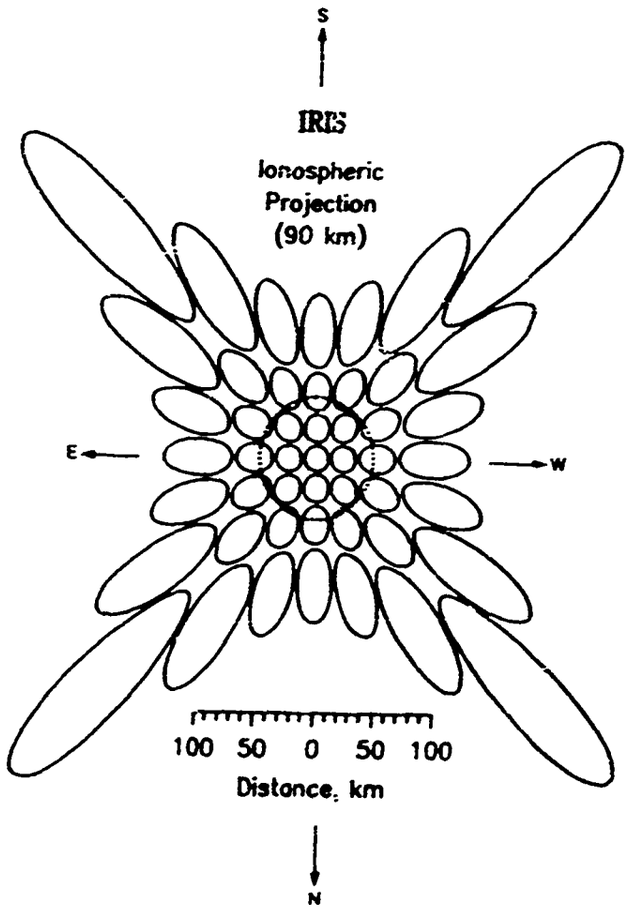


Figure 1. IRIS beams projected to 90 km altitude.

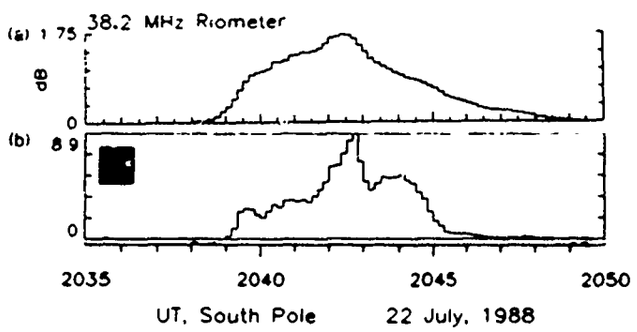


Figure 2. Wide (a) and narrow (b) beam observations of sharp event on 1988 July 22

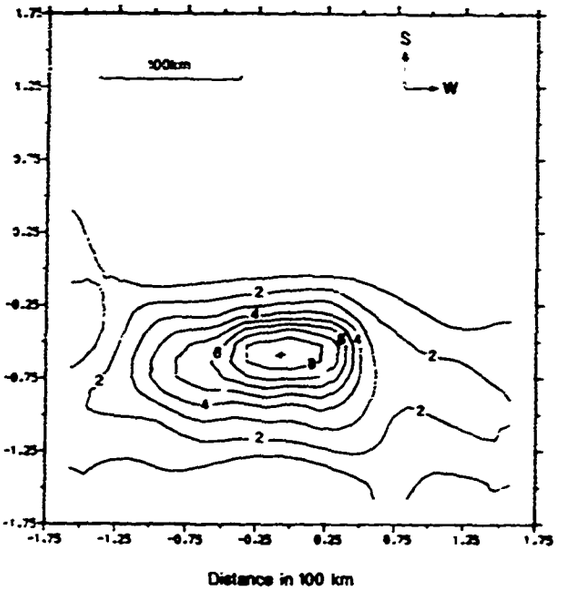


Figure 3. Contour plot of absorption patch at peak of event on 1988 July 22 as viewed from above.

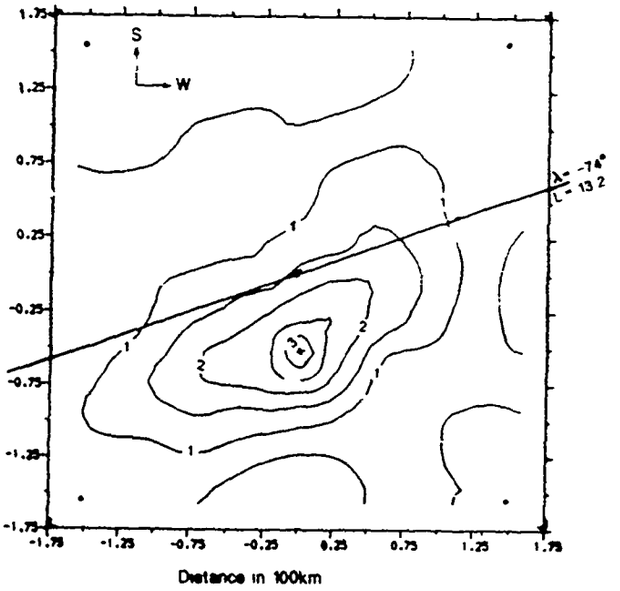


Figure 4. Contour plot of absorption patch on 1988 February 24, showing orientation along L-shell as viewed from above.

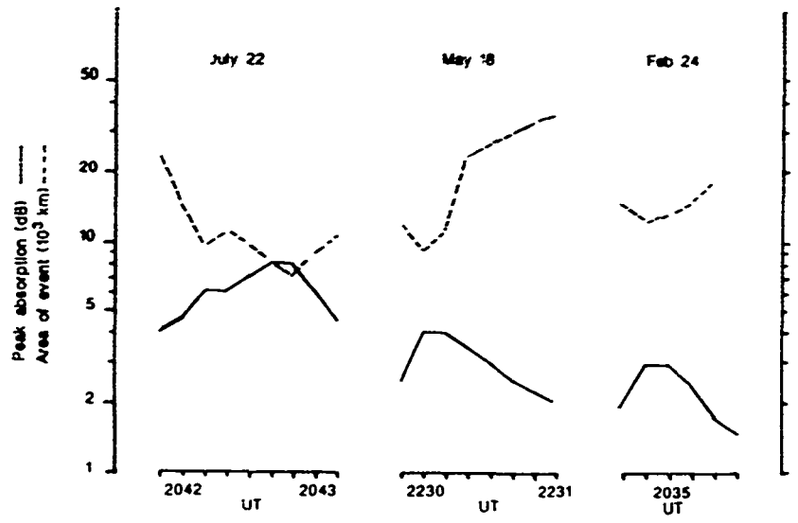


Figure 5. Relations between event magnitude and area.

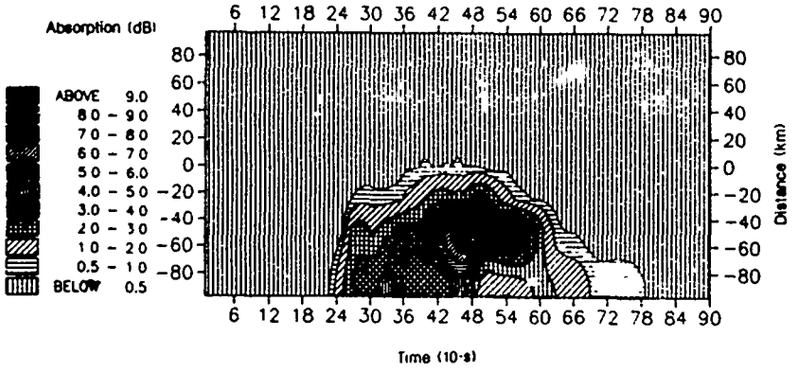


Figure 6. Event of 1988 July 22 plotted against S-N distance and time. South (poleward) is to the top.

AD-P006 273

X-RAY IMAGERY OF ENERGETIC ELECTRON INPUTS TO THE ATMOSPHERE AT HIGH LATITUDES

W. L. Imhof, D. W. Datlowe, E. E. Gaines, J. Mobilia, R. R. Vondrak, H. D. Voss,
and R.M. Robinson
Lockheed Palo Alto Research Laboratory
3251 Hanover Street
Palo Alto, California 94304

ABSTRACT

The precipitation of electrons > 4 keV into the atmosphere occurs at locations spanning a wide range of latitudes and longitudes. The regions where these electrons precipitate have been mapped with a bremsstrahlung x-ray spectrometer in the SEEP payload on the S81-1 spacecraft. Spike-like x-ray emission regions are observed from precipitation originating in the radiation belts. A spike is most often present near the outer edge of the belt, frequently in the form of an arc at a constant invariant latitude. Poleward of the radiation belt boundary, strong arcs are observed as well as weaker patches and widespread events. The arcs are most often singular, but pairs of adjacent arcs are also observed. Detailed energy spectrum and intensity versus latitude profiles of these events have also been measured with electron spectrometers on the satellite. The energy inputs to the atmosphere from electron precipitation >4 keV are generally concentrated in very localized regions in contrast to the more widespread inputs from lower energy electrons. This may produce significant gradients in ionization in the upper atmosphere and thereby affect radio wave propagation. The energy deposition profiles at these restricted positions are presented and compared with the profiles due to cosmic rays which occur at all latitudes and longitudes.

91-09666

INTRODUCTION

The precipitation of energetic electrons into the atmosphere has a number of important consequences associated with the induced ionization and with changes in atmospheric chemistry. For example, it has been demonstrated that electron precipitation can affect radio wave propagation (Helliwell et al., 1973, Imhof et al., 1978, Voss et al., 1984, Goldberg et al., 1987) over a wide range of frequencies, and can also affect ozone concentrations (Brasseur and Solomon, 1984). All of these effects are dependent upon the flux and energy spectrum of the precipitating electrons as well as their location and spatial extent.

Most of our existing knowledge of energetic electron precipitation is based on data acquired with electron spectrometers placed on earth orbiting satellites. Such measurements have provided detailed information on the fluxes and energy spectra, but our knowledge of the spatial extent of electron precipitation events is based primarily on statistical analyses of many separate measurements from satellites. Remote imaging can provide detailed information on the spatial extent of any given event. The imaging can be performed by mapping any one of several quantities produced when electrons enter the atmosphere, including the induced ionization. Radar can provide maps of the ionization (Robinson and Vondrak, 1985), but ground-based facilities can cover only a limited area in any one set of measurements. Fine resolution images of the optical emissions can be obtained from satellites, but images need to be corrected for albedo effects and do not provide multiparameter information on the energy spectra at high energies. Ultraviolet emissions have similar limitations in regard to spectral derivations. When electrons enter the atmosphere and are slowed down, bremsstrahlung x-rays are produced that can be mapped from balloons, rockets, or satellites. Multi-point electron energy spectra and absolute fluxes can be obtained from a satellite using bremsstrahlung x-rays (e.g. Imhof, 1981, Mizera et al., 1984, Datlowe et al., 1988). The use of this technique from satellites is relatively new and x-ray imaging is yet to be used extensively for developing models of electron precipitation. Here, we address the spatial distribution of energetic electrons into the atmosphere. In particular, we shall consider the degree of localization of electron precipitation and also the shapes and spatial extents of the precipitation regions.

PRESENTATION OF DATA

Significant precipitation of high energy electrons into the atmosphere occurs beneath the earth's radiation belts, the most intense precipitation being from the outer belt. The location of the inner and outer radiation belts with respect to the earth is shown schematically in Figure 1. The outer belt extends to an equatorial distance from the center of the earth of approximately 6 earth radii, the outer limit being variable and dependent upon a variety of conditions including the magnetic activity. Between the inner and outer radiation belts is the so-called "slot region", a local minimum in which both the intensities and energy spectra of electrons and protons are quite variable.

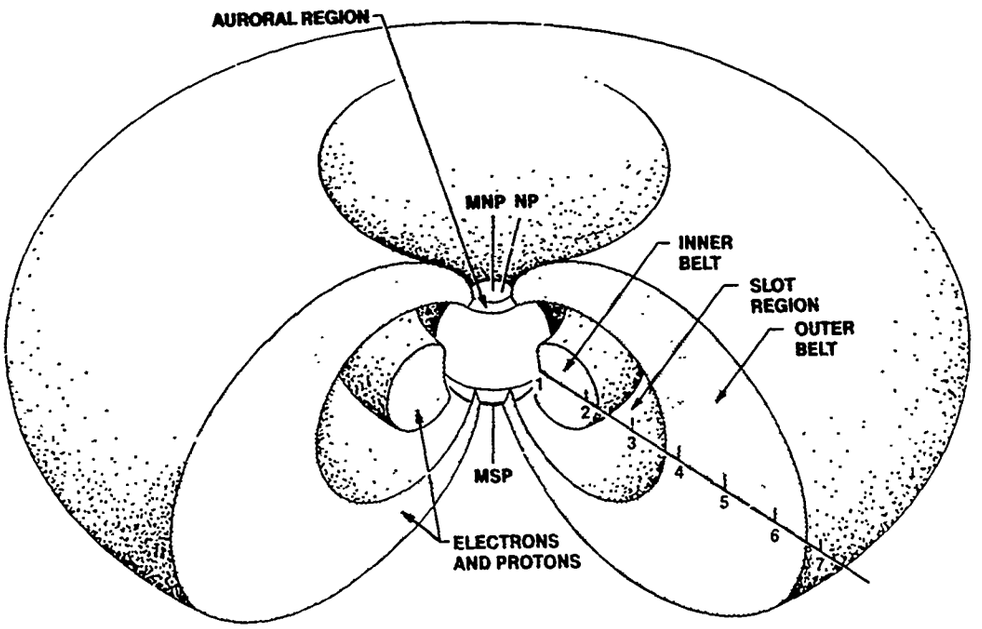


Figure 1
Schematic drawing of the earth's radiation belts.

A spike of precipitating electrons is often observed at the outer edge of the outer radiation belt. Fritz (1970) reported that near midnight the frequency of occurrence is 90 percent. Although electron precipitation commonly occurs in this location, the greatest energy inputs to the atmosphere tend to occur at the less frequently observed spikes considerably equatorward of the outer boundary of the radiation belt.

Measurements of electron precipitation can be improved considerably with the use of remote imaging. Because energetic electrons can produce intense fluxes of bremsstrahlung x-rays, x-ray imaging is an excellent way to remotely sense this type of precipitation. This technique has been used on several satellites. One of the more successful imagers was flown as part of the SEEP payload on the polar-orbiting S81-1 satellite at an altitude of 170 to 270 km. The geometry for the measurements is illustrated schematically in Figure 2. X-rays were mapped in 8 positions to the right and 8 to the left of the satellite path.

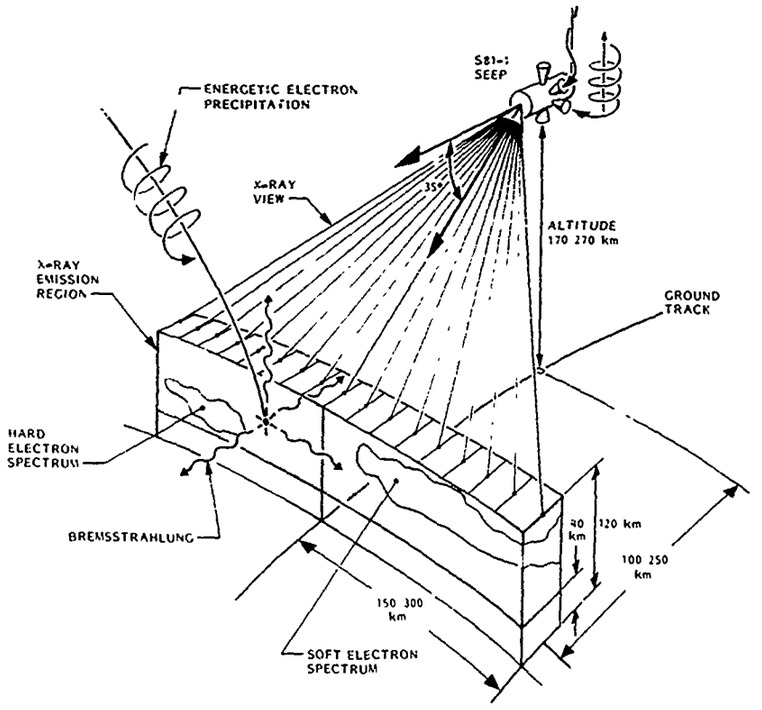


Figure 2
Schematic drawing of the geometry for mapping bremsstrahlung x-rays from the S81-satellite.

Mappings of the electron precipitation near the outer edge of the radiation belt, as obtained with the x-ray imager in the SEEP payload, are shown in Figure 3. Superimposed on the x-ray maps are L-shell contours. As observed frequently at this position in space, the electron precipitation occurred in the shape of an arc at a constant magnetic latitude. With x-ray mappings in longitude performed from the P78-1 satellite it was found that the precipitation arcs near the outer edge of the radiation belts often extend over at least 45° in longitude (Imhof et al., 1990). The widespread arc patterns of the precipitation have important implications for assessing the atmospheric effects. From direct electron measurements it is known that intense precipitation spikes occur well within the outer radiation belt, but they were seldom observed with the SEEP x-ray spectrometer because of the instrument backgrounds associated with the trapped electron population.

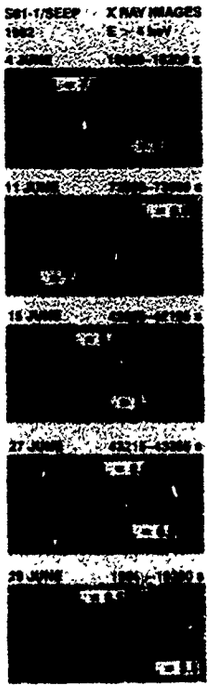


Figure 3

Image of the bremsstrahlung x-rays associated with the electron precipitation near the outer edge of the radiation belt in five cases.

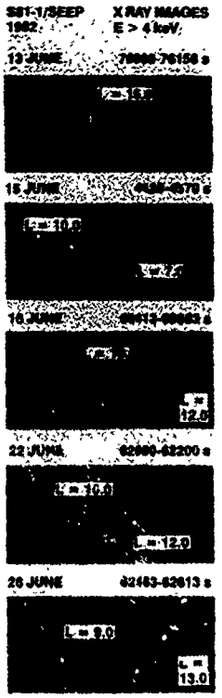


Figure 4

Examples of the image of the bremsstrahlung x-rays associated with electron precipitation poleward of the outer radiation belt.

Let us now consider electron precipitation poleward of the outer radiation belt. This precipitation has been mapped in x-rays and as shown in Figure 4 is often in the form of an arc that is parallel to curves of constant L value. At these high latitudes the energy spectra of the precipitating electrons are generally softer than for those that are precipitating from the radiation belts. Small isolated patches of energetic electron precipitation have been observed in x-rays at high latitudes (Imhof et al., 1985). The patches have typical widths of ~200 km and occur at invariant latitudes between 65° and 83°. They are predominantly present at nighttime, but a few occur in early afternoon hours. In the vicinity of some of the patches additional weaker patches sometimes occur. A few of the precipitation regions are significantly narrower in one direction, consistent with arc-shaped precipitation. It has been concluded that some of the small patches may be parts of larger structures (Imhof et al., 1988).

In addition to arcs and localized patches, energetic electron precipitation occurs in a wide variety of other forms. The patterns sometimes change drastically in a matter of a few minutes. Occasionally, pairs of arcs are observed nearly simultaneously, as shown in Figure 5. Both of these arcs are slightly poleward of the trapped electron population.

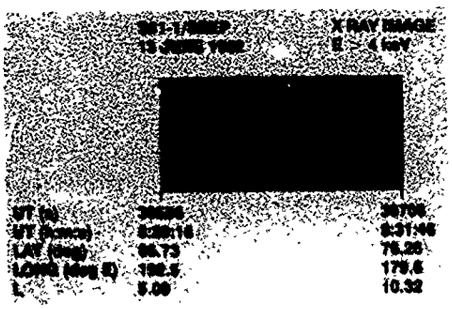


Figure 5

Example of a pair of x-ray arcs observed at a position slightly poleward of the outer radiation belt.

Some electron precipitation at latitudes poleward of the outer radiation belt is localized in position as illustrated by the x-ray image in Figure 6. The x-ray intensity at the center of the small spot observed at 8201S on June 12, 1982 is among the strongest observed on SEEP. The corresponding fluxes of precipitating electrons are therefore quite high. Contour maps of the ion production rate at altitudes of 126, 105, and 90 km calculated from the precipitating electron spectra derived from the x-ray measurements are also shown in the figure.

The calculated ion production rate profile in the atmosphere at the peak region during the event of June 12, 1982 is shown by the solid line in Figure 7. At altitudes below ~80 km the ion production rate is enhanced significantly by the bremsstrahlung x-rays produced by the electrons stopping at higher altitudes. The energy input to the atmosphere for the most intense position is greater and peaks at a higher altitude than most of those in the outer radiation belt. This point is illustrated by the inclusion in the figure of the energy deposition profile for a representative electron precipitation spike at the outer edge of the outer radiation belt (dashed line). When comparing the ion production rate profiles for the outer belt and higher latitude cases it should be remembered that precipitation in the outer belt is often in the form of an arc extending many degrees in longitude, whereas in this case the high latitude precipitation is quite localized in all directions. Therefore, the total energy inputs to the atmosphere for the high latitude patch may not be as much greater than typical lower latitude events as it might at first appear.

Ionization is also produced in the atmosphere by galactic cosmic rays, which are present at all longitudes and latitudes. Some variation occurs with magnetic latitude and with phase of the solar cycle. The ion production rate profile due to cosmic rays is shown in Figure 7 at a magnetic latitude of 65^{o} and during solar minimum. It is apparent that particle precipitation can produce energy deposition at altitudes as low as ~30 km that is larger than the background produced by cosmic rays.

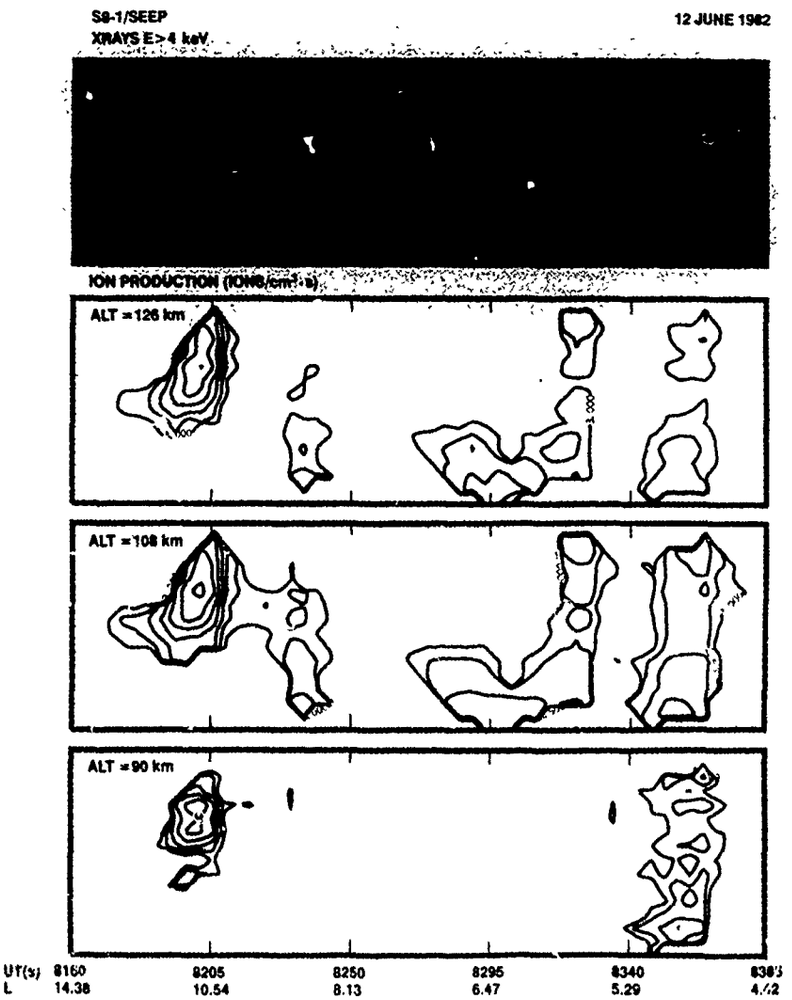


Figure 6

Image of the bremsstrahlung x-rays associated with a precipitation event on June 12, 1982 and contour maps of the ion production rate at selected altitudes of 126 km, 108 km, and 90 km. The minimum contour level is 10^3 pairs/cm^{2}s and there are two contour levels per decade. The maximum ionization is 2 x 10^6 ion pairs/cm^{2}s 108 km observed at 8201s.

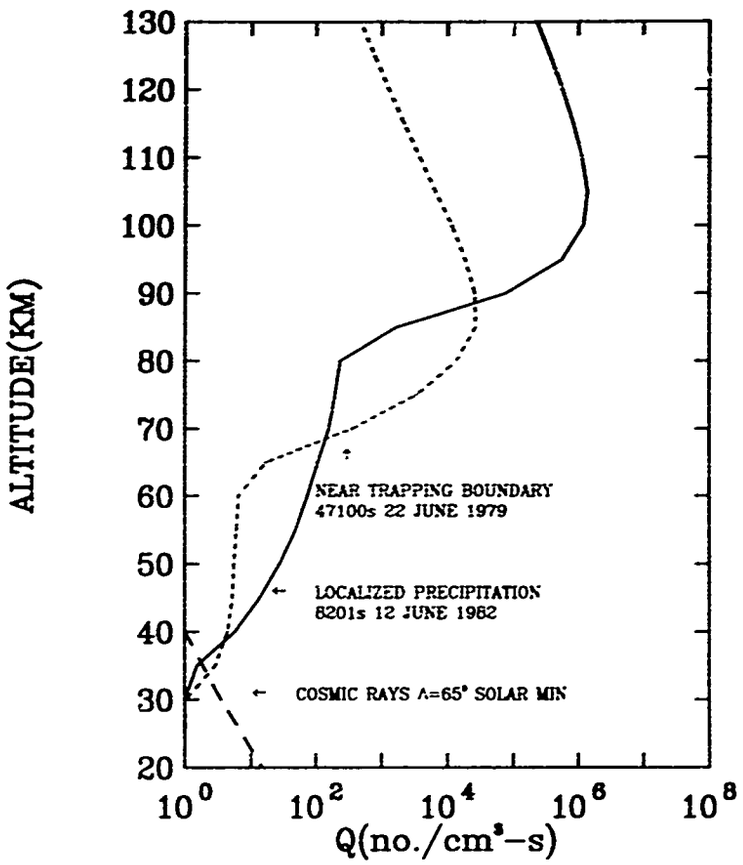


Figure 7

The ion production rate profile at the position of maximum precipitation during the event on June 12, 1982. Also shown are the ion production rate profiles from representative electron precipitation near the boundary of the outer radiation belt and from cosmic rays.

SUMMARY AND CONCLUSIONS

Many examples have been shown of precipitation regions which are either very localized or if extended are shaped in the form of an arc, often parallel to L-shell contours. Even during times of strong precipitation intense electron precipitation is present over only a small fraction of the earth. These patterns are to be contrasted with the inputs from cosmic rays which cover the entire earth. However, in the regions of strong electron precipitation the induced ionization at altitudes above ~40 km is often far greater than that from cosmic rays at the same position. Even at altitudes as low as ~30 km the ionization associated with bremsstrahlung x-rays produced by localized electron precipitation may exceed the ionization associated with cosmic rays.

The examples shown have demonstrated that the electron precipitation often occurs in well-defined patterns. This emphasizes the importance of two-dimensional imaging in monitoring energy inputs from precipitating electrons. Also since the precipitation at a given location can vary rapidly with time, it is important to monitor continually the electron inputs to the atmosphere.

ACKNOWLEDGEMENTS

The SEEP experiment on the S81-1 satellite was sponsored by the Office of Naval Research (contract N00014-79-C-0824). Much of the data analysis presented here was sponsored by the Air Force Office of Scientific Research (contract F-49620-88-C-0072) and by the Office of Naval Research (contract N00014-88-C-0033). The Lockheed Independent Research program provided partial support for the analysis. The efforts of J. P. McGlennon in data reduction are gratefully acknowledged.

REFERENCES

Brasseur, G., and S. Solomon, Aeronomy Of The Middle Atmosphere, 391, D. Reidel, Bingham, MA, 1984.

Datlowe, D.W., W.L. Imhof, and H.D. Voss, X ray spectral images of energetic electrons precipitating in the auroral zone, J. Geophy. Res., 93, 8662, 1988.

Fritz, T.A., Study of the high-latitude outer zone boundary region for >40 keV electrons with satellite INJUN 3, J. Geophys. Res., 75, 538, 1970.

Goldberg, R. A., S. A. Curtis, and J. R. Barcus, Detailed spectra structure of magnetospheric electron bursts precipitated by lightning, J. Geophys. Res., 92, 2505, 1987.

Helliwell, R.A., J. P. Katsufrakis, and M. L. Trimpi, Whistler-induced amplitude perturbation on VLF propagation, J. Geophys., Res. , 78, 4679, 1973.

Imhof, W. L., J. B. Reagan, E. E. Gaines, T. R. Larsen, J. R. Davis, and W. Moler, Coordinated measurements of VLF transmission anomalies and the precipitation of energetic particles into the ionosphere, Radio Science, 13, 717, 1978.

Imhof, W. L., Review of energetic (>20 keV) bremsstrahlung x-ray measurements from satellites, Space Science Reviews, 29, 201, 1981.

Imhof, W. L., E. D. Voss, D.W. Datlowe, and J. Mobilia, Bremsstrahlung x ray images of isolated electron patches at high latitudes, J. Geophys. Res. , 92, 6515, 1985.

Imhof, W.L., H.D. Voss, D.W. Datlowe, and J. Mobilia, Isolated electron precipitation regions at high latitudes, J. Geophys. Res., 93, 2649, 1988.

Imhof, W. L., J. Mobilia, D. W. Datlowe, H. D. Voss, and E. E. Gaines, The longitudle and temporal variations of energetic electron precipitation near the trapping boundary, J. Geophys. Res., (in press), 1990.

Mizera, P.F., D.J. Gorney, and J.L. Roeder, Auroral x-ray images from DMSP-F6, Geophys. Res. Lett., 11, 255, 1984.

Robinson, R.M., and R.R. Vondrak, Characteristics and sources of ionization in the continuous aurora, Radio Science , 20, 447, 1985.

Voss, H. D., W. L. Imhof, J. Mobilia, E. E. Gaines, M. Walt, U. S. Inan, R. A. Helliwell, D. L. Carpenter, J. P. Katsufrakis, and H. C. Chang, Lightning-induced electron precipitation , Nature , 312, 740, 1984.

AD-P006 274

IRREGULARITY ANISOTROPY AND DRIFTS FROM POLAR CAP SCINTILLATIONS DURING THE CURRENT SOLAR MAXIMUM

Sunanda Basu
Boston College
Institute for Space Research
Newton, MA 02159

S. Basu
GL(AFSC)/LIS
Hanscom AFB, MA 01731

Emanoel Costa
Centro de Estudos em Telecommunicacoes
Pontifica Universidade Catolica do Rio de Janeiro
Brazil

91-09967

ABSTRACT

Spaced receiver scintillation measurements made at Thule, Greenland during two sunspot maximum CEDAR High Latitude Plasma Structure (HLPS) campaigns in Dec 1988 and Nov 1989 have been analyzed to yield phase and amplitude scintillation magnitude, decorrelation time, irregularity drifts and anisotropy of the ground diffraction pattern during well defined polar cap patch and sun-aligned arc events. It is found that all these parameters tend to be different in these two classes of events which are controlled by the B_z component of the interplanetary magnetic field.

INTRODUCTION

The current solar maximum has given us a unique opportunity to study plasma structure and dynamics deep within the dark polar cap by utilizing coordinated optical and radio techniques. The on-going NSF initiative in the geo-sciences known as CEDAR, acronym for Coupling, Energetics and Dynamics of Atmospheric Regions, has provided an incentive for several research groups to get together during designated "campaigns" to make simultaneous measurements to determine different sources of High Latitude Plasma Structure (HLPS).

Hitherto, four such HLPS campaigns have been conducted: in Feb 88, Dec 88, Oct-Nov 89, and Feb 90. The first campaign was conducted when the sunspot cycle was low (sunspot number 40), the second and third when the sunspot number was near the peak (approximately 180) and the last one is so recent that not much supporting data is yet available. Careful analysis of the multitechnique data has just begun (Basu et al., 1989; 1990a). However, what has become very evident is the extent to which plasma structuring is controlled by the north-south component B_z of the interplanetary magnetic field. For instance, it is well established that when B_z is negative, large (~ 1000 km) regions of convecting density enhancement known as patches are observed, whereas, when IMF B_z is positive, sun-aligned arcs are found to populate the polar cap (Weber et al., 1984; Buchau et al., 1985). That both patches and arcs are associated with small-scale density and electric field turbulence (~ kms to m) has been determined by a variety of in-situ and remote sensing techniques (Carlson et al., 1988a). Further, using spaced-receiver scintillation measurements during the high sunspot conditions prevailing during the second and third campaigns, we shall show in this paper that the drift, anisotropy and intensity of the several km to hundreds of m irregularities are also indirectly controlled by B_z. The IMF control comes through its effects on the magnetospheric convection pattern and, possibly, also through the specific instability mechanisms which create the plasma structuring for each orientation of B_z.

91

RESULTS

Spaced-receiver scintillation measurements have been conducted at Thule, Greenland using quasi-geostationary satellite transmissions at 250 MHz for several years. One day's observations of irregularity drifts in conjunction with optical observations of patches have been published (Weber et al., 1984). Interpretation of similar measurements conducted at Goose Bay, Labrador has been discussed in detail by Costa et al. (1988). We follow exactly the same technique for the estimation of the parameters of the cross-correlation analysis of the spaced-receiver data obtained at Thule. The receivers were located at the three corners of a right-angled isosceles triangle with smaller sides of 170 m. The received signals were sampled at 50 Hz. The satellite frequency was updated every 168 s, thus creating a loss of lock during a fraction of this time, in the beginning. As a result, and also due to some preprocessing of the received information, data blocks of 72-s length (corresponding to 3600 observations of the signal intensity per channel) approximately centered at the middle of each cycle were used in the correlation analysis. A more detailed description of the receiving system and preprocessing of the data was presented by Basu et al. (1985).

To study the variation in irregularity drift patterns, anisotropy and magnitude of scintillations we choose specific polar cap patch and sun-aligned arc events observed during the second and third HLPS campaigns. In particular, to contrast the behavior of the above mentioned characteristics, we choose one patch and one arc event from the Dec 88 campaign and a similar set from Nov 89. On all days the zenith angle of the observations were high varying between 30-35°. The identification of patches and arcs were made by using supporting information from optical and other radio techniques. Here we will present only the information obtained from the spaced-receiver scintillation measurements.

The amplitude and phase scintillations, S_4 and σ_ϕ, and the decorrelation time, τ, for 50 percent decorrelation for the patch event of Dec 6, 1988 are shown in Figure 1. The axial ratios and pattern velocities obtained from the spaced-receiver analysis are shown in Figures 2 and 3. The results shown are those obtained from the modified-Fedor (MF) algorithm based on a paper by Fedor (1967) and discussed by Costa et al. (1988). The results obtained from the Rino-Livingston algorithm based on work done by Rino and Livingston (1982) and Armstrong and Coles (1972) yields similar results and are not shown to minimize the number of diagrams. Figure 1 shows that the polar cap patch causes fairly intense scintillations between 0018-0136 UT with S_4 = 0.8 in conjunction with phase scintillations exceeding 10 radians (for 82-sec detrend intervals). The other interesting feature of the data is that the decorrelation interval is fairly constant with time varying at most by a factor of 2 from 0.2 to 0.4 sec.

The anisotropy of the diffraction pattern observed in association with patches is shown plotted on a circle in Figure 2 covering corrected geomagnetic latitudes (CGL) from 82° to 90° and corrected geomagnetic local times (CGLT). Axial ratios are represented by the length of a line segment centered at the vertical projection on the ground of the 350-km ionospheric penetration point of the ray path. On the upper right-hand corner of the figure, a line segment corresponding to an axial ratio of 20 is shown. The orientations of the major axes of the anisotropy ellipses are measured by the angle between the radial direction through each penetration point (ind the local northward direction in the polar CGL x CGLT coordinate system) and the correspo ne segment (aligned with the major axis of the ellipse).

The drift velocities of the diffraction pattern measured simultaneously are displayed in Figure 3 using identical polar CGL versus CGLT coordinate system. The magnitude of the velocity vector is represented by the length of a line segment starting at the vertical projection on the ground of the 350-km ionospheric penetration point. The scale for the magnitude is shown on the upper right-hand corner. The orientation of the velocity is measured from north to south in this coordinate system, designated positive in the clockwise direction and negative in the anticlockwise direction. Preliminary comparisons show that the drift velocity of the diffraction pattern is consistent with irregularity drift measurements using other techniques (Costa et al., 1988).

Figures 2 and 3 taken together show the drift velocity to be primarily antisunward with the orientation of the major axes of the anisotropy ellipses being in a direction perpendicular to this antisunward motion. The anisotropy is quite large, many exceeding 10 in magnitude. The drift velocities are also fairly large, most of them exceeding 0.5 km s^{-1}. These drift velocities are consistent with earlier measurements of the dynamics of polar cap patches by different techniques (Weber et al., 1984, 1986; Buchau et al., 1988).

We present scintillation magnitude, decorrelation time, pattern anisotropy and drifts for a second patch event obtained during the third HLPS campaign on Nov 5, 1989, between 2120-0040 UT on Nov 6. Intense scintillations with S_4 = 1 are observed in Figure 4; the decorrelation time being similar to the values obtained earlier except for a few samples which have somewhat larger decorrelation times. Significant 1.2 GHz scintillation is also observed simultaneously using the GPS satellites. (G.J. Bishop, private communication, 1990). Unfortunately, no phase scintillation measurements were available at this time. Figures 5 and 6 show the axial ratio and drift velocity on CGL vs. CGLT coordinate system for

this second patch event. The axial ratios are similar to those seen in Figure 2 while the drift velocities show a westward tilt as compared to an eastward tilt seen in Figure 3.

Scintillation observed during the sun-aligned arc event observed on Dec 12, 1988 at Thule between 1002-1126 UT is shown in Figure 7. The phase and amplitude scintillation is much weaker than that seen in association with patches and the decorrelation interval much more variable. The one large phase scintillation value is obtained in conjunction with a shear in the flow velocity to be presented in Figure 9. The axial ratios shown in Figure 8 are somewhat smaller in the case of arcs as compared to those of patches. The velocity in Figure 9 also shows a reversal from antisunward up to 07 CGLT to sunward beyond that. Such sunward flows during northward B_z conditions were first seen by Burke et al. (1979) and have been confirmed by using combined in-situ and optical techniques by Carlson et al. (1988b).

The scintillation magnitude caused by the second sun-aligned arc event seen on Nov 4, 1989 between 0200-0440 UT is also rather moderate as shown in Figure 10. The axial ratios presented in Figure 11 are small as in Figure 9. The drift velocities themselves shown in Figure 12 have both sunward and antisunward directions of flow thus giving rise to shears in the flow velocity.

As a result of this study we arrive at the preliminary conclusion that intensity of irregularities of scale sizes several km to several hundred m, their flow patterns and the axial ratios of the diffraction pattern are all different for polar cap patches when compared to sun-aligned arc events. Thus we may state that IMF B_z indirectly controls the behavior of even these small-scale irregularities. From a study of several other weaker arc events, it seems probable that such differences between patches and arcs emerge when the particle precipitation causing the arc events are energetic enough to create conductivity structures in the E-region and associated electric field structures. Further careful study is necessary to determine anisotropy and drifts associated with F-region sun-aligned arcs.

It is fairly well established that the gradient drift instability gives rise to structuring in patches (Tsunoda, 1988; Basu et al., 1990b). When taken in conjunction with the results of this study, one can postulate that such irregularities cause diffraction patterns with large axial ratios. On the other hand, velocity shears perpendicular to the magnetic field seem to generate irregularities which create diffraction patterns with somewhat smaller axial ratios. Basu et al. (1988) had earlier discussed irregularity generation in the vicinity of velocity shears at the poleward edge of the auroral oval. It is probable that similar mechanisms can be invoked for intense sun-aligned polar cap arcs. Several different generation mechanisms have been advanced (Keskinen et al., 1988; Basu and Coppi, 1989; Nishikawa et al., 1990). It is hoped that the observed irregularity characteristics described here will provide further impetus to theorists to define the form of irregularities generated by these different mechanisms.

ACKNOWLEDGMENTS

We wish to thank C. Bryant of Boston College and K. Bounar of Radex Inc. for their help with signal analysis and development of computer algorithms. We thank the Danish Commission for Scientific Research in Greenland for permission to conduct ground experiments at Thule Air Base under Project 24-89. The work at AFGL was supported in part by Air Force Office of Scientific Research task 2310G9. The work at Boston College was supported by AFGL contract F19628-90-K-0007 and NSF Grant ATM-8715445.

REFERENCES

Armstrong, J.W. and W.A. Cole (1972), Analysis of three-station interplanetary scintillation, *J. Geophys. Res.*, *77*, 4602.

Basu, B. and B. Coppi (1989), Velocity shear and fluctuations in the auroral regions of the ionosphere, *J. Geophys. Res.*, *94*, 5316.

Basu, Su., S. Basu, E. MacKenzie, and H.E. Whitney (1985), Morphology of phase and intensity scintillations in the auroral oval and polar cap, *Radio Sci.*, *20*, 347.

Basu, S., S. Basu, E. MacKenzie, P.F. Fougere, W.R. Coley, N.C. Maynard, J.D. Winningham, M. Sugiura, W.B. Hanson, and W.R. Hoegy (1988), Simultaneous density and electric field fluctuation spectra associated with velocity shears in the auroral oval, *J. Geophys. Res.*, *93*, 115.

Basu, S. S. Basu, C.E. Valladares, E.J. Weber, J. Buchau, G.J. Bishop, and B.W. Reinisch (1989), Coordinated observations of high latitude ionospheric turbulence, *SPI Conference Proceedings and Reprint Series*, *8*, 137.

Basu, S., S. Basu, E.J. Weber, and G.J. Bishop (1990a), Plasma structuring in the polar cap, *J. Geomag. Geoelec., Japan* (in press).

Basu, S., S. Basu, E. MacKenzie, W.R. Coley, J.R. Sharber, and W.R. Hoegy (1990b), Plasma structuring by the gradient-drift instability at high latitudes and comparison with velocity-shear driven processes, *J. Geophys. Res., 95* (in press).

Buchau, J., E.J. Weber, D.N. Anderson, H.C. Carlson, Jr., J.G. Moore, B.W. Reinisch, and R.C. Livingston (1985), Ionospheric structures in the polar cap: Their origin and relation to 250-MHz scintillation, *Radio Sci., 20*, 325.

Buchau, J., B.W. Reinisch, D.N. Anderson, E.J. Weber, and C. Dozois (1988), Polar cap plasma convection measurements and their relevance to the modeling of the high-latitude ionosphere, *Radio Sci., 23*, 521.

Burke, W.J., M.C. Kelley, R.C. Sagalyn, M. Smiddy, and S.T. Lai (1979), Polar cap electric field structures with a northward interplanetary magnetic field, *Geophys. Res. Lett., 6*, 21.

Carlson, H.C., Jr., E.J. Weber, L.P. Block, and S. Basu (1988a) Satellite, airborne, and radar observations of auroral arcs, *Proceedings of the Symposium on Multipoint Measurements of Magnetospheric Processes, COSPAR, 8*, 9(49).

Carlson, H.C., R.A. Heelis, E.J. Weber, and J.R. Sharber (1988b), Coherent mesoscale convection pattern during northward interplanetary magnetic field, *J. Geophys. Res., 93*, 14501.

Costa, E., P.F. Fougere, and S. Basu (1988), Cross-correlation analysis and interpretation of spaced-receiver measurements, *Radio Sci., 23*, 141.

Fedor, L.S. (1967), A statistical approach to the determination of three-dimensional ionospheric drifts, *J. Geophys. Res., 72*, 5401.

Keskinen, M.J., H.G. Mitchell, J.A. Fedder, P. Satyanarayana, S.T. Zalesak, and J.D. Huba (1988), Nonlinear evolution of the Kelvin-Helmholtz instability in the high-latitude ionosphere, *J. Geophys. Res., 93*, 137.

Nishikawa, K.-I., G. Ganguli, Y.C. Lee, and P.J. Palmadesso (1990), Simulation of electrostatic turbulence due to sheared flows parallel and transverse to the magnetic field, *J. Geophys. Res., 95*, 1029.

Rino, C.L. and R.C. Livingston (1982), On the analysis and interpretation of spaced-receiver measurements of transionospheric radio waves, *Radio Sci., 17*, 845-854.

Tsunoda, R.T. (1988), High-latitude F-region irregularities: a review and synthesis, *Revs. Geophys., 26*, 719.

Weber, E.J., J. Buchau, J.G. Moore, J.R. Sharber, R.C. Livingston, J.D. Winningham, and B.W. Reinisch (1984), F layer ionization patches in the polar cap, *J. Geophys. Res., 89*, 1683.

Weber, E.J., J.A. Klobuchar, J. Buchau, H.C. Carlson, Jr., R.C. Livingston, O. de la Beaujardiere, M. McCready, J.G. Moore, and G.J. Bishop (1986), Polar cap F layer patches: structure and dynamics, *J. Geophys. Res., 91*, 12121.

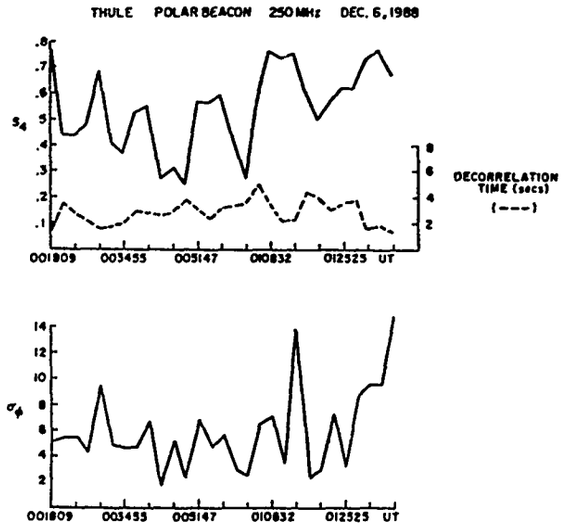


FIGURE 1. Amplitude scintillations (S_4), phase scintillation in radians (σ_ϕ) and decorrelation time at 250 MHz observed at Thule, Greenland during a polar cap patch event on Dec 6, 1988.

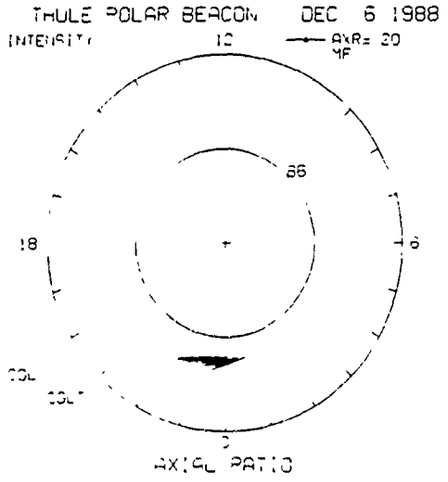


FIGURE 2. Axial ratios and orientations of the major axes of the anisotropy ellipses of the diffraction pattern for the data shown in Figure 1.

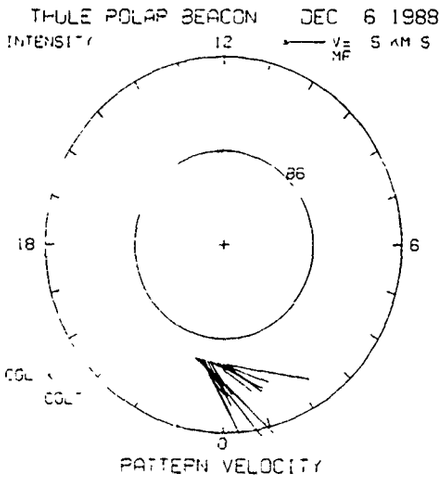


FIGURE 3. Magnitude and directions of the true drift velocity of the diffraction pattern for the data shown in Figure 1.

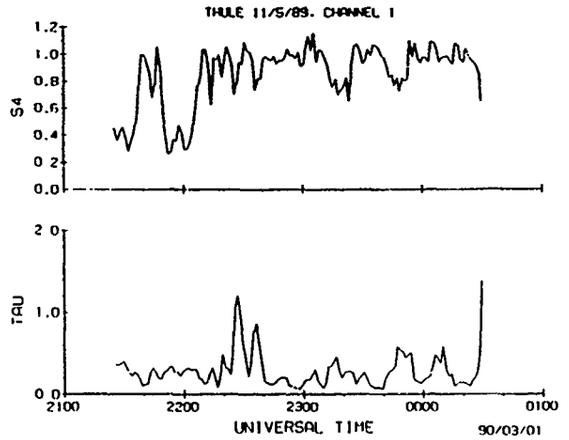


FIGURE 4. Same as in Figure 1 except for the polar cap patch event of Nov 5, 1989.

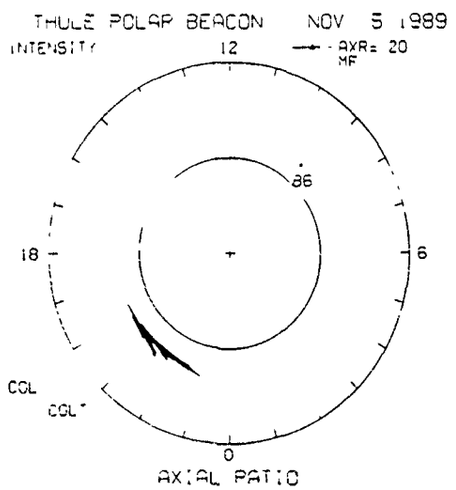


FIGURE 5. Same as in Figure 2 except for the data shown in Figure 4.

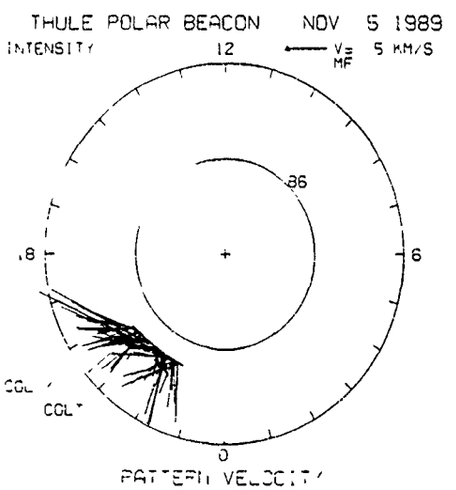


FIGURE 6. Same as in Figure 3 except for the data shown in Figure 4.

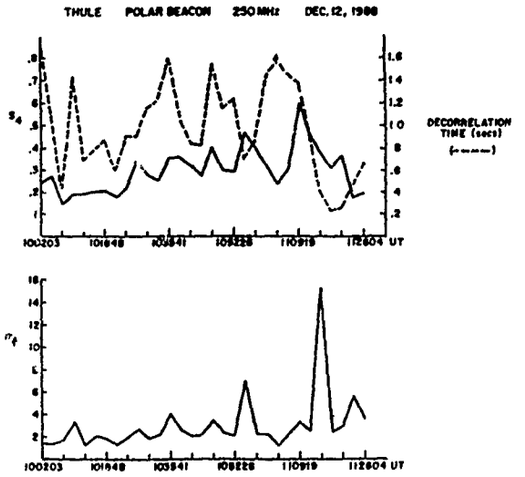


FIGURE 7. Amplitude scintillations (S_4), phase scintillation in radians (σ_ϕ) and decorrelation time at 250 MHz observed at Thule, Greenland during a sun-aligned polar cap arc event on Dec 12, 1988.

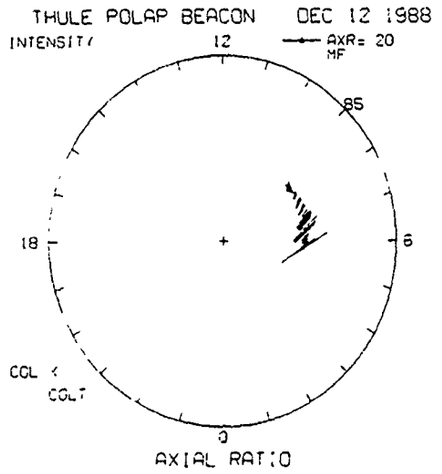


FIGURE 8. Same as in Figure 2 except for the data shown in Figure 7.

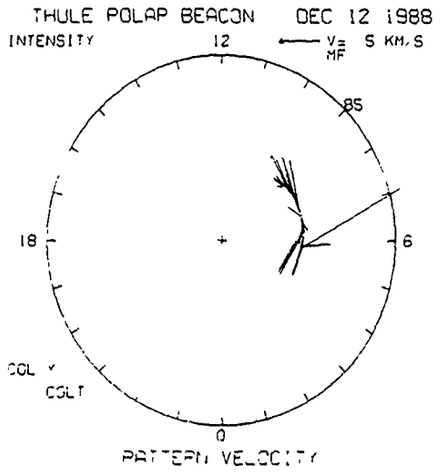


FIGURE 9. Same as in Figure 3 except for the data shown in Figure 7.

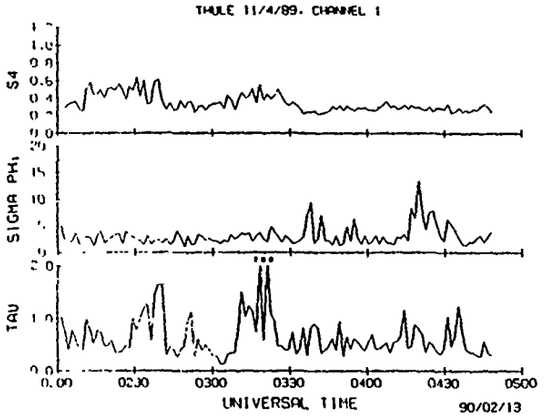


FIGURE 10. Same as in Figure 7 except for the sun-aligned polar cap arc event of Nov 4, 1989.

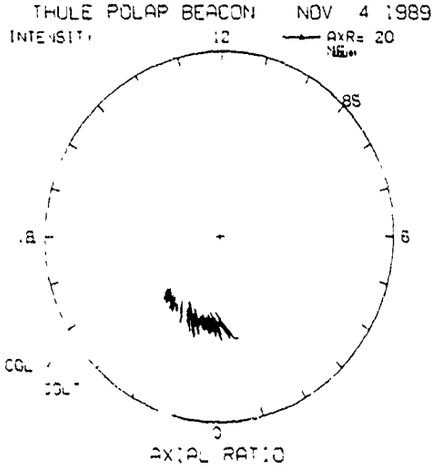


FIGURE 11. Same as in Figure 2 except for the data shown in Figure 10.

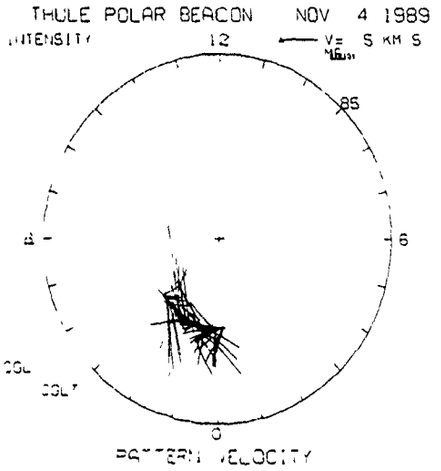


FIGURE 12. Same as in Figure 3 except for the data shown in Figure 10.

AD-P006 275

GOOSE BAY RADAR OBSERVATIONS OF EARTH-REFLECTED ATMOSPHERIC GRAVITY WAVES IN THE HIGH-LATITUDE IONOSPHERE

91-09668

J. M. Ruohoniemi, R. A. Greenwald, K. B. Baker
Applied Physics Laboratory
The Johns Hopkins University
Laurel, Maryland 20723

J. C. Samson
The Department of Physics
University of Alberta
Edmonton, Alberta
Canada

ABSTRACT

The Applied Physics Laboratory of the Johns Hopkins University operates an HF backscatter radar at Goose Bay, Labrador. With this instrument it is possible to observe irregularities in the distribution of ionospheric ionization at E and F region altitudes (100 - 600 km) in the high-latitude (65 - 85°A) ionosphere. Recently it has been established that the passage of atmospheric gravity waves perturbs the ionosphere in ways that are readily detected in returns that reflect off the ionospheric layers. The particular strength of the technique lies in the nearly instantaneous measurement of gravity wave effects over large areas (10^6 km^2). With this information the propagation of gravity waves can be accurately modelled. Generally gravity waves are observed during daylight hours propagating away from the auroral electrojets. The propagation mode involves penetration of wave energy through the lower atmosphere and subsequent reflection by the earth's surface. The frequencies associated with the waves lie in the 0.4 - 0.6 mHz range and the wavelengths vary from 300 to 500 km. The excitation sources appear to lie in the vicinity of the high-latitude electrojets. In this paper we outline the analysis of gravity wave effects on HF propagation and present an example of a modelled gravity wave event.

INTRODUCTION

It has long been known that gravity waves in the neutral atmosphere induce wavelike fluctuations in the layers of the ionosphere [Hines, 1960]. Before their association with gravity waves was established, the ionospheric manifestations were called Travelling Ionospheric Disturbances, or TID's, a term that reflects the dynamical nature of the underlying phenomenon. In this paper we will not distinguish between gravity waves in their neutral atmosphere and ionospheric aspects.

Gravity waves are generally grouped into two broad categories. Large-scale waves (LSWs) have horizontal phase velocities greater than the sound velocity in the lower atmosphere (~ 300 m/s) and wavelengths of the order of 1000 km, while medium-scale waves (MSWs) have phase velocities less than the sound velocity and wavelengths of several hundred kilometers. The two types are further distinguished by their propagation modes; the MSWs propagate mainly within the lower atmosphere while the propagation paths of LSWs waves lie entirely within the thermosphere. The wave oscillation frequencies for both types lie in the 0.2 - 1.0 mHz range. The energy of gravity waves in the high-latitude regions is believed to derive from particle precipitation and ionospheric currents.

Factors that affect the propagation of gravity waves include reflections from the lower part of the thermosphere and the surface of the earth, neutral winds, and molecular dissipation and ion drag in the thermosphere. Downward propagating waves with horizontal phase velocities greater than about 250 m/s tend to be reflected upward before they reach an altitude of 110 km [Richmond, 1978]. However, those with lower horizontal phase

velocities penetrate to the ground and are reflected there [Francis, 1974]. These earth-reflected waves have sufficient dispersion in frequency at F region altitudes to produce oscillatory wave trains even though the sources themselves are impulsive.

In this presentation we will discuss the propagation of MSWs in the earth-reflected mode. These have been identified and modelled using HF backscatter data from the Goose Bay radar. We will illustrate with a case study that the location and onset time of gravity wave excitation can be determined in individual events.

PROCEDURE

The field of view of the HF radar at Goose Bay, Labrador (53.4°N, 60.4°W) takes in northeastern Canada and Greenland (Figure 1). The facility and its operation have been described in detail by Greenwald et al. [1985]. Briefly, the phased-array antenna repeatedly scans a 50° azimuth sector centered on 5° east of geographic north. Within each scan the beam is swept in a clockwise direction through 16 evenly spaced pointing azimuths. The integration time per beam position is usually 6 s so that a scan is completed every 96 s. The range of the backscatter sources is determined by gating the returns with a range step of 45 km. A total of 50 range gates are sampled. The radar can be operated at any frequency in the 10 - 20 MHz range. The backscattered power, Doppler velocity, and spectral width are computed from estimates of the autocorrelation function. On every fourth scan measurements are made using vertical interferometer techniques that result in estimates of the reflection height.

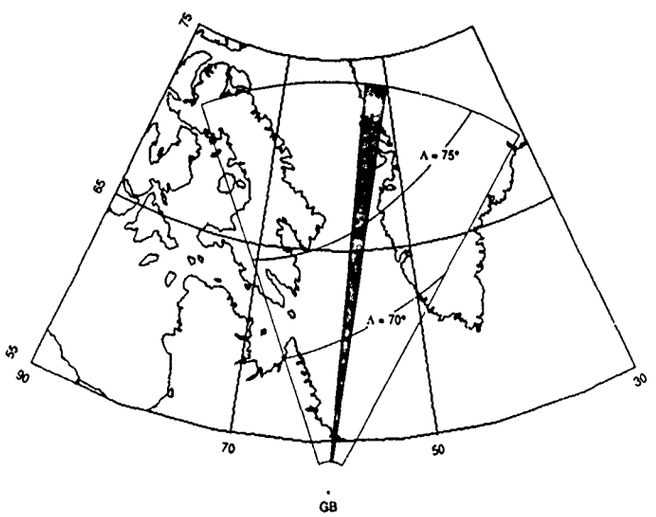


Figure 1. Field of view of the JHU/APL HF radar at Goose Bay, Labrador. The lightly shaded sector represents one of the 16 azimuthal pointing directions of the radar (beam 8, azimuth = 6°) and the darkly shaded inner rectangle indicates the dimensions of a beam-gate sampling cell.

The radar signature of gravity wave activity is seen in the record of ground backscatter returns. These involve oblique reflection of the HF transmissions from the ionosphere followed by backscatter from the earth's surface. Consequently, the ranges to the reflecting points in the ionosphere are less than one-half the measured ranges This also effectively improves the spatial resolution of the gravity wave measurements to ~22 km from the nominal pulse length-determined value of 45 km.

We now consider a sample of Goose Bay data. A form of presentation useful for highlighting gravity wave activity is shown in Figure 2. Here backscattered power is plotted as a function of range and time for just one of the beams of the radar (beam 8). The period is from November 25, 1988 and covers the 12-24 UT interval. Since local time is offset from Universal Time by 4 hours the observations before 20 UT correspond to daylight conditions. Power is shown in dB above noise. Although the details of this plot may be obscure due to the inadequacy of black-and-white reproduction of the color original, it should be clear that quasiperiodic localized intensifications dominate in the 1330-1900 UT

interval. As we shall further discuss, these intensifications are caused by equatorward propagating gravity waves that produce undulations of the bottomside *F* layer. The Doppler velocities associated with the propagating intensity enhancements (*not shown*) have magnitudes which are less than 20 m/s and also exhibit a periodic variation. These velocities result from path length variations brought on by the vertical motions induced by the gravity wave.

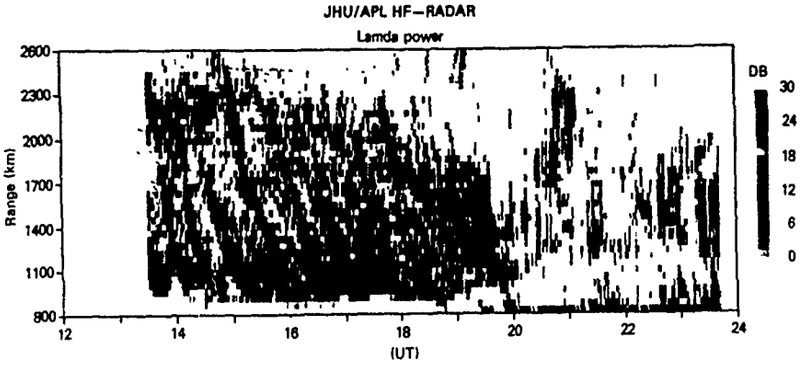


Figure 2. The backscattered power for beam 8 on November 25, 1988. The data are plotted as functions of range and Universal Time (UT).

We turn now to the modelling of HF propagation in an ionosphere perturbed by the passage of an earth-reflected gravity wave. For a full treatment of this problem see Samson et al. [1990]. Briefly, the atmosphere is approximated by a two-layer model with a transition from lower atmosphere to thermosphere at an altitude of 150 km. Downward propagating gravity waves produced by an impulsive source in an auroral electrojet reflect from the earth's surface. This can be modelled as a virtual source at a depth of $-z_s$, where z_s is the source altitude. The reflection coefficient for upward propagating waves at the thermospheric boundary is assumed to be small. Then, as shown by Francis [1974], the various parameters of the gravity wave (pressure, electron density, etc.) have a simple functional form that lends itself to analysis of the wave propagation in terms of constant phase surfaces. Figure 3 shows an example of a map of constant phase surfaces for an earth-reflected gravity wave. The map is adapted to the radar range-time format and represents the data from one beam after mapping to the appropriate reflection points in the ionosphere. The source is assumed to be 2000 km distant from the radar and at an altitude of 130 km. The observation altitude (i.e. the height of the obliquely reflecting ionospheric layer) is assumed to be 250 km.

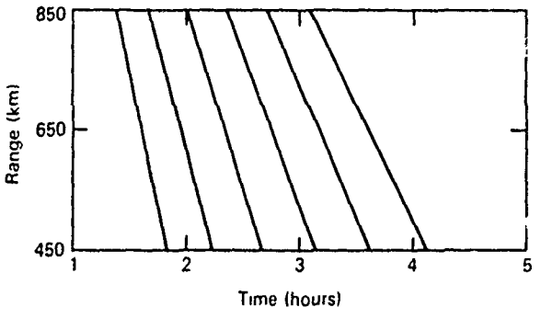


Figure 3. Theoretical map of the surfaces of constant phase for an earth-reflected gravity wave at a range of 2000 km and an altitude of 130 km. The observation altitude is assumed to be 250 km.

Using the theoretical phase map, the range and time of excitation of the gravity wave can be obtained by extending the measured phase surfaces to the position where they intersect on the range-time plane. The same procedure can be followed with experimental data. For these analyses, the wavelength and direction of propagation are estimated by applying a generalized cross-spectral technique [Samson et al., 1990] that incorporates time series data from a spatially distributed subset of beam-gate cells. Once the direction of propagation, range, and time have been determined, the altitude of the source can be estimated by comparing the data with maps of constant phase surface for sources at different altitudes. (As the altitude increases, the spacing between the phase surfaces decreases.) Thus the complete set of wave parameters is obtained.

To confirm that the range-time traces of enhanced backscattered power (Figure 2) do indeed represent the distribution of the constant phase surfaces of the gravity waves (Figure 3), we have computed electron densities for the gravity wave model. The propagation of the HF radar signal through this perturbed ionosphere was then studied by performing a ray-tracing analysis. The results are shown in Figure 4. The left and right hand panels correspond to observations 195 and 210 minutes after excitation, respectively. (Note that, as shown, rays propagate away from the radar site simultaneously over a span of vertical take-off angles. This is due to the relatively broad antenna pattern in the vertical dimension.) The ray paths in the left hand panel show three regions where the rays are focused, at 850 km, 1300 km, and 1900 km ground range. The first is due to the usual focusing near the skip-distance, while the second and third are due to focusing of the rays which are reflected in the ionosphere at ranges and heights corresponding to local minima in the electron density. Rays reflecting from the vicinity of local maxima are defocused instead. When the radar signal is backscattered from the ground, the focused rays show up as enhanced intensities at the corresponding ranges. In essence, these regions of localized intensification correspond (through their association with the density minima in the ionsopheric layers) to surfaces of constant phase of the gravity wave. The right hand panel illustrates the ray path focusing after an additional elapse of 15 minutes. The focusing from the nearer electron density minimum is no longer obvious since it has merged with the skip-distance focusing. The further region is now focused at a ground range of 1600 km and a third region is forming at a range greater than 1600 km. Thus the intensifications are seen to sweep successively through the radar field of view.

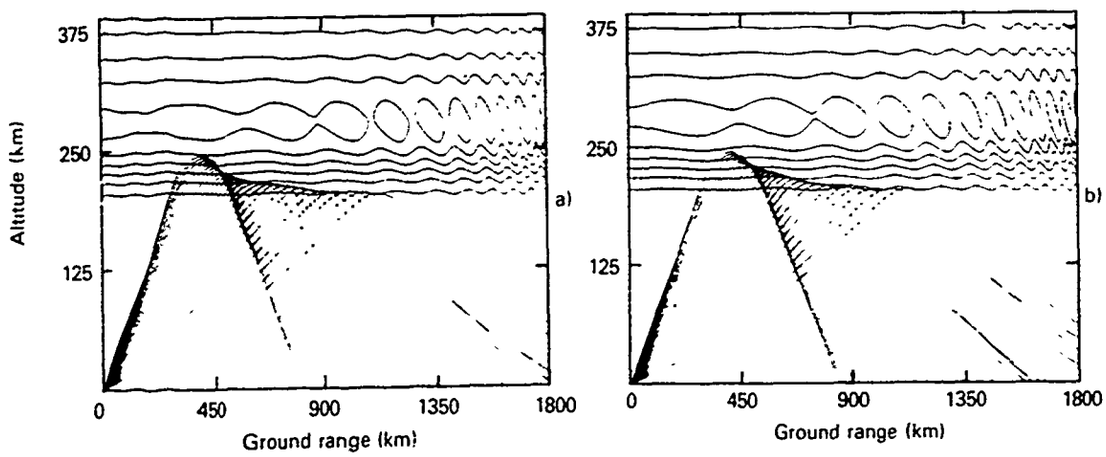


Figure 4. Ray paths of HF radar signal propagation through an ionosphere modulated by the passage of an earth-reflected gravity wave; (a) 195 min after excitation, (b) 210 min after excitation.

CASE STUDY

Gravity wave activity is observed with the Goose Bay radar predominantly in daylight hours when there is sufficient F region ionization to support oblique reflection from the ionosphere. To date, the clearest examples have been seen during late November and early December. Although a number of periods have been extensively analyzed [Samson et al., 1990], for brevity we limit ourselves here to the event of November 25, 1988. A preview of this analysis, complete with color plates, has been given by Samson et al. [1989]

A time series summary of the radar observations in the direction of beam 8 was given in Figure 2. It is informative to isolate the data for a particular range gate. Figure 5 shows the time series of backscattered power and Doppler velocity for the gate that corresponds to ionospheric reflection at a range of 500 km. The data have been detrended to give zero mean. Similiar quasimonochromatic wavetrains are seen, but the velocity appears to lag the power by almost a quarter of a cycle. This phase lag is to be expected, since, with equatorward motion of the gravity wave, the focusing of backscattered power in the vicinity of a density minimum at a given range will precede the peaking of the vertical velocity of the ionospheric layers by this amount. Figure 6 shows the power spectra for power, velocity, and reflection height for the 1330 - 1630 UT interval. All parameters show clear spectral peaks near a frequency of 0.54 mHz (equivalent to a period of 31 min). By application of the cross-spectral techniques described by Samson et al. [1990] it can be demonstrated that there was probably only one source for this event. The wave number spectrum for the backscattered power is shown in Figure 7. The wavelength is about 310 km and the horizontal phase velocity is 180 m/s. The direction of propagation corresponds very nearly to equatorward motion away from the geomagnetic pole.

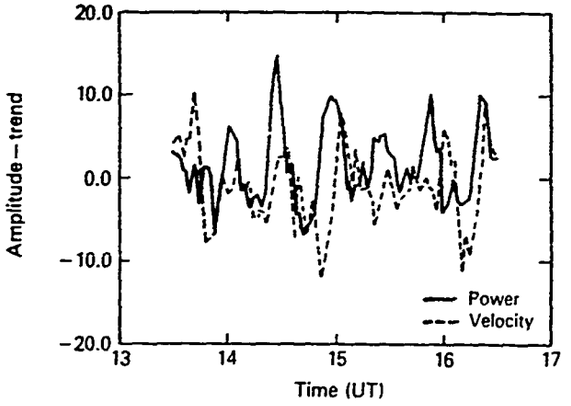


Figure 5. Time series of the backscattered power and Doppler velocity for the event of November 25, 1988 at an ionospheric reflection range of 500 km in the direction of beam 8. The data have been detrended to give a zero mean. The power is in dB and the velocity in m/s.

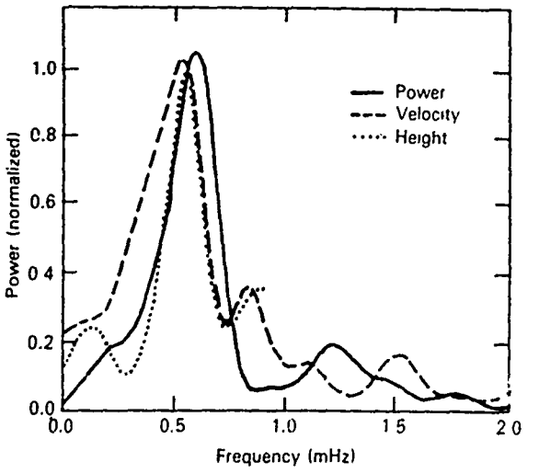


Figure 6. Power spectra of the data for 1330 - 1630 UT on beam 8 at an ionospheric reflection range of 500 km.

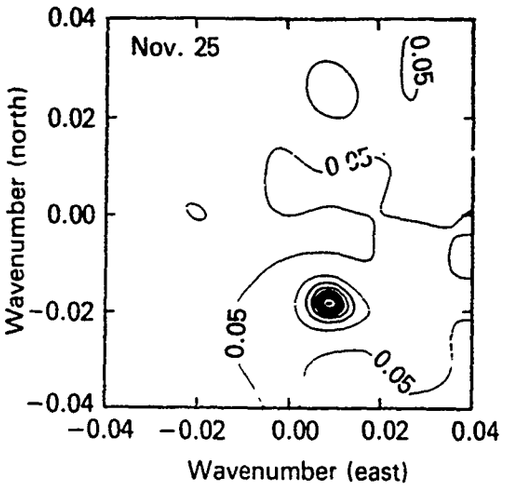


Figure 7. The wave number spectrum for time series of the backscattered power in the 1330 - 1630 UT interval and over the spectral range 0.25 - 0.75 mHz. The wave numbers are expressed in geographic coordinates (east and north). The contours are at intervals of 0.10 unit.

With this information, the experimental phase data can be modelled by the procedure outlined in the previous section. The result is shown in Figure 8. The plot of backscattered power differs from that given in Figure 2 in that a velocity filter has been applied and the mean value and linear trend have been removed. These have the effect of enhancing the phase map representation by isolating the wave-induced variations. The theoretical (straight-line) constant phase traces were obtained by assuming an impulsive source at a range (in the direction of propagation) of 2100 km, at a time of 1211 UT. The altitude of the source was estimated to have been 125 km. The fit through six cycles of the gravity wave (spanning three hours of observations) is remarkably good.

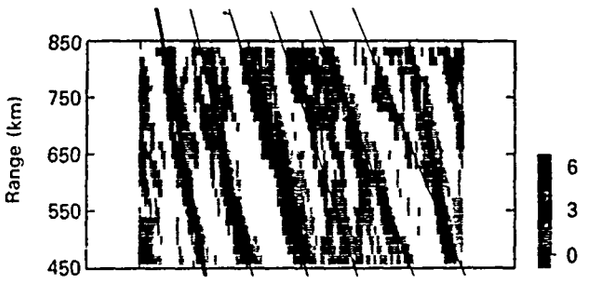


Figure 8. A comparison of the best-fit model phase map with the experimental data recorded for beam 8, on November 25, 1988 The filtered and detrended power data are displayed in terms of dB above noise.

The prediction of a source at a certain location and time can be compared with radar observations of the auroral electrojets. Although the power presentation of Figure 2 is

not ideal for distinguishing between gravity wave effects and backscatter from ionospheric currents, much of the activity at the very greatest ranges in fact originated in a current system. We have analyzed the velocities associated with this scatter in the 1200 - 1300 UT interval to produce the time series plot of plasma convection shown in Figure 9. (For a discussion of the conversion of line-of-sight Doppler velocities into convection velocity vectors, see Ruohoniemi et al. [1989].) This indicates that a shear flow reversal formed at about 1215 UT. The position and time of this development coincide with the source parameters obtained in the preceding analysis. Thus it appears that that the gravity wave activity seen by the Goose Bay in the 1330 - 1630 UT interval was triggered by an earlier shearing of the convective flow (and hence of the ionospheric current) at very high latitudes. We note that the sources of gravity wave activity can generally be associated with auroral electrojet regions observed by the radar.

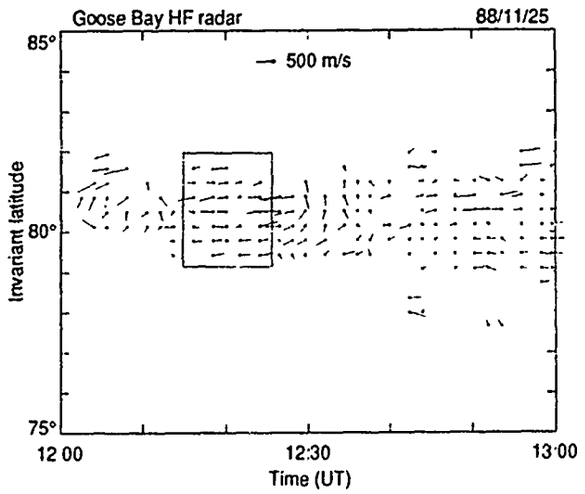


Figure 9. Time series plot of the latitudinal profile of the ionospheric convection for the period 1200 - 1300 UT on November 25, 1988. A shear reversal forms in the flow pattern near 80° invariant latitude at ~1215 UT and lasts for approximately 10 minutes.

CONCLUSION

The HF backscatter technique is a powerful tool for the study of gravity wave phenomena. The high-speed radar measurements at a variety of azimuths and ranges effectively constitute a large and densely packed network of ionospheric 'sensors'. By analyzing the temporal and spatial variations over this network, gravity waves in the high-latitude ionosphere can be identified, their modes of propagation determined, and their sources of excitation located. At the next stage the scope of this work will be extended to include questions of energy transport between the magnetosphere and thermosphere and the regulation of gravity wave activity by magnetospheric processes.

ACKNOWLEDGEMENTS

The Applied Physics Laboratory HF radar at Goose Bay, Labrador is supported in part by the National Science Foundation (NSF) Division of Atmospheric Sciences and the Air Force Office of Scientific Research, Directorate of Atmospheric and Chemical Sciences, under NSF grant ATM-8713982, and, in part, by the National Aeronautics and Space Administration (NASA) under grant NAG5-1099.

REFERENCES

Francis, S. H., A Theory of Medium-Scale Travelling Ionospheric Disturbances, J. Geophys. Res., 69, 5245, 1974.

Greenwald, R. A., K. B. Baker, R. A. Hutchins, and C. Hanuise, An HF Phased-Array Radar for Studying Small-Scale Structure in the High-Latitude Ionosphere, Radio Sci., 20, 63, 1985.

Hines, C. O., Internal Gravity Waves at Ionospheric Heights, Can. J. Phys., 38, 1441, 1960.

Richmond, A. D., Gravity Wave Generation, Propagation, and Dissipation in the Thermosphere, J. Geophys. Res., 83, 4131, 1978.

Ruohoniemi, J. M., R. A. Greenwald, K. B. Baker, J.-P. Villain, C. Hanuise, and J. Kelly, Mapping High-Latitude Plasma Convection with Coherent HF Radars, J. Geophys. Res., 94, 13463, 1989.

Samson, J. C., R. A. Greenwald, J. M. Ruohoniemi, and K. B. Baker, High-Frequency Radar Observations of Atmospheric Gravity Waves in the High-Latitude Ionosphere, Geophys. Res. Lett., 16, 1989.

Samson, J. C., R. A. Greenwald, J. M. Ruohoniemi, A. Frey, and K. B. Baker, Goose Bay Radar Obse-vations of Earth-Reflected Atmospheric Gravity Waves in the High-Latitude Ionosphere, in press, J. Geophys. Res., 1990.

AD-P006 276

COMPUTER SIMULATIONS OF PLASMA STRUCTURE, DYNAMICS, AND TURBULENCE IN THE HIGH LATITUDE IONOSPHERE WITH MAGNETOSPHERIC COUPLING

M.J. KESKINEN, J.D. HUBA, S.T. ZALESAK, AND J.A. FEDDER
SPACE PLASMA BRANCH
PLASMA PHYSICS DIVISION
NAVAL RESEARCH LABORATORY
WASHINGTON, DC 20375-5000

H.G. MITCHELL AND P. SATYANARAYANA
SCIENCE APPLICATIONS INTERNATIONAL CORPORATION
MCLEAN, VIRGINIA 22102

ABSTRACT

The Naval Research Laboratory high latitude ionosphere-magnetosphere mesoscale coupling model has been generalized to include scale-size dependent magnetospheric coupling and E-region Hall and Pedersen conductivity effects. With this model, the nonlinear evolution of plasma interchange-like instabilities in the high latitude F-region ionosphere has been studied using computer simulation techniques. The power spectra of density and electric field irregularities have been computed from the simulations and are found to be in agreement with observations. Applications to density irregularities in and near convecting patches, blobs, and arcs in the high latitude ionosphere are discussed.

INTRODUCTION

Recently, experimental [Basu et al., 1984; Gurnett et al., 1984; Weimer et al., 1985; Basu et al., 1988; Cerisier et al., 1985; Baker et al., 1986; Vickrey et al., 1986; Kintner et al., 1987; Weber et al., 1989; Basu et al., 1990] and theoretical [Mitchell et al., 1985; Lotko et al., 1987; Chaturvedi and Huba, 1987; Keskinen et al., 1988] studies have indicated that the high-latitude ionosphere can be described as a highly dynamic and structured medium containing electric field, magnetic field, and density fluctuations with scale sizes ranging from hundreds of kilometers to centimeters [for recent reviews, see Kintner and Seyler, 1985; Temerin and Kintner, 1988; and Tsunoda, 1988]. The magnitude and extent of structure in the high latitude ionosphere is a function of space and time and can be organized, in general, by geomagnetic latitude and local time [Tsunoda, 1988]. Plasma structure can cause severe propagation effects on C^3I systems operating in and through the ionosphere.

Many source mechanisms have been proposed [see, e.g., Kintner and Seyler, 1985; Temerin and Kintner, 1988; Tsunoda, 1988] to explain features of high latitude ionospheric plasma dynamics, structure and turbulence, e.g., plasma instabilities, currents, particle precipitation, electric fields. Ossakow and Chaturvedi [1979] proposed the current convective instability, an interchange-like plasma instability driven by field-aligned currents, as a source of density irregularities in the high latitude diffuse auroral zone. Keskinen and Ossakow [1982, 1983] have proposed a type of plasma interchange instability, the $\underline{E} \times \underline{B}$ instability, to be a source of high latitude plasma density and electric field fluctuations in the ionosphere for fluctuation scale sizes perpendicular to the geomagnetic field from tens of kilometers to meters. This interchange-like instability, a plasma analogue of the well-known Rayleigh-Taylor instability, which evolves when a heavy fluid is supported against gravity by a lighter fluid, is driven by both a plasma density gradient and an electric field perpendicular to the geomagnetic field. The electric field will result in a relative velocity between convecting ionospheric ions and the neutral thermosphere. The gravitational Rayleigh-Taylor instability has been successfully applied [Ossakow, 1981; Kelley and McClure, 1981] to the explanation of rising bubble and plume-like phenomena during equatorial Spread-F in the low latitude ionosphere. The interchange $\underline{E} \times \underline{B}$ drift instability has been applied to the small scale structure and dynamics in convecting large scale ionospheric density structures recently observed at high latitudes [Vickrey et al., 1980; Muldrew and Vickrey, 1982; Buchau et al., 1983, 1985; Weber et al., 1984, 1986; Foster and Doupnik, 1984; de la Beaujardiere et al., 1985]. Several studies [Sojka and Schunk, 1986; Schunk and Sojka, 1987; Anderson et al., 1987] have modeled features, e.g., peak densities, associated with the large-scale global nature of these density structures. Vickrey and Kelley [1982] have considered the effects of a conducting E layer on the evolution of convecting unstable large scale density structures. At smaller scales, Huba et al. [1983] have studied the

91-09669

linear theory of the interchange $\underline{E} \times \underline{B}$ instability with an inhomogeneous electric field with application to the stability of large scale convecting density structures in the high latitude ionosphere. Keskinen [1983] has investigated the nonlinear evolution of the $\underline{E} \times \underline{B}$ instability with an inhomogeneous electric field.

Mitchell et al. [1985] have studied the nonlinear evolution of the $\underline{E} \times \underline{B}$ drift instability in the high latitude ionosphere with magnetospheric coupling. They found that the primary effect of the magnetosphere is to incorporate ion inertial effects into the development of the interchange-like $\underline{E} \times \underline{B}$ instability. In addition, they found that interchange-like instabilities in the high latitude ionosphere develop in a fundamentally different manner when magnetospheric coupling is included as opposed to the case when magnetosphere-ionosphere coupling is absent. Chaturvedi and Huba [1987] have added three-dimensional effects to the linear theory of interchange instabilities in the high latitude ionosphere. They found that the three-dimensional extent of a large scale convecting ionospheric density structure in the high latitude ionosphere results in a strong (weak) reduction in the growth rate of large (small) scale size fluctuations perpendicular to the geomagnetic field driven unstable by the $\underline{E} \times \underline{B}$ instability.

Recent experimental studies [Basu et al., 1990] have interpreted, in part, simultaneous density and electric field fluctuations measured in regions of strong ionospheric density gradients in and near convecting large scale structures in the polar ionosphere in the light of interchange $\underline{E} \times \underline{B}$ instability processes. Weber et al., [1989] have attributed small scale density and electric field structure on the dusk-side of a convecting polar cap arc to an interchange process. It is now generally believed [Kintner and Seyler, 1985; Temerin and Kintner, 1988' Tsunoda, 1988] that plasma interchange-like instabilities are a major source of electric field and density fluctuations in the high latitude ionosphere for scale sizes of tens of kilometers to meters perpendicular to the geomagnetic field.

However, the three-dimensional nonlinear evolution of interchange instabilities in the high latitude ionosphere and associated scale-size-dependent magnetospheric coupling has not been investigated. In this study we report on the nonlinear evolution of interchange $\underline{E} \times \underline{B}$ instabilities in the high latitude ionosphere with scale-size-dependent magnetospheric coupling.

MODEL EQUATIONS

We model the evolution of the interchange $\underline{E} \times \underline{B}$ instability in the high latitude ionosphere with scale-size-dependent magnetospheric coupling. A fully three-dimensional electromagnetic treatment of the nonlinear evolution of interchange instabilities in the high latitude ionosphere with scale-size-dependent magnetospheric coupling would not be practical either computationally or analytically. We develop a simpler model [Mitchell et al., 1985; Keskinen et al., 1988] which, we believe, retains the principal physical effects of a fully three-dimensional electromagnetic model. We assume that the instability time scales of interest are long compared to the Alfvén transit time across the modeled region, and, as a result, use an electrostatic approximation. The range of time scales beyond which our model is meaningful is on the order of several minutes to tens of minutes. We assume that the plasma is quasi-neutral everywhere and that the electric field maps perfectly along the magnetic field in the modeled region. Due to their high mobility, we take all parallel currents to be carried by the electrons. In addition, we assume that the ions carry the perpendicular currents. As a result, the model may be viewed as a pair of ion layers perpendicular to the magnetic field representing both the ionospheric and magnetospheric contributions. These layers are linked by parallel electron currents, which flow between the layers in order to preserve quasineutrality, and by the mapping of perpendicular electric fields. This model is discussed in more detail in previous studies [Mitchell et al., 1985; Keskinen et al., 1988].

The governing equations are the continuity and momentum equations for each species for each layer, i.e.,

$$\frac{\partial n_\alpha}{\partial t} + \nabla \cdot n_\alpha \underline{v}_\alpha = 0 \tag{1}$$

$$\left(\frac{\partial}{\partial t} + \underline{v}_\alpha \cdot \nabla\right)\underline{v}_\alpha = \frac{e}{m_i}\left(\underline{E} + c^{-1}\underline{v}_\alpha \times \underline{B}\right) - \nu_{in}\,\underline{v}_i \tag{2}$$

where $\alpha = i,e$ denotes species, n_α is the density, m_i the ion mass, v_α is velocity, ν_{in} is the ion-neutral collision frequency, e is the electron charge, c is the speed of light, and $\underline{B}$ is the geomagnetic field. To lowest order, Eq. (2) gives for the electrons

$$\underline{v}_{e\perp} = \frac{c}{B}\,\underline{E} \times \hat{z} \tag{3}$$

and, for the ions,

$$\underline{v}_{i\perp} = \frac{c}{B}\underline{E} \times \hat{z} + \frac{c}{B}\frac{\nu_{in}}{\Omega_i}\underline{E} + \frac{c}{B\Omega_i}[\frac{\partial}{\partial t} + \underline{v} \cdot \nabla]\underline{E} \tag{4}$$

with $\underline{v} \equiv \underline{v}_{e\perp}$ and Ω_i the ion gyrofrequency. The electron perpendicular motion is simply given by an $\underline{E} \times \underline{B}$ drift while the ion perpendicular velocity is composed of an $\underline{E} \times \underline{B}$ drift, a Pedersen drift (second term in Eq. (4)) and a polarization drift (third and fourth terms in Eq. (4)). The perpendicular current density within each layer may be calculated using Eq. (3) and (4) with

$$\underline{J}_\perp = ne(\underline{v}_{i\perp} - \underline{v}_{e\perp}) = \sigma_p\underline{E} + c_m(\frac{\partial}{\partial t} + \underline{v} \cdot \nabla)\underline{E} \tag{5}$$

where $\sigma_p = (nec\nu_{in}/B\Omega_i)$ is the Pedersen conductivity of the layer; $c_m = c^2/4\pi V_A^2$ and $V_A = B/(4\pi nm_i)^{1/2}$ is the Alfvén speed. Mitchell et al. [1985] have shown that electric fields generated by the low-frequency, long-wavelength, fluid-like interchange $\underline{E} \times \underline{B}$ instability generated in the high latitude ionosphere can map far into the magnetosphere on time scales comparable to the interchange $\underline{E} \times \underline{B}$ instability growth time. As a result, a magnetic-field-line-integrated model of the interchange $\underline{E} \times \underline{B}$ instability is appropriate for application to the high latitude near-earth coupled magnetosphere-ionosphere space plasma system. We assume that all currents generated within our modeled region also close within the model. As a result, from quasineutrality we have

$$\int dz\ \nabla \cdot J = 0 \tag{6}$$

giving

$$\nabla \cdot [\Sigma_p\underline{E} + C_m(\frac{\partial}{\partial t} + \underline{v} \cdot \nabla)\underline{E}] = 0 \tag{7}$$

where

$$\Sigma_p = \int_{z_{min}}^{z_{max}} dz\ \sigma_p \tag{8}$$

and

$$C_m = \int_{z_{min}}^{z_{max}} dz\ c_m \tag{9}$$

We note that Eq. (7) now defines a characteristic frequency $\nu = \Sigma_p/C_m$. For fluctuations with characteristic frequency $\partial/\partial t \sim \gamma$, then, $\mathcal{V} \equiv \nu/\gamma$ is dimensionless and is basically the ratio of the integrated Pedersen current and integrated polarization current on the high latitude flux tube of interest. The term C_m is due to cumulative inertia of the plasma along the geomagnetic field and has units of capacitance. In previous studies [Mitchell et al., 1985; Keskinen et al., 1988] we have taken C_m and hence ν to be a constant. In this study we improve this model. In Section III we show that the upper limit of integration z_{max} in the integral for C_m is dependent on $k_\perp$ with $2\pi/k_\perp$ the approximate fluctuation scale size perpendicular to the magnetic field. As a result, C_m and ν are scale-size-dependent. We take the Pedersen conductivity $\Sigma_p \propto \nu_N$ and $C_m \propto N$ with N the integrated plasma density. We can then write, with $\underline{E} = -\nabla\phi$,

$$\frac{\partial N}{\partial t} + \frac{c}{B}\hat{z} \times \nabla\phi \cdot \nabla N = 0 \tag{10}$$

$$\nabla \cdot [\nu N\ \nabla\phi + N(\frac{\partial}{\partial t} + \frac{c}{B}\hat{z} \times \nabla\phi \cdot \nabla)\nabla\phi] = 0 \tag{11}$$

a closed system of equations describing the evolution of low frequency, long wavelength ionospheric interchange instabilities with magnetospheric inertial coupling. In this improved model, we retain a passive magnetospheric layer, i.e., we ignore self-consistent changes in C_m due to magnetospheric convection. This will be a good assumption for incompressible $\underline{E} \times \underline{B}$ flow in the magnetosphere. For electrostatic potential ϕ or density N fluctuations in a magnetic

flux tube with characteristic frequency $\partial/\partial t \sim \gamma$, then, as will be shown in the next section, $\nu/\gamma > 1$ implies ionospheric collisional control while $\nu/\gamma < 1$ suggest magnetospheric inertial control.

ELECTRIC FIELD MAPPING AND LINEAR THEORY

In order to solve Eq. (10)-(11) for the evolution of the interchange $\underline{E} \times \underline{B}$ instability in the high latitude ionosphere with scale-size-dependent magnetospheric coupling, one must first evaluate the quantity $\nu = \Sigma_p/C_m$. Chaturvedi and Huba (1987) have discussed the three-dimensional linear theory of the interchange $\underline{E} \times \underline{B}$ instability in the high latitude ionosphere. For arbitrary background ionospheric density profiles N_0 (x,y,z) with a localized "top-hat" dependence in the z-direction (altitude) and perturbation quantities δN, $\delta\phi$ taken to be of the form $f(z)\exp(\gamma t + ik_x x)$, they found that the electric field mapping distances k_z^{-1} along the geomagnetic field can be written

$$\frac{k_z^2}{k_x^2} = (\alpha R)^{-1}(1 + \gamma/\nu_{in}) \tag{12}$$

where

$$R = (\gamma + k_x^2 D_r)^{-1}(k_x^2 D_r) \tag{13}$$

and $\alpha = \Omega_e\Omega_i/\nu_e\nu_i$, $D_r = \nu_{ei}c^2/\omega_{pe}^2$ with $\omega_{pe}^2 = 4\pi N_0 e^2/m_e$, Ω_e the electron gyrofrequency, Ω_i the ion gyrofrequency, ν_e the electron collision frequency, and ν_{in} is the ion collision frequency. For electrostatic modes, $\gamma << k_x^2 D_r$, Eq. (12) gives $k_z^{-1} \simeq (\Omega_e\Omega_i/\nu_e\nu_i)^{1/2} k_x^{-1} \simeq (\sigma_\parallel/\sigma_\perp)^{1/2} k_x^{-1}$, where $\sigma_\parallel(\sigma_\perp)$ is the parallel (perpendicular) conductivity. For electromagnetic modes $\gamma >> k_x^2 D_r$, we find $k_z^{-1} \simeq V_A/\gamma$ with V_A the Alfven speed. Figure 1 gives typical mapping distances k_z^{-1} for a range of k_x to be expected from the interchange $\underline{E} \times \underline{B}$ instability in the high-latitude ionosphere. We see from Fig. 1 that the long-wavelength fluctuations can map well into the magnetosphere, with the mapping distance V_A/γ determined by inertial processes, while shorter scale size modes are more localized in the ionosphere, with a mapping distance $k_z^{-1} \simeq (\Omega_e\Omega_i/\nu_e\nu_{in})^{1/2} k_x^{-1}$ determined by collisional effects.

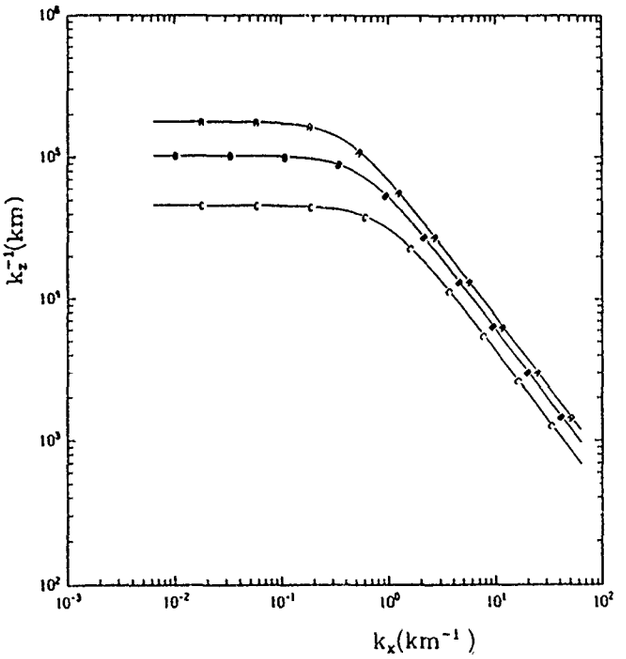


Fig. 1 Plot of the electric field mapping distance k_z^{-1} (km) along the magnetic field vs. the cross-field wavenumber k_x (km^{-1}) using Eq. (12) and (13) with $D_r = 5.6 \times 10^{-2}$ km^2/sec, $\nu_i = 0.01$ sec^{-1}. Curves A, B, C refer to typical growth rates $\gamma = 0.01$, 0.02, 0.03 sec^{-1}.

Figure 2 gives typical values for $\tilde{\nu} \equiv \nu/\gamma_0$, $\gamma_0 = cE_0/BL$, and $L^{-1} = \partial \ln N_0/\partial y$ for the interchange $\underline{E} \times \underline{B}$ instability in the high latitude ionosphere using Figs. (1) and (2) and Eqs. (8)-(9). Here, E_0 is the initial background electric field magnitude. In computing ν we have taken $z_{max} \simeq k_z^{-1}$ in Eq. 9. Since k_z^{-1} is a function of k_x, C_m and ν also will be functions of k_x and, hence, scale-size-dependent. From Fig. 2 one notes that ν increases as one goes from small k_x (long wavelength) to large k_x (shorter wavelength). In previous studies [Mitchell et al., 1985] the value for $\tilde{\nu}$ was taken to be a constant independent of k_x, the fluctuation scale size perpendicular to the geomagnetic field. From Fig. 2 one notes that the low-frequency, long-wavelength interchange modes are more collisional (inertial) for shorter (longer) wavelengths.

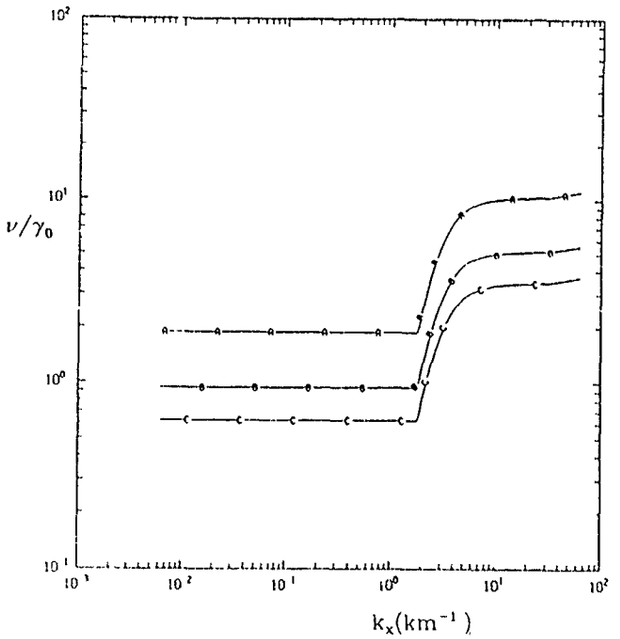


Fig. 2 Plot of $\tilde{\nu} = \nu/\gamma_0 = \Sigma_p/C_m\gamma_0$, $\gamma_0 = cE_0/BL$, with $\Sigma_p = 0.1$ mho for (a) $\gamma_0 = 0.01$ sec^{-1}, (b) $\gamma_0 = 0.02$ sec^{-1}, and (c) $\gamma_0 = 0.03$ sec^{-1}.

NONLINEAR EVOLUTION

Since the nonlinear evolution of interchange instabilities is complicated and not amenable to an exact analytical treatment, we resort to numerical techniques. We now discuss, using numerical simulation techniques, the nonlinear evolution of the interchange $\underline{E} \times \underline{B}$ instability in the high latitude ionosphere with scale-size-dependent magnetospheric coupling, i.e., with $\tilde{\nu}(k_x)$ as given in Fig. 2. Using Fig. 2 we take the following approximate profile for $\tilde{\nu}$

$$\tilde{\nu} = \begin{cases} 0.8 & 2\pi/|\underline{k}| > 3 \text{ km} \quad (14a) \\ 5.4 & 2\pi/|\underline{k}| = 3 \text{ km} \quad (14b) \\ 10 & 2\pi/|\underline{k}| < 3 \text{ km} \quad (14c) \end{cases}$$

Equation (10) is solved numerically using the multidimensional flux-corrected-transport method [Zalesak, 1979], while Eq. (11) is solved using the Cholesky conjugate gradient algorithm of Hain [1980]. The simulations are performed on a 66 x 100 cell grid (x,y) with cell sizes $\Delta x = 0.3$ km and $\Delta y = 0.5$ km. The entire simulation plane is drifting with the $\underline{V}_0 = c\underline{E}_0 \times \underline{B}/B^2$ velocity. Periodic boundary conditions are enforced in the x-direction with Neumann ($\partial\bar{\phi}/\partial y = 0$) conditions at the y-direction boundaries. The interchange instability is initialized using a

slab model for the density of the form $N_0(y)$ [1 + $\varepsilon(x,y)$] with $N_0(y) = N_0[1 + A \tanh(y/L)]$, A = 0.66, and $\varepsilon(x,y)$ is a random function with maximum magnitude $|\varepsilon_{max}| = 0.01$. For this choice of initial density profile, we find $(\partial \ln N_0/\partial y)_{max} = 1.2L^{-1}$ at y/L = -0.25. The gradient scale length L= 25 km with E_{0x} = 20 mV/m giving a plasma convection velocity, in the rest frame of the neutral thermosphere, of 500 m/sec.

Figure 3 displays the evolution of plasma density in a plane perpendicular to the geomagnetic field, due to the nonlinear evolution of the interchange $\underline{E} \times \underline{B}$ instability with scale-size-dependent magnetospheric coupling as modeled with $\nu(k_x)$ given by Eq. (14). In the fully developed state, Fig. 3b, the fingers have elongated with some inertial rolling and vortex formation.

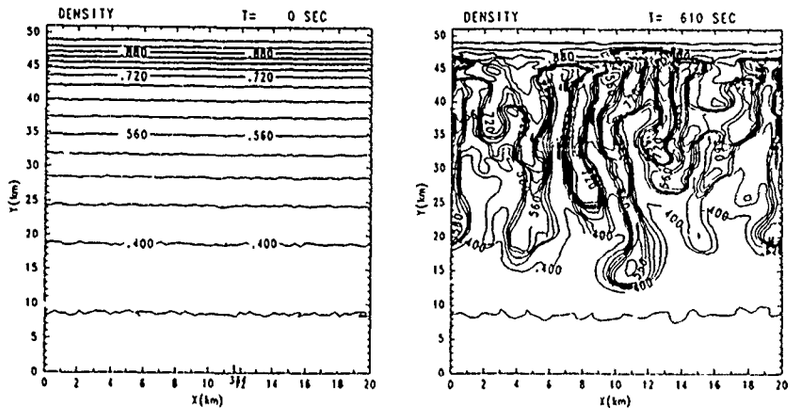


Fig. 3 A plot of the density $N(x,y,t)/N_0$ contours for scale-size-dependent case at two different times (a) t = 0 sec, (b) t = 610 sec. All contours have been multiplied by scale factor of 0.6.

Figure 4 gives the time-averaged spatial power spectra $P(k_x)$ and $P(k_y)$ in the nonlinear regime for the scale-size-dependent case (Eq. (14) corresponding to Fig. 3. We consider a two-dimensional spatial Fourier representation

$$f(x,y,t) = \sum_{k_x,k_y} f_{k_x,k_y}(t)\exp[i(k_x x + k_y y)]$$

with $f = \delta N = N - N_o$, $\delta E_x = E_x - E_o$, $\delta E_y = E_y - E_o$. Here,

$$f_{k_x,k_y} = (2\pi^2) \int dx\, dy\, f(x,y)\exp[-i(k_x x + k_y y)]$$

The transverse averaged power spectra $P(k_x)$ and $P(k_y)$ are defined by

$$P(k_x) = \sum_{k_y} \langle |f_{k_x,k_y}|^2 \rangle$$

$$P(k_y) = \sum_{k_x} \langle |f_{k_x,k_y}|^2 \rangle$$

We have time averaged the spectra computed in the nonlinear regime over a time interval of approximately five linear growth times. For the scale-size-dependent case, this corresponds to t = 500-700 sec. For the density $|\delta N|^2$ we find $P(k_x) \propto k_x^{-n_x}$, $n_x = 2.2 \pm 0.3$ for 1.9 km < $2\pi/k_x$ < 10.3 km and $P(k_y) \propto k_y^{-n_y}$, $n_y = 1.8 \pm 0.3$ for 2.0 km < $2\pi/k_y$ < 15.5 km. For the electric field component $|\delta E_x|^2$, we find $P(k_x) \propto k_x^{-n_x}$, $n_x = 2.3 \pm 0.3$ for 1.9 km < $2\pi/k_x$ < 10.3 km and

$P(k_y) \propto k_y^{-n_y}$, $n_y = 2.1 \pm 0.3$ for 2.5 km $< 2\pi/k_y <$ 24.8 km. For the electric field component $|\delta E_y|^2$ we find $P(k_x) \propto k_x^{-n_x}$, $n_x = 2.2 \pm 0.3$ for 1.9 km $< 2\pi/k_x <$ 10.3 km, and $P(k_y) \propto k_y^{-n_y}$, $n_y = 1.9 \pm 0.3$ for 2.7 km $< 2\pi/k_y <$ 12.4 km. The power spectra for the other two cases, i.e., $\nabla = 0.8$, 10 are displayed in Table 1. In general the spectral indices for the density and electric field for the scale-size-dependent case are intermediate between the cases $\nabla = 0.8$, 10. The spectra for the density and electric field are, for the most part, slightly steeper for the collisional case ($\nabla = 10$) than for the inertial case ($\nabla = 0.8$).

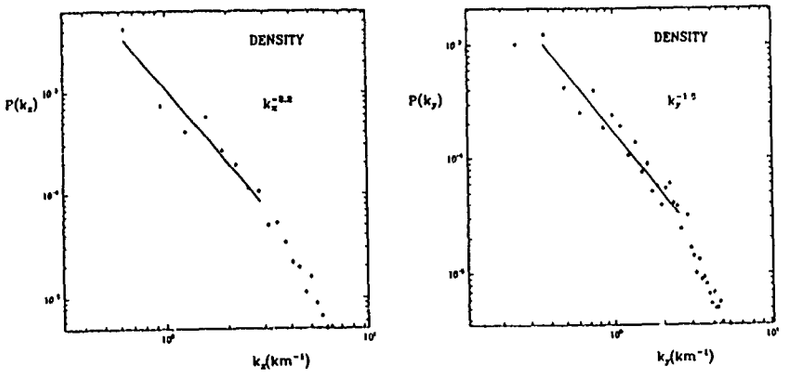


Fig. 4 Plot of time-averaged power spectra (a) $P(k_x)$ and (b) $P(k_y)$ for plasma density δN for scale-size-dependent case. The spectra $P(k_x)$ and $P(k_y)$ are given in arbitrary units.

TABLE 1. COMPARISON OF SPECTRAL INDICES

		$\nabla = 0.8$ (inertial)	∇*	$\nabla = 10$ (collisional)
δN	$P(k_x)$	2.1±0.3	2.2±0.3	2.5±0.2
	$P(k_y)$	1.9±0.2	1.8±0.3	2.0±0.2
δE_x	$P(k_x)$	2.1±0.3	2.3±0.3	2.5±0.3
	$P(k_y)$	1.9±0.2	2.1±0.3	2.1±0.2
δE_y	$P(k_x)$	2.0±0.2	2.2±0.3	2.4±0.2
	$P(k_y)$	1.9±0.3	1.9±0.3	2.1±0.3

*∇ is scale-size-dependent with $\nabla = 0.8$ for $2\pi/|\underline{k}| > 3$ km and $\nabla = 10$ for $2\pi/|\underline{k}| < 3$ km.

In the spectra for the scale-size dependent case weak spectral breaks are occurring at wave numbers $k_B \simeq 3\text{-}4\ km^{-1}$ corresponding to wavelengths $\lambda_B = 2\pi/k_B \simeq 2 - 3$ km, i.e., the spectra are steeper for $\lambda \lesssim \lambda_B$ than for $\lambda \gtrsim \lambda_B$. This spectral break behavior can be explained as a consequence of the collisional spectra being steeper than the inertial spectra. In other words, if one superimposes a collisional ($\nu = 10$) spectrum for $2\pi/k \lesssim 3$ km on a less steep inertial ($\nu = 0.8$) spectrum for $2\pi/k \gtrsim 3$ km, a break scale naturally arises at $2\pi/k \simeq 3$ km. For the most part, we have found that $|\delta N|^2 \propto k^{-n}$, $n \simeq 2$ and $|\delta E|^2 \propto k^{-n}$, $n \simeq 2$ for $k \lesssim 3\ km^{-1}$. The k^{-2} spectral behavior is to be expected due to the steep density gradients generated in the nonlinear evolution of the interchange instability.

SUMMARY AND CONCLUSIONS

We have made a preliminary study of the nonlinear evolution of the interchange $\underline{E} \times \underline{B}$ instability in the high latitude ionosphere with scale-size-dependent magnetospheric coupling. We find that the principal effect of magnetospheric coupling, i.e., increased plasma inertia, is most pronounced for long-wavelength ionospheric fluctuations. We find that the phenomenology and time scales of the nonlinear evolution of ionospheric interchange modes with scale-size-dependent magnetospheric coupling can be characterized as neither purely inertial (magnetospheric control) nor purely collisional (ionospheric control).

Recent studies [Basu et al., 1990], using Dynamics Explorer Satellite data, have measured the simultaneous density and electric field fluctuation power spectra in regions of ionospheric density gradients in and near convecting large scale structures in the polar ionosphere. The horizontal density gradients associated with large scale convecting ionospheric patches, blobs, and structures can be unstable to the interchange $\underline{E} \times \underline{B}$ instability as these structures convect through the thermosphere. Basu et al. find that the density δN and electric field δE spectra can be represented by power laws k^{-n} with the spectral index for density (electric field) peaking near $n \simeq 1.9 \pm 0.3$ for noon-midnight passes across the polar cap ionosphere. For approximate dusk-dawn passes, they find that $n \simeq 1.9 \pm 0.3$ for the density and $n \simeq 1.9 - 2.6$ for the electric field. In addition, Cerisier et al. [1985] have measured the density and electric field spectra in regions of large scale ionospheric plasma density gradients. They, too, find power law behavior with spectral index $n \simeq 1.9$ for the density and $n \simeq 1.8$ for the electric field component parallel to the spacecraft trajectory. For the interchange $\underline{E} \times \underline{B}$ instability we find the power spectrum of the density δN fluctuation $P(k_x) \propto k_x^{-n_x}$, $n_x = 2.2 \pm 0.3$ and $P(k_y) \propto k_y^{-n_y}$, $n_y = 1.8 \pm 0.3$. For electric field δE_x spectra, we compute $P(k_x) \propto k_y^{-n_x}$, $n_x = 2.3 \pm 0.3$ and $P(k_y) \propto k_y^{-n_y}$, $n_y = 2.1 \pm 0.3$ while for the component δE_y we find $P(k_x) \propto k_x^{-n_x}$, $n_x = 2.2 \pm 0.3$ and $P(k_y) \propto k_y^{-n_y}$, $n_y = 1.9 \pm 0.3$. Here the x-direction is perpendicular to the initial ionospheric plasma density gradient. Within the error bounds, these theoretically computed spectral indices are in agreement with those derived from satellite observations. Exact agreement among experimental and theoretical spectral indices may not be possible since the initial conditions, time history, and spatial evolution of large scale polar and auroral ionospheric density structures is not known. Furthermore, plasma density inhomogeneities in the high-latitude F-region ionosphere can last for hours because of small recombination rates and diffusion rates. The range of time scales over which our model calculations are meaningful is from a few minutes to about ten minutes. As a result, the very long time dynamics and evolution of such inhomogeneities is not clear and will be reserved for a future report. In addition, we have neglected in our theoretical model several effects, e.g., Hall conductivity, particle precipitation, thermospheric wind structure, and a more realistic magnetosphere, which could be important for the evolution of turbulence and structure in the high latitude ionosphere.

In the future, we hope to consider E-region and associated Hall conductivity effects, longer time scale effects, and a more detailed magnetospheric response in order to better describe small scale turbulence and structure in the high latitude magnetosphere-ionosphere-thermosphere coupled system.

ACKNOWLEDGMENTS: This work was supported by NASA and ONR.

REFERENCES

Anderson, D.N., J. Buchau, and R.A. Heelis, Origin of density enhancements in the winter polar cap ionosphere, in The Effect of the Ionosphere on Communication, Navigation, and Surveillance Systems, ed. J.M. Goodman, U.S. Government Printing Office, Washington, DC, 1987.

Baker, K.B., R.A. Greenwald, A.D.M. Walker, P.F. Bythrow, L.J. Zanetti, T.A. Potemra, D.N. Hardy, P.J. Rich, and C.L. Rino, A case study of plasma processes in the dayside cleft, J. Geophys. Res., 91, 3130, 1986.

Basu, Su., S. Basu, E. MacKenzie, W.R. Coley, W.B. Hanson, and C.S. Lin, F-region electron density irregularity spectra near auroral acceleration and shear regions, J. Geophys. Res., 89, 5554-5564, 1984.

Basu, Su., S. Basu, E. MacKenzie, P.F. Fougere, W.R. Coley, N.C. Maynard, J.D. Winningham, M. Sugiura, W.B. Hanson, and W.R. Hoegy, Simultaneous density and electric field fluctuation spectra associated with velocity shears in the auroral oval, J. Geophys. Res., 93, 115-136, 1988.

Basu, Su., Sa. Basu, E. MacKenzie, W.R. Coley, J.R. Sharber, and W.R. Hoegy, Plasma structuring by the gradient-drift instability at high latitudes and comparison with velocity-shear driven processes, J. Geophys. Res., 1990 (in press).

Buchau, J., B.W. Reinisch, E.J. Weber, and J.G. Moore, Structure and dynamics of the winter polar cap F region, Radio Sci., 18, 995, 1983.

Buchau, J., E.J. Weber, D.N. Anderson, H.C. Carlson, Jr., J.G. Moore, B.W. Reinisch, and R.C. Livingston, Ionospheric structures in the polar cap: Their origin and relation to 250-MHz scintillation, Radio Sci., 20, 325, 1985.

Cerisier, J.C., J.J. Berthelier, and C. Beghin, Unstable density gradients in the high latitude ionosphere, Radio Sci., 20, 755, 1985.

Chaturvedi, P.K. and J.D. Huba, The interchange instability in high latitude plasma blobs, J. Geophys. Res., 92, 3357, 1987.

de la Beaujardiere, O., J.D. Craven, V.B. Wickwar, G. Candal, J.M. Holt, L.A. Frank, L.H. Brace, D.S. Evans, and J.D. Winningham, Universal time dependence of night-time F region densities at high latitudes, J. Geophys. Res., 90, 4319, 1985.

Foster, J.C. and J.R. Doupnik, Plasma convection in the vicinity of the dayside cleft, J. Geophys. Res., 89, 9107, 1984.

Gurnett, D.A., R.L. Huff, J.D. Menietti, J.L. Burch, J.D. Winningham, and S.D. Shawhan, Correlated low frequency electric and magnetic noise along the auroral field lines, J. Geophys. Res., 89, 8971, 1984.

Hain, K., A non-recursive incomplete Cholesky decomposition method for the solution of linear equations with a sparse matrix, NRL Memorandum Report 4264, Naval Research Laboratory, Washington, DC, 1980.

Hassam, A.B., W. Hall, J.D. Huba, and M.J. Keskinen, Spectral characteristics of interchange turbulence, J. Geophys. Res., 91, 13513, 1986.

Huba, J.D., S.L. Ossakow, P. Satyanarayana, and P.N. Guzdar, Linear theory of the $\underline{E} \times \underline{B}$ instability with an inhomogeneous electric field, J. Geophys. Res., 88, 425, 1983.

Kelley, M.C. and J.P. McClure, Equatorial spread F: A review of recent experimental results, J. Atmos. Terr. Phys., 43, 427, 1981.

Keskinen, M.J. and S.L. Ossakow, On the spatial power spectrum of the $\underline{E} \times \underline{B}$ gradient drift instability in ionospheric plasma clouds, J. Geophys. Res., 86, 6947, 1981.

Keskinen, M.J. and S.L. Ossakow, Nonlinear evolution of plasma enhancements in the auroral ionosphere, 1, Long wavelength irregularities, J. Geophys. Res., 89, 144, 1982.

Keskinen, M.J. and S.L. Ossakow, Nonlinear evolution of convecting plasma enhancements in the auroral ionosphere, 2, Small scale irregularities, J. Geophys. Res., 88, 474, 1983.

Keskinen, M.J., H.G. Mitchell, J.A. Fedder, P. Satyanarayana, S.T. Zalesak, and J.D. Huba, Nonlinear evolution of the Kelvin-Helmholtz instability in the high-latitude ionosphere, J. Geophys. Res., 93, 137, 1988.

Kintner, P.M. and C.E. Seyler, The status of observations and theory of high latitude ionospheric and magnetospheric plasma turbulence, Space Sci. Rev., 41, 91, 1985.

Kintner, P.M., M.C. Kelley, G. Holmgren, H. Koskinen, G. Gustafsson, and J. LaBelle, Detection of spatial density irregularities with the Viking plasma wave interferometer, Geophys. Res. Lett., 14, 467, 1987.

Lotko, W., B.U.O. Sonnerup, and R.L. Lysak, Nonsteady boundary layer flow including ionospheric drag and parallel electric fields, J. Geophys. Res., 92, 8635, 1987.

Mitchell, H.G., J.A. Fedder, M.J. Keskinen, and S.T. Zalesak, A simulation of high latitude F layer instabilities in the presence of magnetosphere-ionosphere coupling, Geophys. Res. Lett., 12, 283, 1985.

Muldrew, D.B. and J.F. Vickrey, High latitude F region irregularities observed simultaneously with ISIS-1 and the Chatanika radar, J. Geophys. Res., 87, 907, 1982.

Ossakow, S.L. and P.K. Chaturvedi, Current convective instability in the diffuse aurora, Geophys. Res. Lett., 6, 332, 1979.

Ossakow, S.L., Equatorial spread F- A review, J. Atmos. Terr. Phys., 43, 437. 1981.

Schunk, R.W. and J.J. Sojka, A theoretical study of the lifetime and transport of large ionospheric density structures, J. Geophys. Res., 92, 12343, 1987.

Sojka, J.J. and R.W. Schunk, A theoretical study of the production and decay of localized electron density enhancements in the polar ionosphere, J. Geophys. Res., 91, 3245, 1986.

Temerin, M. and P.J. Kintner, Review of ionospheric turbulence, Proc. Chapman Conference on Plasma Waves and Instabilities in Magnetospheres and Comets, Sendai, Japan, October 1987, AGU monograph (in press).

Tsunoda, R.T., High-latitude F-region irregularities: A review and synthesis, Physica Scripta, 1988 (in press).

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.

Vickrey, J.F. and M.C. Kelley, The effects of a conducting E layer on classical F region cross-field plasma diffusion, J. Geophys. Res., 87, 4461, 1982.

Weber, E.J., M.C. Kelley, J.O. Ballenthin, S. Basu, H C. Carlson, J.R. Fleischman, D.A. Hardy, N.C. Maynard, R.F. Pfaff, P. Rodriguez, R.E. Sullivan, and M. Smiddy, Rocket measurements within a polar cap arc: plasma particle, and electric circuit parameters, J. Geophys. Res., 94, 6692, 1989.

AD-P006 277

SCINTILLATION NEAR THE F-LAYER TROUGH OVER NORTHERN EUROPE

L. Kersley, S. E. Pryse and C. D. Russell
Department of Physics
University College of Wales
Aberystwyth, U.K.

ABSTRACT

Results are presented of scintillation observations made during a two and a half year period at Lerwick in the Shetland Islands using more than 19000 passes of NNSS satellites. Examples of scintillation morphology, in the region near the scintillation boundary and the F-layer trough, for both amplitude and phase are discussed using exceedence levels for the S_4 and σ_ϕ indices respectively. The equatorwards advancement of the scintillation boundary in response to enhanced solar activity during the increasing phase of the solar cycle is shown to be a dominant feature in the observations.

INTRODUCTION

The high-latitude ionosphere is a highly structured medium containing irregularities with a wide range of scale sizes. The irregularities are of interest not only to an understanding of the physical processes controlling the ionospheric plasma but also because of their influence on radio systems propagating in the ionospheric medium. The high-latitude irregularity region is often marked by a sharp equatorwards boundary. For irregularities with scales in the sub-kilometre regime, which are responsible for scintillation of transionospheric radio signals, this boundary has been termed the scintillation boundary. Many observations were made in the 1960's, often using transmissions on 20 MHz and 40 MHz from the BE-B satellite, of scintillations in the vicinity of the boundary, with complementary studies being undertaken in both hemispheres. A characteristic feature of much of this early work was the recognition of the equatorwards motion of the high-latitude irregularity zone in response to enhanced magnetic activity.

The scintillation boundary was found to occur at similar geomagnetic latitudes to those of the minimum in electron density and content associated with the so-called mid-latitude trough. AARONS (1970) noted similarities in response to magnetic activity of the scintillation boundary and the trough, though in a later paper AARONS and ALLEN (1971) commented on discrepancies in the statistics of diurnal behaviour of the two, especially in the early morning sector. Using simultaneous measurements of boundary and trough, KERSLEY et al. (1972) found that between 1500 LT and 0200 LT the scintillation boundary was on average some 1.5° equatorwards of the trough minimum but after 0400 LT the boundary was polewards of the trough, the mean difference increasing with local time. In a subsequent paper KERSLEY et al. (1975) showed that the trough reached a lowest latitude around 0600 LT whereas for the scintillation boundary the minimum latitude occurred at 2300 LT.

The current experiment forms a systematic study of amplitude and phase scintillation in the region of the scintillation boundary and the F-layer trough over northern Europe. It makes use of the abundance of suitable transmissions available from the multi-satellite NNSS system. In the initial phase of this experiment observations were made during a two-year period of scintillations in the auroral zone from a site at Kiruna in northern Sweden. The results of this earlier study were reported by KERSLEY et al. (1987 and 1988). The current paper presents results from a complementary set of long-term observations made during the second phase of the programme at a site near the scintillation boundary and the trough. Sample results are presented here in a format which allows direct comparison with those given in the earlier papers covering the observations in the auroral zone.

91-09670

EXPERIMENT

The observations reported here were made at Lerwick (60.15 N, 1.15 W) in the Shetland Islands from July 1987 to November 1989 using the phase coherent transmissions on 150 and 400 MHz from NNSS satellites. The constellation of up to 7 active satellites in 1100 km polar orbits allows study of scintillations as a function of latitude. Details of the experimental arrangement, including data acquisition and processing, have been given in earlier papers. In the current phase of the project observations have been made of more than 19000 satellite passes and a data base formed comprising records for each 20s of satellite pass totalling more than 750,000 records. Amplitude scintillation is characterised by an S_4 index, the square root of the normalised variance of signal intensity, while phase scintillation of the 150 MHz signal with respect to the 400 MHz reference is represented by σ_ϕ, the root mean square phase variation. In each case a detrend filter with cut-off at 0.2 Hz is used. In the later observations of the current series measurements of the trends in differential carrier phase have been obtained allowing estimates to be made of the variation of total electron content as a function of latitude thus enabling identification of the features of the trough.

SCINTILLATION OCCURRENCE

The sample of figures presented here for the Lerwick observations are in the form of percentage exceedence of specified thresholds of S_4 (for 150 MHz) and σ_φ for amplitude and phase respectively. Direct comparison can be made with the earlier results obtained from the auroral zone studies (KERSLEY et al., 1987 and 1988).

The variation of scintillation occurrence as a function of corrected geomagnetic latitude shows certain characteristic features, illustrated by the examples of Figure 1 which is for amplitude scintillations from satellites within ± 20° longitude of Lerwick. The features are an enhancement of scintillation near the latitude of the station and an increase towards higher latitudes. In addition to the role of the scintillation boundary, both of the above features are partially associated with geometrical effects, the first arising from the field-aligned nature of the irregularities and the second having a contribution from low elevation angle propagation. A striking difference in occurrence levels of scintillation is apparent in the plots shown here which represent extremes in the data set. While the observations for November 1987 show very low overall occurrence levels, the results for March 1989, although of the same general form, have vastly enhanced scintillation occurrence.

The occurrence of scintillation as a function of geomagnetic latitude (at the 350 km ionospheric point of the ray path) for the two and a half year data set (for passes ± 20° longitude) is shown in Figures 2 and 3 for amplitude and phase respectively. For amplitude at 150 MHz all scintillation with $S_4 > 0.2$ is included, while $\sigma_\varphi \geq 25$ degrees has been chosen for the phase fluctuation of the 150 MHz signal with respect to the 400 MHz reference. This value of phase scintillation threshold not only allows direct comparison with the earlier NNSS auroral zone observations but, by allowing for different detrend filter cut-offs, enables the present results to be compared with those of other workers (see KERSLEY et al., 1987).

The results displayed in Figures 2 and 3 show marked differences from the earlier auroral zone observations. The effect of the field-aligned enhancement is now no longer the dominant feature in the plots. Indeed, the equatorward advance of scintillation occurrence in the winter 1988/89 and spring 1989 and again in November 1989 is the most significant aspect of the results. It is suggested that this advance is a consequence of enhanced solar activity, with the movement of the scintillation boundary in response to the approach of the maximum of the solar cycle. The earlier observations in the auroral zone, which were made near solar minimum in 1984 to 1986 showed a clear seasonal dependence with more scintillation in summer than winter. While some evidence of a similar trend can be seen in the current observations, with a reduction of activity at higher latitude in winter 1987/8, it is certainly not sustained in the later measurements.

As examples of diurnal variations the plots for autumn 1987 and autumn 1989 are shown in Figures 4a and b. The nighttime advance to lower latitudes of scintillation occurrence is apparent in both plots but here again the greatly increased levels in the later year are the dominant feature. Plots for other seasons show similar general trends except that in winter the nighttime enhancement is extended in temporal extent. These diurnal plots have been obtained from observations of satellite passes within ± 10° longitude of Lerwick. Broadly similar features are to be found in the diurnal variations of the phase data.

The effect of magnetic activity on scintillation occurrence can be seen in the diurnal plots of Figures 5a and b. These are for amplitude scintillation using data from summer 1989. Under quiet magnetic conditions scintillation maximises in the post midnight hours but when magnetic activity is disturbed the equatorwards advance is apparent with high levels of occurrence covering the middle latitude range for much of the day.

In the earlier auroral zone observations plots from passes covering a wide range of longitudes were used in conjunction with model studies to demonstrate that there was a component in the irregularities having a sheet-like as opposed to field-aligned rod-like form. Similar studies have been undertaken for the present observations. The occurrence plots, of which Figure 6 for autumn 1989 is an example, show little evidence of either the concentric rings about the field-aligned point characteristic of rod-like irregularities or the saddle-shaped contours formed by sheets. While plots of the kind shown in Figure 1 do show field-aligned enhancements for some months there is little evidence for irregularity structures with large axial ratios of the kind found in the earlier observations at auroral zone latitudes. It is concluded that the irregularities in the vicinity of the boundary have anisotropies with much smaller elongation in the magnetic field direction.

It must be appreciated that movements of the scintillation boundary represent a complicating factor in the interpretation of the bulk statistics shown here. Work is continuing in an attempt to try to isolate the effects of boundary motion from the observations.

CONCLUSIONS

Examples have been presented of amplitude and phase scintillation occurrence obtained from observations of more than 19000 NNSS satellite passes at a site in the vicinity of the scintillation boundary and the F-layer trough. The response of scintillation activity to the advancing solar cycle is a dominant feature of the observations.

ACKNOWLEDGEMENTS

This work has been supported by the UK Radio Propagation Programme, while the sponsorship of the Air Force Geophysics Laboratory through Grant AFOSR-87-0378 is gratefully acknowledged. Thanks are due to the Superintendent and staff of the Meteorological Observatory, Lerwick, for the use of their facilities.

REFERENCES

Aarons, J. (1970) The high latitude F-region irregularity structure during the October 30 to November 4, 1968 magnetic storm, Radio Sci., 5, 59.

Aarons, J. and R. S. Allen (1971) Scintillation boundary during quiet and disturbed magnetic conditions, J. Geophys. Res., 76, 170.

Kersley, L., D. B. Jenkins and K. J. Edwards (1972) Relative movements of the mid-latitude trough and scintillation boundary, Nature, 239, 11.

Kersley, L., A. P. van Eyken and K. J. Edwards (1975) Ionospheric mid-latitude trough and abrupt scintillation boundary, Nature, 254, 312.

Kersley, L., S. E. Pryse and N. S. Wheadon (1987) Amplitude and phase scintillation at high latitudes over northern Europe, in The Effect of the Ionosphere on Communication, Navigation and Surveillance Systems, ed. J. M. Goodman, NRL, Washington, D.C., USA, 591.

Kersley, L., S. E. Pryse and N. S. Wheadon (1988) Amplitude and phase scintillation at high latitudes over northern Europe, Radio Sci., 23, 320.

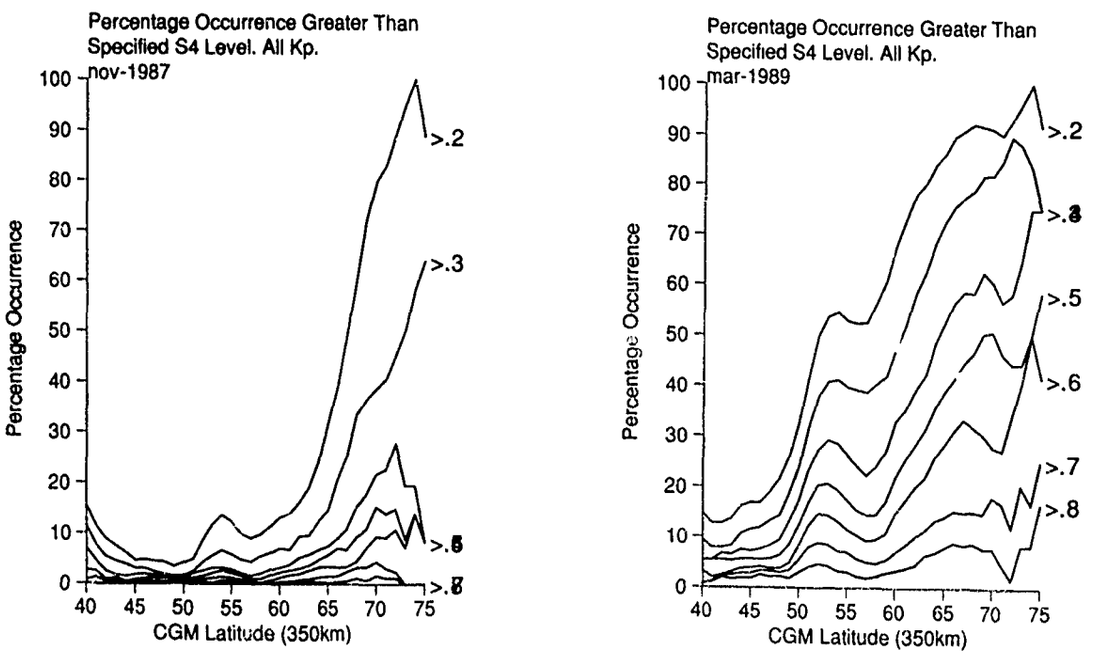


Figure 1.

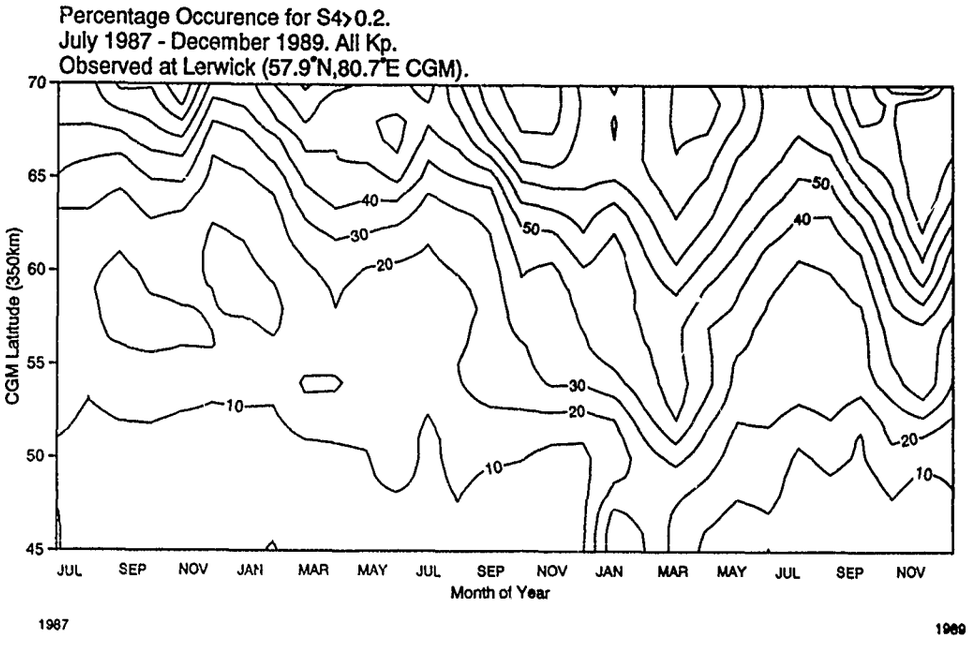


Figure 2.

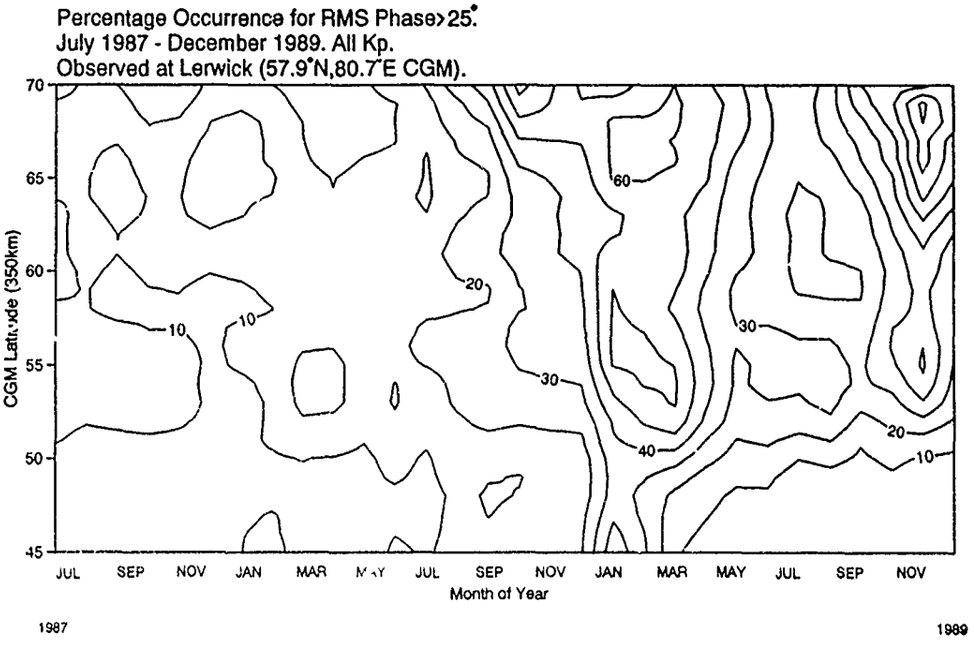


Figure 3.

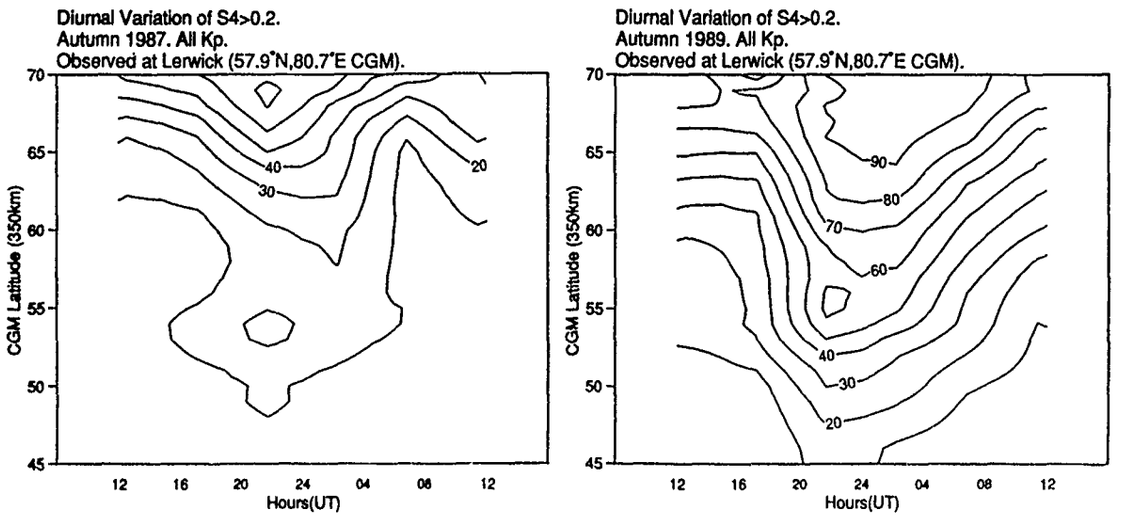


Figure 4a and b

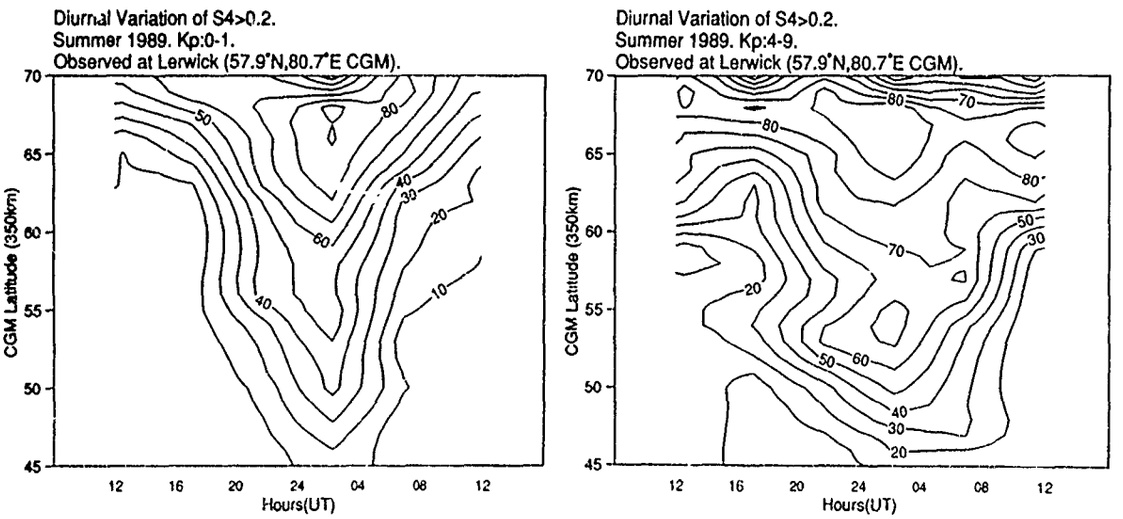


Figure 5a and b.

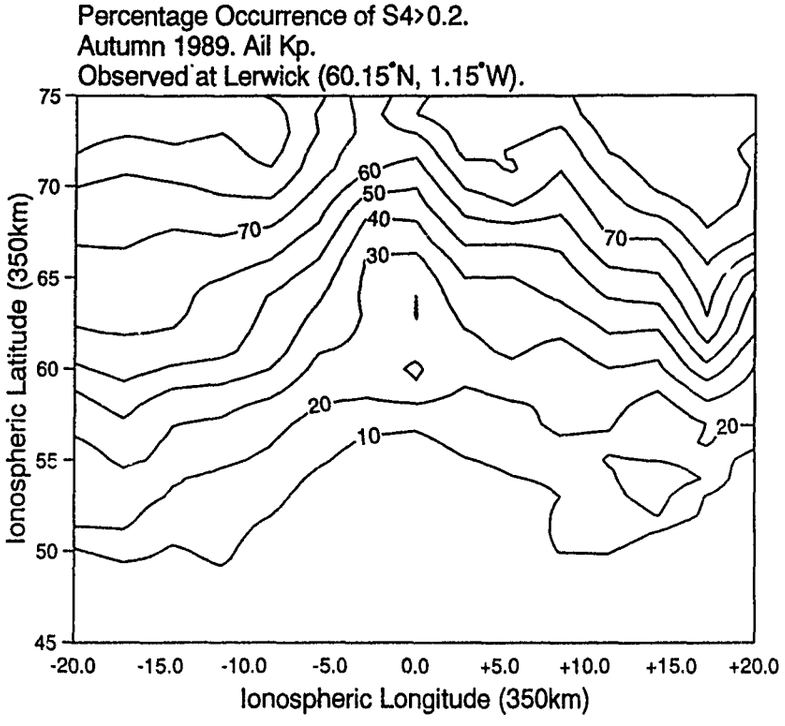

Percentage Occurrence of S4>0.2.
Autumn 1989. Ail Kp.
Observed at Lerwick (60.15°N, 1.15°W).
75
70
65
60
55
50
45
Ionospheric Latitude (350km)
60
50
40
30
70
70
20
20
10
-20.0
-15.0
-10.0
-5.0
0.0
+5.0
+10.0
+15.0
+20.0
Ionospheric Longitude (350km)

Figure 6.

AD-P006 278

COHERENCE SCALES OF WAVEFIELD DURING PROPAGATION THROUGH NATURALLY DISTURBED IONOSPHERE IN THE POLAR CAP, AURORAL AND EQUATORIAL REGIONS

S. Basu
GL(AFSC)/LIS
Hanscom AFB, MA 01731

Sunanda Basu
Boston College
Institute for Space Research
Newton, MA 02159

R.C. Livingston
SRI International
Menlo Park, CA 94025

91-09674

ABSTRACT

Phase and intensity scintillation measurements have been made at low latitudes in the equatorial anomaly region, and at high latitudes in the auroral oval and the polar cap regions, using phase coherent transmissions at 250 MHz from stationary and near stationary satellites. The observations pertain to periods of high solar activity when intense scintillation activity is recorded at each of the above observing sites. This data set has been utilized to study the reduction of coherence times of intensity and complex amplitude scintillation with increasing strength of scattering. Estimates of coherence scales of intensity and complex amplitude scintillation at 250 MHz are provided which indicate that coherent scales of scintillation are typically of the order of hundreds of meters at high latitudes but approach values as small as tens of meters in the equatorial anomaly region. The phase spectral index in the nightside auroral oval is observed to be much steeper (p_ϕ = -4) than those typically observed in the equatorial (p_ϕ = -2.4) or polar cap regions ($p_\phi \approx$ -2.3). It shows the importance of large scale phase variations in the nightside auroral oval. Under strong scatter conditions, the coherence times of complex amplitude scintillation are shown to asymptotically approach a value which is 1.4 times the coherence time of intensity scintillation. This result is consistent with the theoretical predictions for Rayleigh statistics.

1. INTRODUCTION

Radio waves from satellites during their passage through the ionospheric irregularities of electron density suffer random phase changes. Due to the diffractive process, the emerging wavefront with random phase changes across it also develops amplitude fluctuations as it travels towards the observer on ground. The temporal variations of phase and amplitude, known as radio wave scintillations, are recorded by an observer on ground due to the relative motion of the observer, the satellite and the ionospheric irregularities.

The magnitude of phase and amplitude scintillations is primarily controlled by the observing geometry, the integrated electron density deviation of the irregularities, their anisotropy and spatial wavenumber spectrum. In addition, the magnitude of phase fluctuations (or rms phase deviation) measured on ground depends on the frequency of truncation corresponding to the data detrend interval. On the other hand, the intensity scintillation data are intrinsically Fresnel (or high pass) filtered by the diffraction process. The normalized second moment of signal intensity, termed as the S_4 index, provides a measure of the observed intensity scintillations (Briggs and Parkin, 1963). For practical purposes, weak or forward scattering occurs when $S_4 < 0.6$ and the onset of multiple scattering occurs when $S_4 \geq$

0.6. For systems applications, the depth of fading below the median level is used as a severity of scintillations. The S_4 index of 0.6 demarcating the weak and strong scatter corresponds to a fading level of 10 dB (Whitney, 1974).

The temporal structure of phase scintillations is primarily dictated by the data truncation frequency, mentioned earlier. Further, the phase scintillation structure fails to map the irregularity structure in the case of very strong scattering when cycle slips start to occur (Rino and Owen, 1984). The temporal structure of intensity scintillation, in the case of weak scattering, is dictated by the Fresnel frequency $v/\sqrt{2\lambda z}$, where λ is the radio wavelength, z the effective distance between the irregularity layer and the observer and v is the relative motion between the observer, the satellite and the irregularities. For strong scattering, the temporal structure of intensity scintillation does not map the irregularity wavenumber spectrum owing to the effects of refractive and/or diffractive scattering of the irregularities (Booker and MajidiAhi, 1981). Under these circumstances, the scintillation structure on ground develops scales finer than those present in the ionosphere, the exact relationship being governed by the strength of scattering and the form of irregularity wavenumber spectrum in the ionosphere.

Scintillations cause both enhancements and fadings about the median signal level. The fading level is of concern to systems as the performance is degraded when the fadings exceed the specified fade margin of the link. As such, strong scatter conditions causing deep fadings engage the attention of engineers. The degree of degradation depends not only on the magnitude of the fading relative to the margin, but the duration of fades, the rate of fading, the type of modulation and the criteria of acceptability. In short, the temporal structure of scintillations under strong scatter conditions needs to be documented. The information on the depth of scintillations coupled with a knowledge of the temporal structure of scintillations allows a systems of engineer to assess the probability of a loss of phase lock, increase of overall system noise and loss of messages.

On a global scale, the degradations due to scintillations are most serious for propagation paths which intercept the low latitude irregularity belt around the magnetic equator and the high latitude environment encompassing the auroral oval and the polar cap regions. In view of the increase of ionization density in the ionosphere under solar maximum conditions, the random phase changes introduced by the irregularities become most pronounced at this time which result in severe scintillation activity.

In this paper, a study of the temporal structure of scintillations is undertaken. The observations pertain to the severely disturbed environments of the equatorial, auroral and polar cap regions during the current solar maximum period. The signal sources used in the study are either stationary or near-stationary satellites so that the relative drift arises from the motion of the irregularities with respect to the observer on ground. The temporal structures are studied from the intensity and phase scintillation spectra and are quantified in terms of the coherence times of signal intensity and the corresponding complex signal.

2. OBSERVATIONAL MATERIAL

The paper is based on observations made in the equatorial, auroral and polar cap regions during the solar maximum period. Our station at Ascension Island performs round-the-clock multi-frequency (UHF and L-band) intensity scintillation observations in the equatorial region near the southern crest of the equatorial anomaly using transmissions from the Marisat satellite. At this location (geographic lat: 7.4°S; long: 14.4°W; dip latitude: 17°S), during the solar maximum period, the F-region ionization density attains high values in the post-sunset period which coincides with the generation of intense F-region irregularities. This combination produces one of the most intense irregularity environments and, as a result, Ascension Island records severe scintillation activity during solar maximum. This study also utilizes high latitude phase and amplitude scintillation observations made with 250-MHz phase coherent transmissions of the near-stationary polar beacon satellites with a computer-controlled receiver (Basu et al., 1985). The high latitude observing sites are located at Sondrestrom, Greenland (geographic lat: 67°N; long: 51°W; corrected geomagnetic latitude: 74°) corresponding to the dayside cusp and nightside auroral oval and Thule, Greenland (geographic lat: 77°N; long: 69°W ; corrected geomagnetic latitude: 85°) situated deep within the polar cap.

3. RESULTS

Figure 1 reproduced from Basu et al. (1987) illustrates the variation of the temporal structure of scintillations with the strength of scattering as observed at Ascension Island. At the highest frequency of 3954 MHz, the Marisat transmissions suffered weak scattering (S_4 = 0.31) and the power spectrum of scintillations can be simply related to the wavenumber spectrum of ionospheric irregularities. The irregularity environment was intense enough to cause strong scattering not only at 257 MHz but at 1541 MHz as well. It may be noted that the fading rate increases with decrease of frequency as a result of the

increase in the strength of scattering. The decrease in the coherence time of scintillations is reflected in a progressive broadening of the spectrum with a decrease in frequency.

The coherence time of intensity scintillation, τ_I, defined (Rino and Owen, 1981) as the 50% decorrelation time, may be expressed as,

$$\frac{<I(t)\,I(t+\tau_I)>-1}{<I(t)^2>} = 0.5 \qquad (1)$$

where I(t) is the signal intensity normalized to unity. In the limiting case of strong scattering, Rino and Owen (1981) have developed an explicit relationship between τ_I and the phase scintillation structure function defined by the phase spectral strength, T, which corresponds to the power spectral density (psd) of phase power spectrum at 1 Hz and the phase spectral slope p. The functional relationship is

$$\tau_I^{\,p-1} = K\,C(p)\,T \qquad (2)$$

where K is a constant, C(p) a function of phase spectral index, p, which is constrained to lie between 1 and 3.

Figure 2 shows a plot of the coherence time of intensity scintillations at 244 MHz, observed at Ascension Island, on a logarithmic scale against the phase spectral strength (decibels) for strong scintillations ($S_4 > 0.6$). The data were acquired with a phase-locked computer-controlled receiver using the 244 MHz phase coherent transmissions from the Fleetsatcom satellite. Figure 2 shows that, in the strong scatter regime, with increased phase spectral strength, the intensity scintillations develop finer structures of the wave field on the ground. The decrease of coherence time from about 1 sec to .05 sec occurs when the phase spectral strength increases from -20 dB to -2.5 dB. From a knowledge of ionospheric irregularity drift, the coherence times may be translated to coherence scales of the wave field on ground for applications to space diversity receiving systems. In the equatorial region strong scatter conditions are encountered in the early evening hours when the ionospheric drift is typically of the order of 200 ms^{-1} from the west to the east. Therefore, the coherence times of 1 sec to .05 sec for 244-MHz scintillations correspond to scales varying between 200 m and 10 m. Further, the best fit straight line in the scatter diagram of Figure 2 implies a phase spectral slope of -2.4 in accordance with Eq. (2).

Figure 3 shows two 164-sec samples of intensity and phase scintillation data obtained at Sondrestrom, Greenland using the 250-MHz transmissions from the polar beacon satellite. The data were acquired on 5-6 December 1988 when the station was located in the nightside auroral oval and the daily average sunspot number was 144. The power spectra of intensity and phase scintillations obtained by using the FFT algorithm and Blackman-Harris windows are also shown. The S_4 index of 0.3 and rms phase deviation of 3.8 radians implies that weak scatter conditions prevailed. The high frequency roll-off commences beyond the Fresnel frequency of 0.5 Hz. Figure 4 shows another data sample acquired after about one hour when auroral conditions prevailed. Intense scintillation activity yielding saturated intensity scintillations with $S_4 = 1.09$ and rms phase deviation of 17 radians were obtained. The roll-off portion of the intensity spectrum does not commence before 2 Hz and the consequent broadening of the intensity scintillation spectrum implies short coherence times under strong scatter conditions. It should be noted, however, that at auroral locations ionospheric drift may vary from 100 ms^{-1} to 1000 ms^{-1} which can transport even the weak scatter spectrum in Figure 3 to higher frequencies while the spectral slope is preserved. In Figure 4 the flat-topped intensity spectrum coupled with high S_4 index confirms that strong scatter conditions indeed prevailed. However, contrary to the equatorial case, the high frequency roll-off at auroral locations that control the coherence times depends both on the strength of scattering as well as the prevailing drift.

The above data set from Sondrestrom was used to obtain a plot (Figure 5) of coherence times of intensity scintillations at 250 MHz against phase spectral strengths for $S_4 > 0.6$. The decrease of coherence time with increase of phase spectral strength is observed as in Figure 2 except that coherence times less than 0.1 sec are not obtained. As demonstrated with Figure 2, the slope of the best-fit straight line through the scatter plot of Figure 5 was used to derive a phase spectral index of -4.3. The slopes of the weak scatter spectra shown in Figure 3 agree extremely well with this value.

It should be remarked that during this period, spaced receiver scintillation measurements indicated drift speeds in excess of 1 km/sec which is at least a factor of 5 larger than the equatorial drifts mentioned earlier. As a result, the coherence scales at auroral locations even under strong scatter conditions are much larger and do not approach values less than 100 m.

The data set from Sondrestrom was analyzed to obtain the complex coherence time. The phase scintillation data was filtered with a high pass filter having a cut-off frequency of 0.1 Hz. The filtered phase data and the intensity data were combined to generate the complex amplitude. The variation of

the coherence time of complex amplitude with phase spectral strength is similar to that of the intensity coherence time. Figure 6 shows the relationship between the coherence times of complex amplitude and intensity. We observe that the coherence times of complex amplitude asymptotically approach values equal to $\sqrt{2}$ times the intensity coherence times under intense scattering conditions. This follows the theoretical prediction for the case of Rayleigh scattering.

During the November, 1989 CEDAR-HLPS campaign in the solar maximum period, multi-technique measurements of ionospheric structures at high latitudes were performed. At Thule, Greenland, intensity and phase scintillation measurements were made using the 250-MHz transmissions from the polar beacon satellites. Long-lasting intense scintillation activity was recorded in the winter polar cap when large patches of structured F-region plasma intercepted the communication link. Figure 7 shows the transit of such an ionization patch through the raypath between 0930 UT and 1100 UT on November 4, 1989. The time variation of the S_4 index of scintillation, rms phase deviation over 82-second detrend interval and intensity coherence times are shown. The data set ($S_4 > 0.6$) acquired during this period was used to study the variation of intensity coherence time with phase spectral strength (Figure 8). The best fit straight line through the scatter plot implies a phase spectral index of -2.3. The spectral index is considerably smaller than that obtained in the auroral oval. This has considerable geophysical implications regarding the effects of auroral E-region ionization on F-region scintillation.

The final diagram (Figure 9) investigates the relationship between the coherence times of complex signal and intensity scintillation in the winter polar cap. It is observed that the relationship approaches the theoretical limit specified for Rayleigh scattering.

ACKNOWLEDGMENTS

We wish to thank C. Bryant of Boston College and K. Bounar of Radex Inc. for their help with signal analysis and development of computer algorithms. We thank the Danish Commission for Scientific Research in Greenland for permission to conduct ground experiments at Thule Air Base under Project 24-89. The work at AFGL was supported in part by Air Force Office of Scientific Research task 2310G9. The work at Boston College was supported by AFGL contract F19628-90-K-0007.

REFERENCES

Basu, S., S. Basu, J.P. McClure, W.B. Hanson, and H.E. Whitney (1983), High resolution topside in-situ data of electron densities and VHF/GHz scintillations in the equatorial region, *J. Geophys. Res., 88*, 7075.

Basu, S., S. Basu, E. MacKenzie, and H.E. Whitney (1985) Morphology of phase and intensity scintillations in the auroral oval and polar cap, *Radio Sci., 20*, 947.

Basu, S., E. MacKenzie, S. Basu, E. Costa, P.F. Fougere, H.C. Carlson, and H.E. Whitney (1987), 250 MHz/GHz scintillation parameters in the equatorial, polar and auroral environments, *IEEE Journal on Selected Areas in Communications, SAC-5*, 100.

Booker, H.G. and G. MajidiAhi (1981), Theory of refractive scattering in scintillation phenomena, *J. Atmos. Terr. Phys.*, 1199.

Briggs, B.H. and I.A. Parkin (1963), On the variation of radio star and satellite scintillation with zenith angle, *J. Atmos. Terr. Phys., 25*, 339.

Rino, C.L. and J. Owen (1981), On the temporal coherence loss of strongly scintillating signals, *Radio Sci., 16*, 31.

Rino, C.L. and J. Owen (1984), Numerical simulations of intensity scintillation using the power law phase screen model, *Radio Sci., 19*, 891.

Whitney, H.E. (1974), Notes on the relationship of scintillation index to probability distributions and their uses for systems design, *Rep. AFCRL-TR-74-0004, AD778092*, Air Force Cambridge Research Laboratories, Bedford, MA.

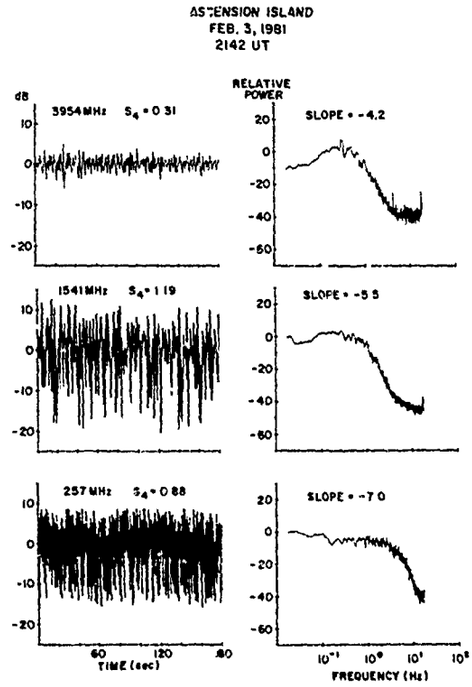


FIGURE 1. A 3-min data segment of scintillation observed at 3954, 1541 and 257 MHz at Ascension Island and their respective spectra.

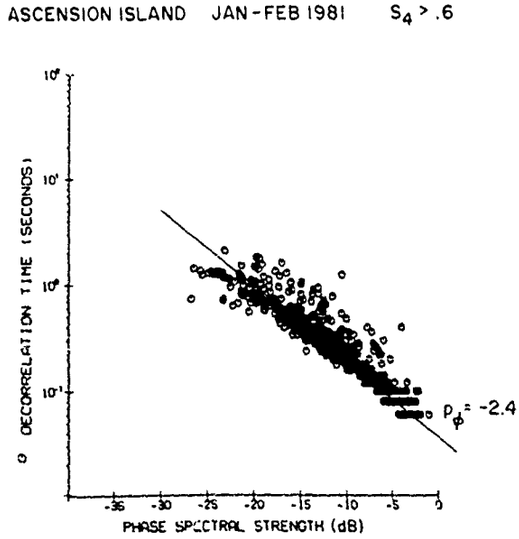


FIGURE 2. Scatter plot illustrates decrease of decorrelation time of intensity scintillation at 244 MHz with phase spectral strength observed at Ascension Island. The functional dependence of the two parameters is used to derive a theoretical estimate of phase spectral index $p_\phi = -2.4$.

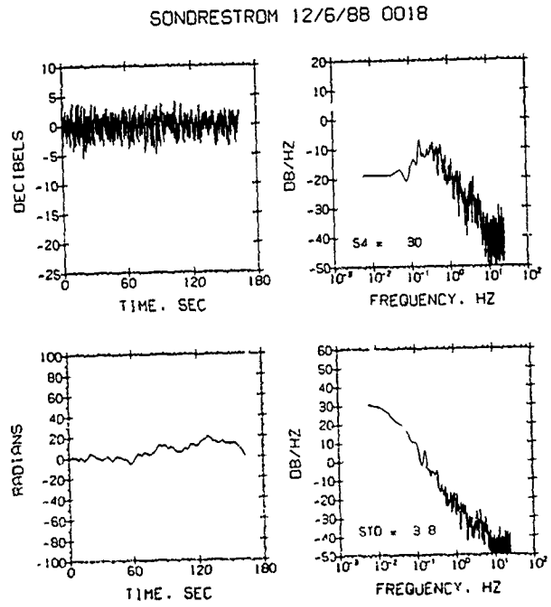


FIGURE 3. Weak intensity (S_4 < 0.6) and phase scintillation data recorded at 250 MHz at Sondrestrom, Greenland and their respective spectra.

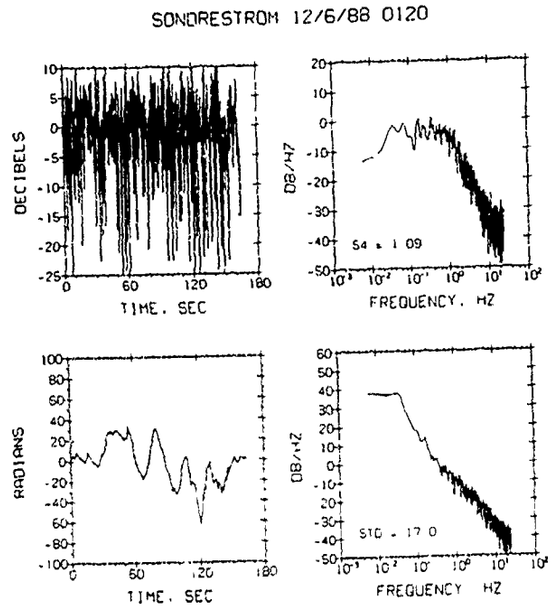


FIGURE 4. Strong intensity ($S_4 \geq 0.6$) and phase scintillation recorded at 250 MHz at Sondrestrom and their respective spectra.

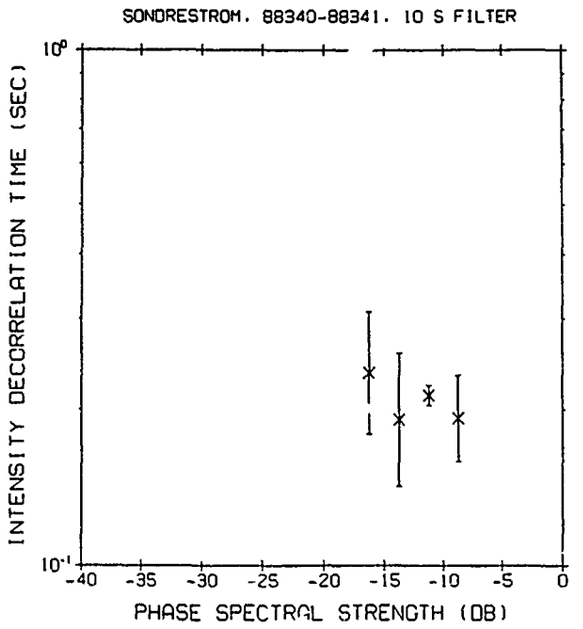


FIGURE 5. Histogram of intensity decorrelation time at 250 MHz for $S_4 > 0.6$ observed at Sondrestrom shown against phase spectral strength.

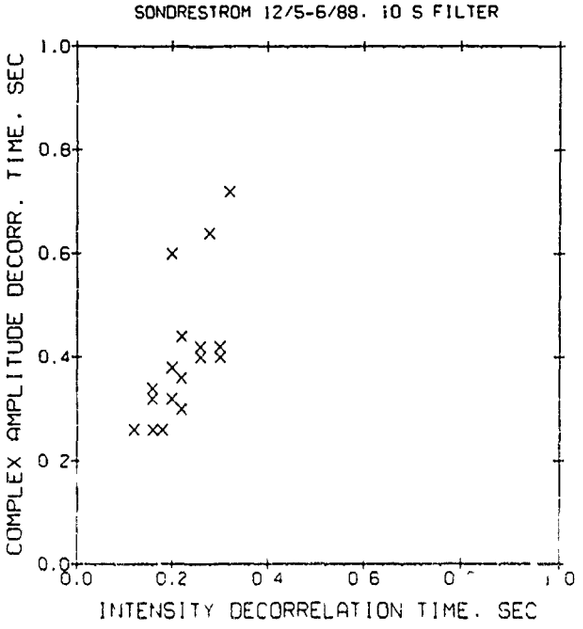


FIGURE 6. Dependence of decorrelation times of complex amplitude scintillation on decorrelation times of intensity scintillation observed at Sondrestrom.

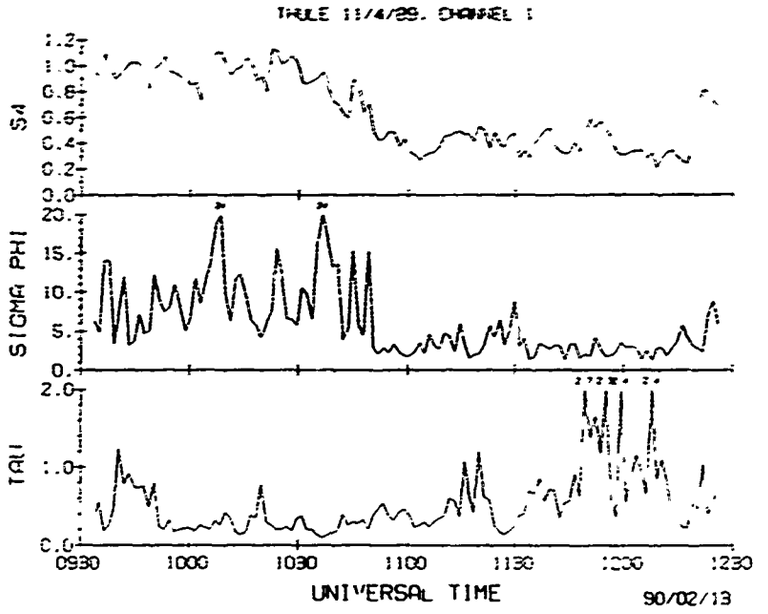


FIGURE 7. Illustrates the variation of intensity scintillation (S_4), rms phase deviation (SIGMA PHI) and decorrelation time (TAU) at 250 MHz observed at Thulc, Greenland during the encounter of a polar cap patch between 0930-1100 UT.

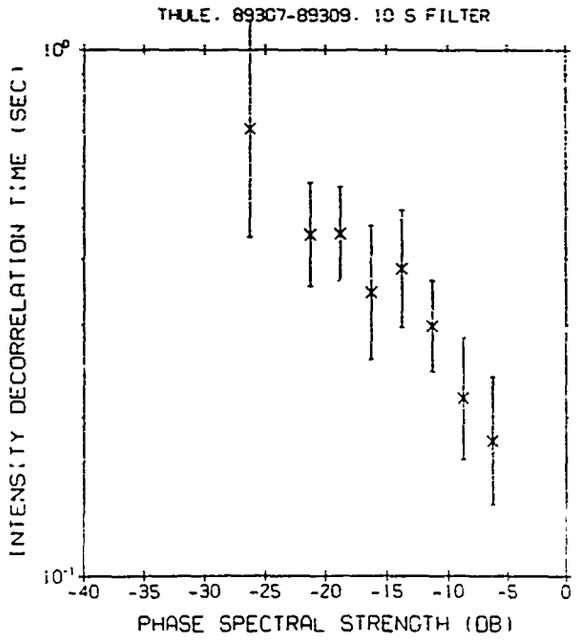


FIGURE 8. Histogram showing variation of decorrelation times of 250-MHz intensity scintillation with phase spectral strength for $S_4 > 0.6$ observed at Thule, Greenland.

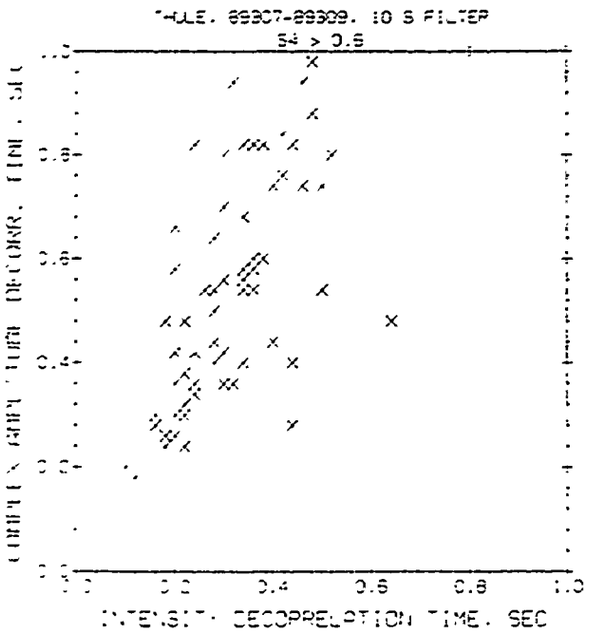


FIGURE 9. Dependence of decorrelation times of complex amplitude scintillation on decorrelation times of intensity scintillation observed at Thule.

AD-P006 279

THEORETICAL SCALING LAWS FOR THE SPECTRUM OF DENSITY IRREGULARITIES IN THE HIGH AND LOW LATITUDE IONOSPHERE

M.J. KESKINEN
SPACE PLASMA BRANCH
PLASMA PHYSICS DIVISION
NAVAL RESEARCH LABORATORY
WASHINGTON, DC 20375-5000

ABSTRACT

Constraints on and scaling laws associated with the spectrum of density irregularities in the high and low latitude ionosphere are derived using conservation laws implied by the fundamental nonlinear plasma fluid equations describing low frequency, long wavelength ionospheric plasma dynamics and structure. For the high latitude case we discuss the spectrum implied by interchange-like plasma instabilities and apply our results to convecting ionospheric plasma enhancements, blobs, patches, and polar cap arcs. For the low latitude ionosphere, we derive scaling laws for the density spectrum associated with the Rayleigh-Taylor instability and make applications to equatorial spread-F. For both the high and low latitude cases, we distinguish the spectral behavior to be expected in both the lower and upper ionosphere.

INTRODUCTION

In the last several years a considerable amount of both experimental [Basu et al., 1984; Basu et al, 1988; Gurnett et al., 1984; Kintner et al., 1987; Weimer et al., 1985; Vickrey et al., 1980; Vickrey et al., 1986; Cerisier et al., 1985; Baker et al., 1986; Curtis et al., 1982] and theoretical [Mitchell et al., 1985; Lysak and Carlson 1981; Lotko et al., 1987; Chaturvedi and Huba, 1987; Keskinen et al., 1988] research has been directed to the origin, modeling, and general interpretation of plasma dynamics and structure in the high latitude ionosphere. It is now known [for recent reviews, see Kintner and Seyler, 1985; Temerin and Kintner, 1988; and Tsunoda, 1988] that the high and low latitude ionosphere can be characterized as a turbulent and structured plasma containing density, electric field, and magnetic field fluctuations with scale sizes ranging from approximately hundreds of kilometers to centimeters. This structure can severely degrade C^3I systems operating in and through the ionosphere.

Several different source mechanisms [Kintner and Seyler, 1985; Temerin and Kintner, 1988; Tsunoda, 1988] have been proposed to account for high latitude magnetospheric and ionospheric plasma turbulence, e.g., electric fields, particle precipitation, and plasma instabilities. The relative importance of each mechanism depends on several factors, e.g., geomagnetic latitude, seasonal effects, and scale size perpendicular to the geomagnetic field. Ossakow and Chaturvedi [1979] proposed the current-convective instability, an interchange-like instability driven by field-aligned currents, to explain features of ionospheric irregularities occurring in the diffuse aurora. Keskinen and Ossakow [1982, 1983] have proposed the $\underline{E} \times \underline{B}$ drift instability to be a source of high latitude ionospheric plasma density and electric field fluctuations for scale sizes from tens of kilometers to meters. This interchange-like instability, a plasma analogue of the well-known Rayleigh-Taylor instability which results when a heavy fluid is supported against gravity by a lighter fluid, is driven by both a plasma density gradient and an electric field perpendicular to the geomagnetic field. The electric field will result in a relative velocity between convecting high latitude ionospheric plasma ions and the neutral thermosphere. The $\underline{E} \times \underline{B}$ drift instability has been applied to the small scale dynamics, stability and evolution of convecting large scale ionospheric density structures recently observed in the high latitude ionosphere [Vickrey et al., 1980; Muldrew and Vickrey 1982; Buchau et al., 1983, 1985; Weber et al., 1984, 1986; Foster and Doupnik, 1984; de la Beaujardiere et al., 1985]. Several studies [Sojka and Schunk, 1986; Schunk and Sojka, 1987; Anderson et al., 1987] have modeled the large scale, global features of these plasma density enhancements and structures. Vickrey and Kelley [1982] have considered the effects of a conducting E layer on the evolution of convecting large scale ionospheric density structures. At smaller scales, Huba et al [1983] have studied the linear theory of the $\underline{E} \times \underline{B}$ instability with an inhomogeneous electric field. Mitchell et al. [1985] have investigated the nonlinear evolution of the $\underline{E} \times \underline{B}$ instability in the high latitude ionosphere with magnetospheric coupling. They find that the

91-09671

primary effect of the magnetosphere is to incorporate inertia into the development of the $\underline{E} \times \underline{B}$ instability. In addition they find that plasma interchange-like instabilities in the high latitude ionosphere develop in a fundamentally different manner when magnetospheric coupling is included as opposed to the case when magnetospheric coupling is absent. Chaturvedi and Huba [1987] have added three-dimensional effects to the linear theory of the $\underline{E} \times \underline{B}$ instability in the high latitude ionosphere. Recent studies [Basu et al., 1990] have interpreted simultaneous density and electric field fluctuations in regions of ionospheric density gradients in and near convecting large scale structures in the polar ionosphere in light of an interchange $\underline{E} \times \underline{B}$ instability process. In addition, Weber et al [1989] have attributed the generation mechanism of electric field and density fluctuations on the duskside of a typical convecting polar cap arc to be due to an interchange-like instability. It is now generally believed [Temerin and Kintner, 1988; Tsunoda, 1988] that plasma interchange-like instabi``ties are a major source of density and electric field fluctuations in the high latitude ionosphere in the scale size regime of a few tens of kilometers to tens of meters perpendicular to the geomagnetic field.

However, the spectrum of density or electric field fluctuations associated with the nonlinear evolution of the $\underline{E} \times \underline{B}$ instability in the coupled magnetosphere-ionosphere system has not been analytically investigated in detail. In this study we derive constraints on and scaling laws associated with the nonlinear spectrum of the $\underline{E} \times \underline{B}$ interchange instability as it may occur in the high latitude ionosphere-magnetosphere coupled system. These constraints are derived by analysis of the fundamental nonlinear equations governing the $\underline{E} \times \underline{B}$ instability. We also apply these techniques to the low latitude ionosphere by analyzing another form of interchange instability, the Rayleigh-Taylor instability in equatorial spread F. We then discuss and summarize our results.

MODEL EQUATIONS AND ANALYSIS

We consider the high latitude ionospheric plasma to be a weakly ionized low $\beta = 8\pi n_e(T_e + T_i)/B_o^2$ plasma where n_e is the electron plasma density, $T_e(T_i)$ the electron (ion) temperature, and B_o the geomagnetic field. For our coordinate system we take the ambient density gradient to be in the x-direction, i.e., $n = n_o(x)$, the ambient electric field to be in the y-direction, i.e., $\underline{E} = E_o\hat{y}$, and the geomagnetic field to be in the z-direction $\underline{B} = B_o\hat{z}$. The equations used in this analysis are continuity, momentum, and charge conservation:

$$\frac{\partial n_\alpha}{\partial t} + \nabla\cdot n_\alpha \underline{V}_\alpha = 0 \tag{1}$$

$$0 = -\frac{e}{m_e}\left(\underline{E} + c^{-1}\underline{V}_e \times \underline{B}\right) - (T_e/m_e)(\nabla n/n) - \nu_{ei}(\underline{V}_e - \underline{V}_i) \tag{2}$$

$$\left(\frac{\partial}{\partial t} + \underline{V}\cdot\nabla\right)\underline{V}_i = \frac{e}{m_i}\left(\underline{E} + c^{-1}\underline{V}_i \times \underline{B}\right) - (T_i/m_i)(\nabla n/n) - \nu_{ie}(\underline{V}_i - \underline{V}_e) - \nu_{in}\underline{V}_i \tag{3}$$

$$\nabla\cdot\underline{J} = \nabla\cdot\left[n(\underline{V}_i - \underline{V}_e)\right] = 0 \tag{4}$$

Here, n_α is the electron (α=e) or ion (α=i) density, $\underline{V}_\alpha$ is the electron or ion velocity, $\underline{E}$ is the electric field, e is the electron charge, m_α is the electron or ion mass, c is the speed of light, T_α is the electron or ion temperature, $\nu_{ei}(\nu_{ie})$ is the electron-ion (ion-electron) collision frequency, and ν_{in} is the ion-neutral collision frequency.

We now perturb Eq. (1)-(4) about an equilibrium and let $n_\alpha = n_0 + \delta n$, $\underline{E} = \underline{E}_o - \nabla\delta\phi$, and $\underline{V}_\alpha = \underline{V}_{o\alpha} + \delta\underline{V}$ with δn, $\delta\underline{E}$, $\delta\underline{V} \propto \exp\,[i(k_x x + k_y y) - \omega t]$ where $(k_x^2 + k_y^2)^{1/2} L > 1$ and $L^{-1} = n_0^{-1}(\partial n_0/\partial x)$. After transforming to a reference frame moving with velocity $\underline{V}_o = (c\underline{E}_o \times \underline{B})/B^2 = (cE_o/B)\hat{x}$, Eqs. (2) and (3) yield.

$$\underline{V}_e = \underline{V}_{eo} + \delta\underline{V}_e \tag{5}$$

$$\underline{V}_i = \underline{V}_{io} + \delta\underline{V}_i \tag{6}$$

with

$$\underline{V}_{eo} = -(v_e^2/\Omega_e)(\partial \ln n_o/\partial x)\hat{y}$$

$$\delta\underline{V}_e = -\frac{c}{B}\nabla\,\delta\tilde{\phi} \times \hat{z} - \frac{\nu_{ei}}{\Omega_e}\frac{v_i^2}{\Omega_i}\left(1 + \frac{T_e}{T_i}\right)\nabla\frac{\delta n}{n_o}$$

$$\underline{V}_{io} = \left[\left(v_i^2/\Omega_i\right)(\partial \ln n_o/\partial x) + (\nu_i/\Omega_i)(cE_o/B)\right]\hat{y}$$

$$\delta\underline{V}_i = -\frac{c}{B}\nabla\delta\psi \times \hat{z} - \frac{\nu_i}{\Omega_i}\frac{c}{B}\nabla\delta\psi - \frac{\nu_{ie}}{\Omega_i}\frac{v_i^2}{\Omega_i}\left(1 + \frac{T_e}{T_i}\right)\nabla\frac{\delta n}{n_o} - \frac{c}{B}\frac{1}{\Omega_i}\left(\frac{\partial}{\partial t} + \underline{V}_{io}\cdot\nabla\right)\nabla\psi$$

where $\delta\tilde{\phi} = \delta\phi - (T_e/e)(\delta n/n_o)$, $\delta\psi = \delta\phi + (T_i/e)(\delta n/n_o)$, $\Omega_e(\Omega_i)$ is the electron (ion) gyrofrequency, and $v_e(v_i)$ is the electron (ion) thermal velocity. Mitchell et al [1985] has shown that electric fields generated by the interchange $\underline{E} \times \underline{B}$ instability in the high latitude ionosphere can map far into the magnetosphere on time scales comparable to the $\underline{E} \times \underline{B}$ instability growth time. As a result, a magnetic-field-line integrated model of the interchange $\underline{E} \times \underline{B}$ instability is appropriate for application to the high latitude, near-earth space plasma [Mitchell et al., 1985]. Using Eqs. (1) and (4) one can then write [Mitchell et al., 1985],

$$\frac{\partial \Sigma_p}{\partial t} + \nabla\cdot(\Sigma_p\underline{V}) = D\nabla^2\Sigma_p \tag{7}$$

$$\nabla\cdot\left[\Sigma_p\nabla\tilde{\phi} + C_m\left(\frac{\partial}{\partial t} + \underline{V}\cdot\nabla\right)\nabla\tilde{\phi} + \frac{T_i}{e}\nabla\Sigma_p\right] = 0 \tag{8}$$

$$\underline{V} = -\frac{c}{B}\nabla\tilde{\phi} \times \hat{z} \tag{9}$$

with

$$\Sigma_p = \int dz\,\frac{nec\nu_i}{B\Omega_i} = \int dz\,\sigma_p$$

$$C_m = \frac{1}{4\pi}\int dz\,(c^2/V_A^2)$$

with $V_A = B/(4\pi nm_i)^{1/2}$ the Alfvén speed, $D = (\nu_{ei}v_i^2/\Omega_e\Omega_i)[1+T_e/T_i]$, and $\underline{E} = -\nabla\tilde{\phi}$ with $\tilde{\phi} = \phi - (T_e/e)\ln n$. Within the context of this model, Eq. (8) then defines an effective frequency $\nu = \Sigma_p/C_m$. For fluctuations with characteristic frequency $\partial/\partial t \sim \gamma$, then ν/γ is essentially the flux-tube-integrated ratio of the Pedersen current in the ionosphere and polarization current in the magnetosphere. The quantity $\nu/\gamma < 1$ implies magnetospheric control for small Pedersen conductivities, large inertial capacitances, and large fluctuation growth rates. On the other

hand, $\nu/\gamma > 1$ implies ionospheric control for large Pedersen conductances, small inertial capacitances, and small fluctuation growth rates. In order to solve Eq. (7)-(8) one must specify σ_p and V_A along a specified flux tube. We make the assumption that the magnetosphere, to lowest order, is uniform and incompressible. As a result, a continuity equation for the plasma density in the magnetosphere is not necessary. We further take the Pedersen conductivity $\Sigma_p \propto \nu N$ and the inertial capacitance $C_m \propto N$ with $N = \int n\, dz$ the integrated plasma density. Eqs. (7)-(9) can be further simplified giving:

$$\frac{\partial N}{\partial t} + \nabla\cdot N\underline{V} = D\nabla^2 N \tag{10}$$

$$\nabla\cdot\left[\nu N\nabla\tilde{\phi} + n\left(\frac{\partial}{\partial t} + \underline{V}\cdot\nabla\right)\nabla\tilde{\phi} + (T_i/e)\nu\nabla N\right] = 0 \tag{11}$$

$$\underline{V} = -\frac{c}{B}\,\nabla\tilde{\phi}\times\hat{z} \tag{12}$$

Eq. (10-(12) are now a closed system in N and $\tilde{\phi}$.

Eq. (10)-(11) have been used, in the limit T_i, $T_e = 0$, to study the nonlinear evolution of the interchange $\underline{E}\times\underline{B}$ instability in the high latitude ionosphere by Mitchell et al. [1985]. They find that the $\underline{E}\times\underline{B}$ instability develops in a fundamentally different manner depending on whether the collisional ($\nu/\gamma > 1$) or inertial ($\nu/\gamma < 1$) limit is taken. Here γ is the linear growth rate of the $\underline{E}\times\underline{B}$ instability. We now proceed to derive scaling laws associated with spectrum of the interchange $\underline{E}\times\underline{B}$ instability in the inertial limit. Expanding $N = n_o(x) + \delta n(x,y)$ and $\tilde{\phi} = \phi_o + \delta\phi(x, y)$ we find

$$\frac{\partial\delta n}{\partial t} + \frac{c}{B}\,\hat{z}\times\nabla\delta\phi\cdot\nabla n_o + \frac{c}{B}\,\hat{z}\times\nabla\delta\phi\cdot\nabla\delta n = D\nabla^2\delta n \tag{13}$$

Multiplying Eq. (13) by δn and integrating over all x and y with $d^2x = dxdy$ we find

$$\frac{1}{2}\,\frac{\partial}{\partial t}\int d^2x\,(\delta n)^2 + \frac{c}{B}\int d^2x\,\delta n\,\hat{z}\times\nabla\delta\phi\cdot\nabla n_o$$
$$+ \frac{c}{B}\int d^2x\,\delta n\,\hat{z}\times\nabla\delta\phi\cdot\nabla\delta n - D\int d^2x\,\delta n\nabla^2\delta n = 0 \tag{14}$$

In Eq. (14), the second term on the left-hand side, resembling a source term, is proportional to the initial density gradient while the fourth term, resembling a sink, is proportional to the cross-field diffusion. The third term, which is nonlinear and represents a mode-mode coupling among the fluctuations, vanishes upon averaging over d^2x since

$$\int d^2x\,\delta n\,\hat{z}\times\nabla\delta\phi\cdot\delta n = \frac{1}{2}\int d^2x\,\hat{z}\times\nabla\delta\phi\cdot\nabla(\delta n)^2$$

$$= \frac{1}{2}\int d^2x\,\nabla\cdot\left[(\delta n)^2\hat{z}\times\nabla\delta\phi\right] = 0$$

since $(\delta n)^2 \to 0$ as $x, y \to \infty$ for a finite sized ionosphere where $\nabla\cdot[(\delta n)^2\,\hat{z}\times\nabla\delta\phi] = \hat{z}\times\nabla\delta\phi\cdot\nabla(\delta n)^2$. We Fourier expand δn and $\delta\phi$ as follows:

$$\begin{pmatrix}\delta n(x,y)\\ \delta\phi(x,y)\end{pmatrix} = \int d^2k\begin{pmatrix}\delta n_{\underline{k}}\\ \delta\phi_{\underline{k}}\end{pmatrix}\exp\, i\underline{k}\cdot\underline{x}$$

with $\underline{k} = k_x\hat{x} + k_y\hat{y}$. As a result, Eq. (14) can be written

$$\frac{1}{2}\frac{\partial}{\partial t}\int d^2k I_{\underline{k}} + i\frac{c}{B}\int d^2k\, \hat{z} \times \underline{k}\cdot\nabla n_o \delta n_{-\underline{k}}\delta\phi_{\underline{k}} + D\int d^2k k^2 I_{\underline{k}} = 0 \tag{15}$$

with $I_{\underline{k}} = |\delta n_{\underline{k}}|^2$. Eq. (11) gives

$$\delta\phi_{\underline{k}} = \left[n_o k^2(i\omega - \nu)\right]^{-1}\left[\nu\,(i\underline{k}\cdot\underline{E}o + (T_i/e)k^2)\right]\delta n_{\underline{k}} \tag{16}$$

which can be inserted into Eq. (15). We not that Eq. (15) using Eq. (16) can be analyzed either in the collisional ($\nu > \omega$) limit, as was done by Keskinen and Ossakow [1981], or the inertial ($\nu < \omega$) limit. In the inertial limit we have

$$\frac{1}{2}\frac{\partial}{\partial t}\int d^2k I_{\underline{k}} = \int d^2k \gamma_{\underline{k}} I_{\underline{k}} \tag{17}$$

with

$$\gamma_{\underline{k}} = \left[\left(\nu c/Bk^2 n_o\right)\left(\underline{k}\cdot\underline{E}_o\,\hat{z}\times\underline{k}\cdot\nabla n_o\right)\right]^{1/2} - \frac{Dk^2}{2} = \left(\nu cE_o\cos^2\theta/BL\right)^{1/2} - \frac{Dk^2}{2}$$

with $L^{-1} = (1/n_o)(\partial n_o/\partial x)$ and θ is the angle defined by $\underline{k}$ and $\underline{E}_o$.

Eq. (17) can be written in polar coordinates ($k^2 = k_x^2 + k_y^2$, $\theta = \tan^{-1} k_x/k_y$) assuming steady state conditions in the following manner:

$$\int_{k_{min}}^{k_c} dk\, k \int_0^{2\pi} d\theta\, \gamma_{\underline{k},g}\, I(k,\theta) = \int_{k_c}^{k_{max}} dk\, k \int_0^{2\pi} d\theta\, \gamma_{\underline{k},d} I(k,\theta) \tag{18}$$

where $\gamma_{k,g} = (\nu\cos^2\theta cE_o/BL)^{1/2}$ is the growth rate and $\gamma_{k,d} = -Dk^2$ is the damping rate. As a result, the spectral power generated in the unstable range of wave numbers between k_{min} and k_c, the critical wave number, is balanced by the power dissipated in the stable range of wave numbers between k_c and k_{max} under steady state conditions.

In the following we assume a general form for the density spectrum $I(k,\theta)$ and use Eq. (18) to find constraints on the parameters used to specify $I(k,\theta)$. To our knowledge, the steady state spectrum $I(k,\theta)$ of the $\underline{E}\times\underline{B}$ instability in the high latitude ionosphere, in the inertial regime, has not been studied in detail. Keskinen and Ossakow [1982, 1983] have studied aspects of the spectrum in the collisional regime. However, Hassam et al. [1986] have investigated the nonlinear spectrum of the gravitational Rayleigh-Taylor instability, a closely related instability to the $\underline{E}\times\underline{B}$ instability, in both the inertial and collisional regimes. Hassam et al. [1986] show that the spectrum in the inertial regime is more isotropic than in the collisional regime. We take, consistent with the computations of Hassam et al. [1986], the spectrum to be of the form

$$I(k,\theta) = I_o f(\theta)\,(1 + (k/k_o)^2)^{-(n+1)/2}$$

with $f(\theta) = \cos^m\theta$ and $m > 0$. Substituting $I(k,\theta)$ into Eq. (18) we find

$$\left(\frac{k_o}{k_{max}}\right)^{n-1} = \frac{3-n}{n-1}\,\frac{\Gamma\left(\frac{m+2}{2}\right)\Gamma\left(\frac{m+2}{2}\right)}{\Gamma\left(\frac{m+1}{2}\right)\Gamma\left(\frac{m+3}{2}\right)}\,\frac{\left(\nu cE_o/BL\right)^{1/2}}{Dk_{max}^2/2}$$

where $\Gamma(q)$ is the gamma function defined by

$$\Gamma(q) = \int_0^\infty dy\, y^{q-1}e^{-y}$$

and we have assumed $n > 1$. From Eq. (19) we see that $1 < n < 3$. We note that the scaling in the inertial limit for k_o as given by Eq. (19) is different from the collisional limit [Keskinen and Ossakow, 1981] where $(k_o/k_{max}) \propto (cE_o/BLD\, k_{max}^2)^{1/n-1}$.

In order to verify the spectral scaling laws outlined in this paper, density power spectra in or near convecting patches and plasma enhancements in the high latitude ionosphere must be measured. The spectral index, outer scale k_o, and inner scale k_{max} wavenumbers would then need to be computed. The scaling of k_o/k_{max} with the ambient electric field E_o in the neutral frame, say, could then be investigated. Taking $n = 2$, $m = 2$, $cE_o/B = 500$ m/s, $D = 1$ m^2/S, $L = 10$km and $2\pi/k_{max} = 10$ m [R.T. Tsunoda, 1989] we find from Eq. (19) $2\pi/k_o \approx 600$ m which is consistent with recent observations [R.T. Tsunoda, 1989].

The gravitational Rayleigh-Taylor instability, thought to be responsible for rising bubble and plumelike phenomena in the equatorial ionosphere, [Ossakow, 1981; Kelley and McClure, 1981] is analogous to the interchange $\underline{E} \times \underline{B}$ instability. The Rayleigh-Taylor instability can be described by equations identical to Eq. (1)-(4) except that the driving electric field $\underline{E}_o = E_o\underline{y}$ is replaced by gravity $\underline{g} = -g\underline{x}$. The formulae for k_o in Eq. (19) which are applicable to interchange instability due to the gravitational Rayleigh-Taylor instability may then be computed simply by taking $E_o \rightarrow Bg/c\nu_{in}$. Recently, experimental observations gathered during the PLUMEX rocket campaign [Rino et al, 1981; Kelley et al., 1982] have been further analyzed [R.C. Livingston, private communication]. These data indicate that the power spectrum of density fluctuations in equatorial Spread-F usually contains a break at a wavelength λ_B which is altitude dependent. The power spectrum is steeper (shallower) for wavelengths greater (less) than the break wavelength. Typical power spectra $P_k \propto k^{-n}$ have spectral indices $n = 1.5 - 3$ and are computed in the altitude interval of approximately 275-500 km. Furthermore, the data indicate [R.C. Livingston private communication] that λ_B is directly proportional to altitude, i.e., the break wavelength λ_B is large (small) at high (low) altitudes with $\lambda_B \approx 0.2 - 1$ km in the altitude interval $z \approx 275$-500 km. In addition, it was found [R. C. Livingston, private communication] that the altitude dependence given by ν_{in}, the ion-neutral collision frequency, provides a good fit to the data, i.e., $\lambda_B \propto \nu_{in}^{-1}$ or $k_B \propto \nu_{in}$ with $k_B = 2\pi/\lambda_B$. Using Eq. (19) and the previous results of Keskinen and Ossakow [1981] we can then write for the gravitational Rayleigh-Taylor instability, $(k_o/k_{max})^{n-1} \propto (g/L)^{1/2}/Dk_{max}^2$ in the inertial limit, i.e., $\nu_{in} < (g/L)^{1/2}$ while $(k_o/k_{max})^{n-1} \propto g/\nu_{in}LDk_{max}^2$ in the collisional limit $\nu_{in} > (g/L)^{1/2}$. These scalings can be rewritten as $(k_o/k_{max})^{n-1} \propto \gamma_G/\gamma_D(k_{max})$ where γ_G is the growth rate and $\gamma_D(k_{max})$ is the damping rate for mode k evaluated at $k = k_{max}$. As a result $k_o^I/k_o^C \propto \gamma_G^I/\gamma_G^C$ where $k_o^I(k_o^C)$ is the spectral break wavenumber in the inertial (collisional) limit and $\gamma_G^I = (g/L)^{1/2}$ is the growth rate in the inertial limit with $\gamma_G^C = g/\nu_{in}L$ the growth rate in the collisional limit. Since $\gamma_G^I < \gamma_G^C$, typically, $k_o^I/k_o^C < 1$, i.e., the spectral break wavelength $\lambda_o \propto 2\pi/k_o$ is larger in the inertial regime (high altitudes) than in the collisional regime (low altitudes) in agreement with the observations of Livingston et al. Furthermore, the scaling of the spectral break wavenumber with altitude $k_o^I/k_o^C \propto \gamma_G^I/\gamma_G^C \propto \nu_{in}$ is also consistent with the PLUMEX data [R.C. Livingston, private communication].

SUMMARY

We have derived constraints and scaling relationships for the power spectrum of the inertial $\underline{E} \times \underline{B}$ drift instability in the weakly nonlinear regime in the high latitude ionosphere by analyzing conservation laws implied by the fundamental plasma fluid equations. By assuming $I(k,\theta) \propto f(\theta)(1 + (k/k_o)^2)^{-(n+1)/2}$ where $I(k,\theta)$ is the two-dimensional density power spectrum of the $\underline{E} \times \underline{B}$ instability, $\cos\theta = \underline{k}\cdot\underline{E}_o$, $\underline{E}_o$ is the ambient electric field, $f(\theta) = \cos^m\theta$, and k_o is the outer scale we find that $1 < n < 3$. Furthermore, we derive the scaling of the ratio of outer (k_o) and inner (k_{max}) scales $k_o/k_{max} \propto (\nu cE_o/BLD^2k_{max}^4)^{1/2(n-1)}$ with k_{max} the inner scale. Using typical values of ν, E_o, L, D, and k_{max} we find values for k_o and n which are consistent with observations. Table I summarizes our results for the E x B interchange instability as applied to the high latitude ionosphere. Table 2 displays our results for the Rayleigh-Taylor interchange instability in the low latitude ionosphere.

ACKNOWLEDGMENTS: We thank S.L. Ossakow and J.D. Huba for useful discussions. This was supported by DNA and ONR.

References

Anderson, D.N., J. Buchau, and R.A. Heelis, Origin of density enhancements in the winter polar cap ionosphere in The Effect of the Ionosphere on Communication Navigation, and Surveillance Systems, ed. J.M. Goodman, U.S. Government Printing Office, Washington, DC, 1987.

Baker, K.B., R.A. Greenwald, A.D.M. Walker, P.F. Bythrow, L.J. Zanetti, T.A. Potemra, D.N. Hardy, F.J. Rich, and C.L. Rino, A case study of plasma processes in the dayside cleft, J. Geophys. Res. 91, 3030, 1986.

Basu, Su., S. Basu, E. MacKenzie, W.R. Coley, W.B. Hanson, and C.S. Lin, F-region electron density irregularity spectra near auroral acceleration and shear regions. J. Geophys. Res. 89, 5554-5564, 1984.

Basu, Su., S. Basu, E. MacKenzie, P.F. Fougere, W.R. Coley, N.C. Maynard, J.D. Winningham, M. Sugiura, W.B. Hanson and W.R. Hoegy, Simultaneous density and electric field fluctuation spectra associated with velocity shears in the auroral oval, J. Geophys. Res. 93, 115-136, 1988.

Basu, Su., S. Basu, E. MacKenzie, W.R. Coley, J.R. Sharber, and W.R. Hoegy, Plasma structuring by the gradient-drift instability at high latitudes and comparison with velocity-shear driven processes, J. Geophys. Res., 1990 (in press).

Buchau, J. B.W. Reinisch, E.J. Weber, and J.G. Moore, Structure and dynamics of the winter polar cap F region, Radio Sci. 18, 994, 1983.

Buchau, J., E.J. Weber, D.N. Anderson, H.C. Carlson, Jr., J.G. Moore, B.W. Reinisch, and R.C. Livingston, Ionospheric structures in the polar cap: their origin and relation to 250-MHz scintillation, Radio Sci. 20, 325, 1985.

Cerisier, J.C. J.J. Berthelier and C. Beghin, Unstable density gradients in the high latitude ionosphere, Radio Sci. 20, 755, 1985.

Chaturvedi, P.K. and J.D. Huba, The interchange instability in high latitude plasma blobs, J. Geophys. Res. 92, 32357, 1987.

Curtis, S.A. W.R. Hoegy, L.H. Brace, N.C. Maynard, M. Sugiura, and J.D. Winningham, DE-2 cusp observations: Role of plasma instabilities in topside ionospheric heating and density fluctuations, Geophys. REs. Lett. 9, 9997, 1982.

de la Beaujardiere, O., J.D. Craven, V.B. Wickwar, G. Caudal, J.M. Holt, L.A. Frank, L.H. Brace, D.S. Evans, and J.D. Winningham, Universal time dependence of nighttime F-region densities at high latitudes, J.Geophys. Res. 90, 4319, 1984

Foster, J.C. and J.R. Doupnik, Plasma convection in the vicinity of the dayside cleft, J. Geophys. Res. 89, 9107, 1984.

Gurnett, D.A., R.L. Huff, J.D. Menietti, J.L. Burch, J.D. Winningham, and S.D. Shawhan, Correlated low frequency electric and magnetic noise along the auroral field lines, J. Geophys. Res. 89, 8971, 1984.

Hassam, A.B. W. Hall, J.D. Huba, and M.J. Keskinen, Spectral characteristics of interchange turbulence, J. Geophys. Res. 91, 13513, 1986.

Huba, J.D., S.L. Ossakow, P. Satyanarayana, and P.N. Guzdar, Linear theory of the $\underline{E} \times \underline{B}$ instability with an inhomogeneous electric field, J. Geophys. Res. 88, 425, 1983.

Kelley, M.C. and J.P. McClure, Equatorial spread F: A review of recent experimental results, J. Atmos. Terr. Phys. 43, 427, 1981.

Kelley, M.C., R. Pfaff, K.D. Baker, J.C. Ulwick, R. Livingston, C. Rino, and R. Tsunoda, Simultaneous rocket probe and radar measurements of equatorial spread F-transitional and short wavelength results, J. Geophys. Res. 87, 1575, 1982.

Keskinen, M.J. and S.L. Ossakow, On the spatial power spectrum of the $\underline{E} \times \underline{B}$ gradient drift instability in ionospheric plasma clouds, J. Geophys. Res. 86, 6947, 1981.

Keskinen, M.J. and S.L. Ossakow, Nonlinear evolution of plasma enhancements in the auroral ionosphere, 1 Long wavelength irregularities, J. Geophys. Res. 87, 144, 1982.

Keskinen, M.J. and S.L. Ossakow, Nonlinear evolution of convecting plasma enhancements in the auroral ionosphere, 2, Small scale irregularities, J. Geophys. Res. 88, 474, 1983.

Keskinen, M.J., H.G. Mitchell, J.A. Fedder, P. Satyanarayana, S.T. Zalesak, and J.D. Huba, Nonlinear evolution of the Kelvin-Helmholtz instability in the high-latitude ionosphere, J. Geophys Res. 93, 137, 1988.

Kintner, P.M. and C.F. Seyler, The status of observations and theory of high latitude ionospheric and magnetospheric plasma turbulence, Space Sci. Rev., 41, 91, 1985.

Kintner, P.M., M.C. Kelley, G. Holmgren, H. Koskinen, G. Gustafsson, and J.LaBelle, Detection of spatial density irregularities with the Viking plasma wave interferometer, Geophys. Res. Lett. 14, 467, 1987.

Lotko, W. B.U. O. Sonnerup, and R.L. Lysak, Nonsteady boundary layer flow including ionospheric drag and parallel electric fields, J. Geophys. Res. 92, 8635, 1987.

Lysak, R.L. and C.N. Carlson, The effect of microscopic turbulence on magnetosphere-ionosphere coupling Geophys. Res. Lett., 8, 269, 1981.

Mitchell, H.G., J.A. Fedder, M.J. Keskinen, and S.T. Zalesak, A simulation of high latitude F layer instabilities in the presence of magnetosphere-ionosphere coupling, Geophys. Res. Lett. 12, 283, 1985.

Muldrew, D.B. and J.F. Vickrey, High latitude F-region irregularities observed simultaneously with ISIS-1 and the Chatanika radar, J. Geophys. Res. 87, 907, 1982.

Ossakow, S.L. and P.K. Chaturvedi, Current convective instability in the diffuse aurora, Geophys. Res. Lett., 6, 332, 1979.

Ossakow, S.L., Equatorial spread F - A review, J. Atmos. Terr Phys., 43, 437, 1981.

Rino, C.L., R.T. Tsunoda, J. Petriceks, R.C. Livingston, M.C. Kelley, and K.D. Baker, Simultaneous rocket-borne beacon and in situ measurements of equatorial spread F-intermediate wavelength results, J. Geophys. Res. 86, 2411, 1981.

Schunk, R.W. And J.J. Sojka, A theoretical study of the lifetime and transport of large ionospheric density structures, J. Geophys. Res., 92, 12343, 1987.

Sojka, J.J. and R.W. Schunk, A theoretical study of the production and decay of localized electron density enhancements in the polar ionosphere, J. Geophys. Res. 91, 3245, 1986.

Temerin, M. and P.J. Kintner, Review of ionospheric turbulence, Proc. Chapman Conference on Plasma Waves and Instabilities in Magnetospheres and Comets, Sendai, Japan, October 1987, AGU Monograph (in press).

Tsunoda, R.T. High-latitude F-region irregularities: a review and synthesis, Rev. Geophys. Space Phys., (in press), 1989.

Vickrey, J.F. and M.C. Kelley, The effects of a conducting E layer on classical F region cross-field plasma diffusion, J. Geophys. Res. 87, 4461, 1982.

Vickrey, J.F., C.L. Rino, and T.A Potemra, Chatanikal Triad observations of unstable ionization enhancements in the auroral F region, Geophys. Res. Lett. 7, 789, 1980.

Vickrey, J.F., R.C. Livingston. N.B. Walker, T.A. Potemra, R.A. Heelis, M.C Kelley and F.J. Rich, On the current-voltage relationship of the magnetospheric generator at intermediate spatial scales. Geophys. Res. Lett. 13, 495, 1986.

Weber, E.J., J. Buchau, J.G. Moore, J.R. Sharber, R.C. Livingston, J.D. Winningham, and B.W. Reinisch, F-layer ionization patches in the polar cap, J. Geophys. Res. 89, 1683, 1984.

Weber, E.J., J.A. Klobuchar, J. Buchau, H.C. Carlson, R.C. Livingston, O. de la Beaujardiere, M. McCready, J.E. Moore, and G.J. Bishop, Polar Cap F layer patches: Structure and dynamics, J. Geophys. Res. 91, 12121, 1986.

Weber, E.J., M.C. Kelley, J.O. Bellenthin, S. Basu, H.C. Carlson, J.R. Fleischman, D.A. Hardy, N.C. Maynard, R.F. Pfaff, P. Rodriguez, R.F. Sheehan, and M. Smiddy, Rocket measurements within a polar cap arc: plasma, particle, and electric circuit parameters, J. Geophys. Res. 94, 6692, 1989.

Weimer, D.R., C.K. Goertz, D.A. Gurnett, N.C. Maynard, and J.L. Burch, Auroral zone electric fields from DE 1 and and 2 at magnetic conjunctions, J. Geophys. Res. 90, 7479, 1985.

TABLE I: INERTIAL AND COLLISIONAL E × B INSTABILITY

Collisionality/Altitude	n	k_o	
Inertial/High [$\nu < 4cE_o/BL$]	$1 < n < 3$	$\left(\frac{k_o}{k_{max}}\right)^{n-1} = \frac{3-n}{n-1} \frac{\Gamma\left(\frac{m+2}{2}\right)\Gamma\left(\frac{m+2}{2}\right)}{\Gamma\left(\frac{m+1}{2}\right)\Gamma\left(\frac{m+3}{2}\right)}$	$\frac{(\nu cE_o/BL)^{1/2}}{Dk_{max}^2/2}$
Collisional/Low [$\nu > 4cE_o/BL$]	$1 < n < 3$	$\left(\frac{k_o}{k_{max}}\right)^{n-1} = \frac{3-n}{n-1} \frac{\Gamma\left(m+\frac{1}{2}\right)\Gamma\left(\frac{m+4}{2}\right)}{\Gamma\left(\frac{m+2}{2}\right)\Gamma\left(m+\frac{3}{2}\right)}$	$\frac{cE_o/BL}{Dk_{max}^2/2}$

TABLE II: INERTIAL AND COLLISIONAL RAYLEIGH-TAYLOR INSTABILITY

Collisionality/Altitude	n	k_o	
Inertial/High [$\nu_i < (g/L)^{1/2}$]	$1 < n < 3$	$\left(\frac{k_o}{k_{max}}\right)^{n-1} = \frac{3-n}{n-1} \frac{\Gamma\left(\frac{m+2}{2}\right)\Gamma\left(\frac{m+2}{2}\right)}{\Gamma\left(\frac{m+1}{2}\right)\Gamma\left(\frac{m+3}{2}\right)}$	$\frac{(g/L)^{1/2}}{Dk_{max}^2/2}$
Collisional/Low [$\nu_i > (g/L)^{1/2}$]	$1 < n < 3$	$\left(\frac{k_o}{k_{max}}\right)^{n-1} = \frac{3-n}{n-1} \frac{\Gamma\left(m+\frac{1}{2}\right)\Gamma\left(\frac{m+4}{2}\right)}{\Gamma\left(\frac{m+2}{2}\right)\Gamma\left(m+\frac{3}{2}\right)}$	$\frac{g/\nu_i L}{Dk_{max}^2/2}$

AD-P006 280

A GENERAL CHANNEL MODEL FOR RF PROPAGATION THROUGH STRUCTURED IONIZATION

Roger A. Dana
Mission Research Corporation
Santa Barbara, California

Leon A. Wittwer
Defense Nuclear Agency
Washington, D.C.

INTRODUCTION

Design and evaluation of radio frequency (RF) systems that must operate through ionospheric disturbances resulting from high altitude nuclear detonations requires an accurate channel model. Such a model can be used to construct realizations of the received signal for use in digital simulations of transionospheric communications links and radars or for use in hardware channel simulators.

It is well known that the first order statistics of the received RF signal after propagation through strongly scattering ionospheric disturbances are accurately described by the Rayleigh probability distribution for the amplitude and a uniform distribution for the phase. Equivalently, the two orthogonal components of the received electric field are independent, zero mean normal random processes. Given the second order statistics of the received signal, standard statistical techniques can be used to generate realizations of the channel impulse response function from which the received signal can be constructed.

The second order statistics of transionospheric RF scintillation are described by the two-position, two-frequency, two-time mutual coherence function of the received signal. The derivation of the mutual coherence function, which is a solution of Maxwell's equations, requires a model for the temporal and spatial variations of the electron density fluctuations in the ionosphere. Under Taylor's frozen-in hypothesis, the electron density fluctuations are described as a rigid structure that drifts past the line-of-sight. There is then a deterministic relationship between spatial and temporal fluctuations in the electron density. This model is accurate when the ionization has formed a thin layer of striations aligned with the geomagnetic field lines. Before striations have formed or when there are multiple scattering layers in the ionosphere with different relative velocities, a turbulent model may be more appropriate. In the fully turbulent case, the spatial and temporal fluctuations of the electron density are uncorrelated. Reality should lie somewhere between these two limiting models.

This paper describes a general model which varies smoothly between the frozen-in and turbulent models. The starting point for the general model is a description of the electron density fluctuations in the ionosphere that includes space-time correlation. Then the mutual coherence function for the general model is presented. Examples are given which illustrate the effects on the received signal due to the variation from the turbulent to frozen-in models.

91-09672

MUTUAL COHERENCE FUNCTION

The derivation of the mutual coherence function starts with Maxwell's equations from which the parabolic wave equation is derived. The parabolic wave equation can be solved to give the received electric field for a specific electron density distribution in the ionosphere. However, the electron density distribution is a random process so the received electric field is also a random process. The parabolic wave equation is therefore used to derive an equation for the two-position, two-frequency, two-time mutual coherence function of the electric field, $\Gamma(\Delta r,\Delta\omega,\Delta t)$. The solution of the differential equation for Γ, which is also a solution of Maxwell's equations, then provides a description of the second-order statistics of the received electric field. The Fourier transform of the mutual coherence function is the Generalized Power Spectral Density (GPSD) of the received signal.

Consider a monochromatic spherical wave with an electric field $E(r,\omega,t)$ which is a function of position r, carrier radian frequency ω, and time t. The wave originates from a transmitter located at $r = (0,0,-z_t)$ and propagates in free space in the positive z direction until it is incident on an irregularly ionized layer which extends from $0 < z < L$ and is infinite in the x-y plane. After emerging from the layer at $z = L$, the wave propagates in free space to a receiver located at $r = (0,0,z_r)$. This geometry is illustrated in Figure 1. As the wave propagates, its phase varies as $i\omega t - i\langle k\rangle z$, so we can write

$$E(r,\omega,t) = U(r,\omega,t)\exp\left[i\left(\omega t - \int_{-z_t}^{z}\langle k(z')\rangle\, dz'\right)\right] \tag{1}$$

where

$$\langle k\rangle^2 = \frac{\omega^2}{c^2}\left[1 - \frac{4\pi r_e c^2\langle n_e\rangle}{\omega^2}\right], \tag{2}$$

c is the speed of light in a vacuum, $\langle n_e\rangle$ is the mean free electron density in the ionosphere, and r_e is the classical electron radius. Under the small angle scattering assumption, it can be shown that U obeys the parabolic wave equation [Tatarskii 1971]:

$$\nabla_\perp^2 U - 2i\langle k\rangle\frac{\partial U}{\partial z} - \langle k\rangle^2\left[1 - \frac{4\pi r_e c^2 n_e(r,t)}{\omega^2}\right]U = 0 \tag{3}$$

where $n_e(r,t)$ is the free electron density which is a function of space and time.

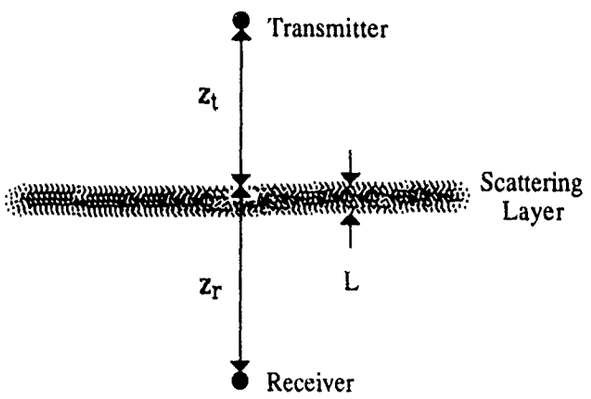


Figure 1. Propagation geometry.

The statistical nature of $n_e(r,t)$ requires a statistical approach to solving the parabolic wave equation. Thus a differential equation for the mutual coherence function,

$$\Gamma(\Delta r,\Delta\omega,\Delta t) = \langle U(r_1,\omega_1,t_1)U^*(r_2,\omega_2,t_2)\rangle \tag{4}$$

is derived from the parabolic wave equation assuming small angle scattering, wide-sense statistically stationary processes, and a model for the temporal fluctuations of n_e. The details of this derivation may be found in Tatarskii [1971] or Dana [1986].

The solution of the differential equation for the mutual coherence function depends on structure function $A(\rho)$ [Sreenivasiah, Ishimaru, and Hong 1976]:

$$A(\rho) = \int_{-\infty}^{\infty} \exp\,(i\mathbf{K}_\perp\cdot\rho)\, S_e(\mathbf{K}_\perp,K_z=0)\, \frac{d^2\mathbf{K}_\perp}{(2\pi)^2} \tag{5}$$

where ρ is a position vector in a plane normal to the line-of-sight with components x and y, and $S_e(\mathbf{K})$ is the three-dimensional power spectral density of the electron density fluctuations in the ionosphere. The structure function is then the correlation of the electron density fluctuations in a plane normal to the propagation direction. It is these fluctuations that cause diffraction and scintillation in the received signal. A power-law form for S_e is usually assumed [Wittwer 1979] which results in a Bessel function form for the structure function. Because the structure function occurs in the exponent of the solution for the mutual coherence function, the quadratic phase approximation is made to produce a mathematically tractable result. This quadratic phase approximation simplifies the expression for $A(\rho)$ to

$$A(\rho) = A_0\left[1 - A_2\left(\frac{x^2}{L_x^2} + \frac{y^2}{L_y^2}\right)\right]. \tag{6}$$

The constants A_0 and A_2 are determined by functional form of S_e and by the phase variance imparted on the wave in the ionosphere. The quantities L_x and L_y are electron density fluctuation scale sizes in the plane normal to the line-of-sight.

Although much information is available on the spatial fluctuations of n_e (see, for example, Wittwer 1979), there is little information on small time-scale fluctuations of n_e. Thus S_e does not, in general, contain Doppler frequency information, and a model is required to describe the temporal fluctuations. A general model for temporal dependence of the quadratic form of the structure function is

$$A(\rho,t) = A_0\left[1 - A_2\left(\frac{x^2}{L_x^2} + \frac{y^2}{L_y^2} + \frac{t^2}{T_0^2} - 2\frac{C_{xt}\,x\,t}{L_xT_0} - 2\frac{C_{yt}\,y\,t}{L_yT_0}\right)\right]. \tag{7}$$

The quantity T_0 can be thought of as a "decorrelation time" of the electron density fluctuations in the ionosphere. The degree of space-time correlation is determined by the coefficients C_{xt} and C_{yt}.

This form for the structure function reduces to the frozen-in model if, for example, C_{xt} is set to unity and C_{yt} is set to zero. Then the x dependence can be written as $(x-vt)^2/L_x^2$ where the velocity is equal to L_x/T_0. The turbulent model corresponds to the case where both C_{xt} and C_{yt} are zero.

Given a quadratic form for the structure function, a closed-form solution can be obtained for the mutual coherence function [Knepp 1983, Dana 1986]. This solution reduces to a mathematically tractable form when it is assumed that the scattering occurs in a thin layer. With these assumptions, the mutual coherence function for the general model is

$$\Gamma(x,y,\omega,t) = \frac{\exp\left[-\frac{1}{2}\left(\frac{\sigma_\phi \omega}{\omega_0}\right)^2\right] \exp\left[-(1 - C_{xt}^2 - C_{yt}^2)\left(\frac{t}{\tau_0}\right)^2\right]}{\left[1 + i\frac{\omega \Lambda_x}{\omega_{coh}}\right]^{\frac{1}{2}}\left[1 + i\frac{\omega \Lambda_y}{\omega_{coh}}\right]^{\frac{1}{2}}}$$

$$\times \exp\left[-\frac{\left(\frac{x}{\ell_x} - C_{xt}\frac{t}{\tau_0}\right)^2}{1 + i\frac{\omega \Lambda_y}{\omega_{coh}}}\right] \exp\left[-\frac{\left(\frac{y}{\ell_y} - C_{yt}\frac{t}{\tau_0}\right)^2}{1 + i\frac{\omega \Lambda_x}{\omega_{coh}}}\right] \quad (8)$$

where σ_ϕ is the standard deviation of the phase fluctuations imparted on the wave in the ionosphere, ω_0 is the angular frequency of the RF carrier, and

$$\Lambda_x = \left[\frac{2\ell_x^4}{\ell_x^4 + \ell_y^4}\right]^{\frac{1}{2}} \qquad \Lambda_y = \left[\frac{2\ell_y^4}{\ell_x^4 + \ell_y^4}\right]^{\frac{1}{2}} . \quad (9)$$

The decorrelation distances ℓ_x and ℓ_y are given by the expressions

$$\ell_x = \frac{(z_t + z_r)L_x}{z_t \sigma_\phi \sqrt{A_2}} \qquad \ell_y = \frac{(z_t + z_r)L_y}{z_t \sigma_\phi \sqrt{A_2}} , \quad (10)$$

the decorrelation time is given by the expression

$$\tau_0 = \frac{T_0}{\sigma_\phi \sqrt{A_2}} , \quad (11)$$

and the coherence bandwidth ω_{coh} is given by the expression

$$\omega_{coh} = \frac{\Lambda_y \omega_0^2 L_x^2}{2c\sigma_\phi^2 A_2} \frac{z_t + z_r}{z_t z_r} . \quad (12)$$

GENERALIZED POWER SPECTRAL DENSITY

The generalized power spectral density $S(\mathbf{K}_\perp,\tau,\omega_D)$ of the signal incident on the plane of the receiver is the Fourier transform of the mutual coherence function:

$$S(\mathbf{K}_\perp,\tau,\omega_D) = \int_{-\infty}^{\infty} d^2\rho \int_{-\infty}^{\infty} \frac{d\omega}{2\pi} \int_{-\infty}^{\infty} dt\, \Gamma(\rho,\omega,t) \exp\left[-i(\mathbf{K}_\perp \cdot \rho - \omega\tau - \omega_D t)\right] \quad (13)$$

where angle-of-arrival $\mathbf{K}_\perp$ is the Fourier transform pa . of position ρ in the x-y plane, delay τ is the Fourier transform pair of relative carrier frequency ω, and Doppler frequency ω_D is the Fourier transform pair of relative time t. The quantity $S(\mathbf{K}_\perp,\tau,\omega_D)(d^2\mathbf{K}_\perp/4\pi^2)d\tau(d\omega_D/2\pi)$ is equal to the mean signal power arriving with angles-of-arrival in the interval $\mathbf{K}_\perp/4\pi^2$ to $(\mathbf{K}_\perp+d^2\mathbf{K}_\perp)/4\pi^2$, with delays relative to a nominal propagation time in the interval τ to $\tau+d\tau$, and with Doppler frequencies in the interval $\omega_D/2\pi$ to $(\omega_D+ d\omega_D)/2\pi$. The delay dependence of the GPSD is a consequence of the fact that some of the signal energy takes a dog-leg path through the ionosphere from the transmitter to the receiver and arrives later than the signal energy that propagates straight through the ionosphere.

In general, the GPSD can be written as the product of a Doppler spectrum $S_D(\omega_D)$ and an angle-delay spectrum $S_{K\tau}(\mathbf{K}_\perp,\tau)$:

$$S(\mathbf{K}_\perp,\tau,\omega_D) = S_D(\omega_D)\ S_{K\tau}(\mathbf{K}_\perp,\tau)\ . \tag{14}$$

After performing the integrals indicated in Equation 13, the Doppler spectrum for the general model is

$$S_D(\omega_D) = \frac{\sqrt{\pi}\tau_0}{\sqrt{1 - C_{xt}^2 - C_{yt}^2}} \exp\left[-\frac{(\tau_0\omega_D - C_{xt}K_x\ell_x - C_{yt}K_y\ell_y)^2}{4(1 - C_{xt}^2 - C_{yt}^2)}\right] \tag{15}$$

In terms of the components of $\mathbf{K}_\perp$ (K_x and K_y), the angle-delay part of the GPSD is

$$S_{K\tau}(K_x,K_y,\tau) = \left[\frac{\pi}{2}\right]^{\frac{1}{2}} \ell_x\ell_y\,\alpha\omega_{coh} \exp\left[-\frac{K_x^2\ell_x^2}{4} - \frac{K_y^2\ell_y^2}{4}\right]$$

$$\times \exp\left\{-\frac{\alpha^2}{2}\left[\omega_{coh}\tau - \frac{\Lambda_y(K_x^2+K_y^2)\ell_x^2}{4}\right]^2\right\} \tag{16}$$

where the delay parameter α is defined to be

$$\alpha = \frac{\omega_0}{\sigma_\phi\omega_{coh}}\ . \tag{17}$$

The value of α is quite large under strong scattering conditions. The components of $\mathbf{K}_\perp$ are related to the scattering angles θ_x and θ_y about the x and y axis respectively by the relations

$$K_x = \frac{2\pi\sin(\theta_x)}{\lambda} \qquad K_y = \frac{2\pi\sin(\theta_y)}{\lambda}\ . \tag{18}$$

It should be noted that the range of delay in Equation 16 is from $-\infty$ to $+\infty$. However, this delay is relative to some nominal propagation time, and the value of the GPSD rapidly approaches zero for decreasing negative values of $\omega_{coh}\tau$. In the limit that α is infinity, the GPSD is non-zero only for positive values of delay. Thus Equation 16 presents no real problem with causality.

It is interesting to examine the limits of the general model Doppler spectrum in order to show that this model does indeed encompass both the frozen-in and fully turbulent models. These limits for the Doppler spectra are:

$$\underset{\substack{C_{xt}\to 1\\ C_{yt}\to 0}}{\text{Limit}}\ S_D(\omega_D) = 2\pi\,\tau_0\,\delta(\tau_0\omega_D - K_x\ell_x) \qquad \text{(Frozen-in Model)} \tag{19}$$

and

$$\underset{\substack{C_{xt}\to 0\\ C_{yt}\to 0}}{\text{Limit}}\ S_D(\omega_D) = \sqrt{\pi}\,\tau_0 \exp\left[-\frac{\tau_0^2\omega_D^2}{4}\right]\ . \qquad \text{(Turbulent Model)} \tag{20}$$

For the frozen-in model, the delta-function relationship between Doppler frequency and K_x is what is obtained by assuming that the random diffraction pattern of the signal is "frozen" and drifts in the x direction past the receiver. For the turbulent model, the Doppler spectrum is independent of $\mathbf{K}_\perp$ so the temporal and spatial variations in the received signal are also independent.

When the two decorrelation distances are equal ($\ell_x = \ell_y = \ell_0$), the scattering is isotropic about the line-of-sight, and the GPSD takes a somewhat simpler form. A three-dimensional plot of the isotropic one-dimension angle-delay part of the GPSD (Equation 16 integrated over one component of $\mathbf{K}_\perp$) is shown in Figure 2. This plot shows the mean received power density as a function of normalized angle $K\ell_0$ and normalized delay $\omega_{coh}\tau$. The vertical axis is linear with arbitrary units.

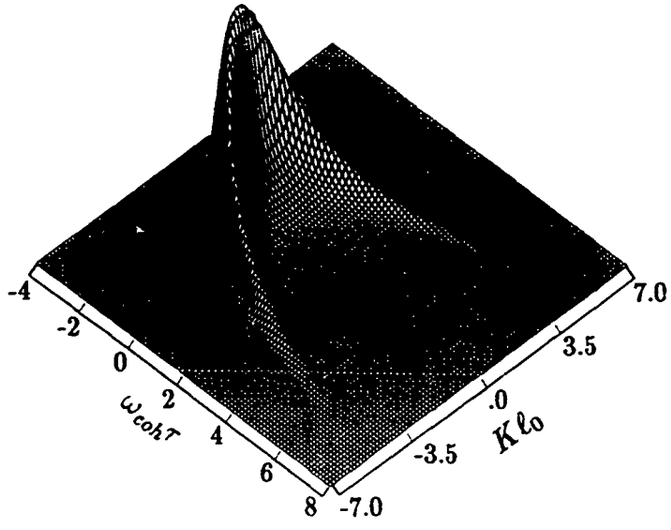


Figure 2. Angle-delay generalized power spectral density.

It can be seen that the power arriving at large angles is also the power arriving at long delays. The power arriving at long delays thus has higher spatial frequency components than power arriving at short delays. When there is strong space-time correlation (i.e. when $\sqrt{C_{xt}^2 + C_{yt}^2}$ is approximately equal to unity) these higher spatial frequency components correspond to higher Doppler frequency components. The signal arriving at long delays then varies more rapidly in time than the signal arriving at short delays.

Another view of the GPSD can be obtained by considering the delay-Doppler scattering function (Equation 14 integrated over $\mathbf{K}_\perp$). For the turbulent model, the scattering function is separable into a function of Doppler frequency times a function of delay. This is not the case for the frozen-in model.

A comparison of the scattering functions for the frozen-in and turbulent models is shown in Figure 3. The frozen-in scattering function is just a reproduction of Figure 2 with normalized angle $K\ell_0$ replaced with normalized Doppler frequency $\tau_0\omega_D$. This is a consequence of the delta-function relationship between angle and Doppler frequency for the frozen-in model. For this model the signal at long delays has correspondingly large Doppler shifts, and a wing-like structure is seen in the scattering function. The turbulent model scattering function does not exhibit these Doppler wings because the Doppler spectrum is the same at all delays. Both functions have exactly the same power density at each delay. The difference in appearance of the figures is due to the fact that the turbulent model signal at long delay is more spread out in Doppler frequency and is therefore less obvious.

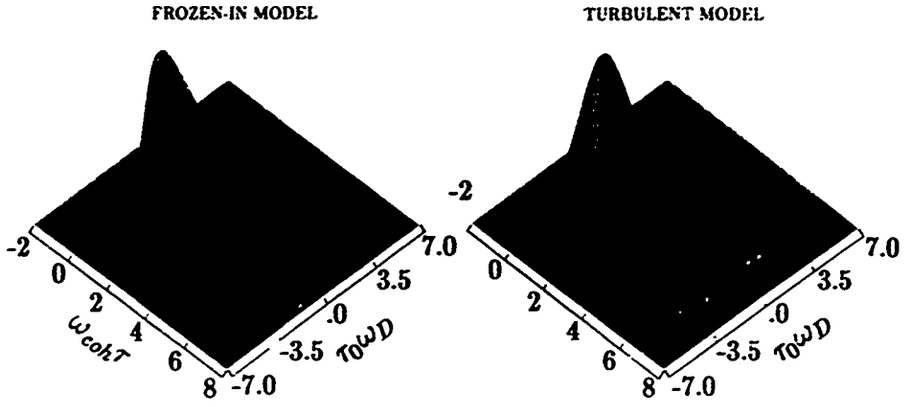


Figure 3. Comparison of scattering functions for the frozen-in and turbulent models.

A progression of scattering functions for the general model is shown in Figure 4. The space-time correlation coefficient C_{xt} varies from 0.99 for the scattering function in the upper left to 0.7 for the scattering function in the lower right. The scattering function for C_{xt} equal to 0.99 is essentially identical to that for the frozen-in model, and the scattering function for C_{xt} equal to 0.7 is essentially identical to that for the turbulent model. For intermediate values of C_{xt}, the scattering functions still exhibit Doppler wings but the wings have broader Doppler spectra as C_{xt} decreases.

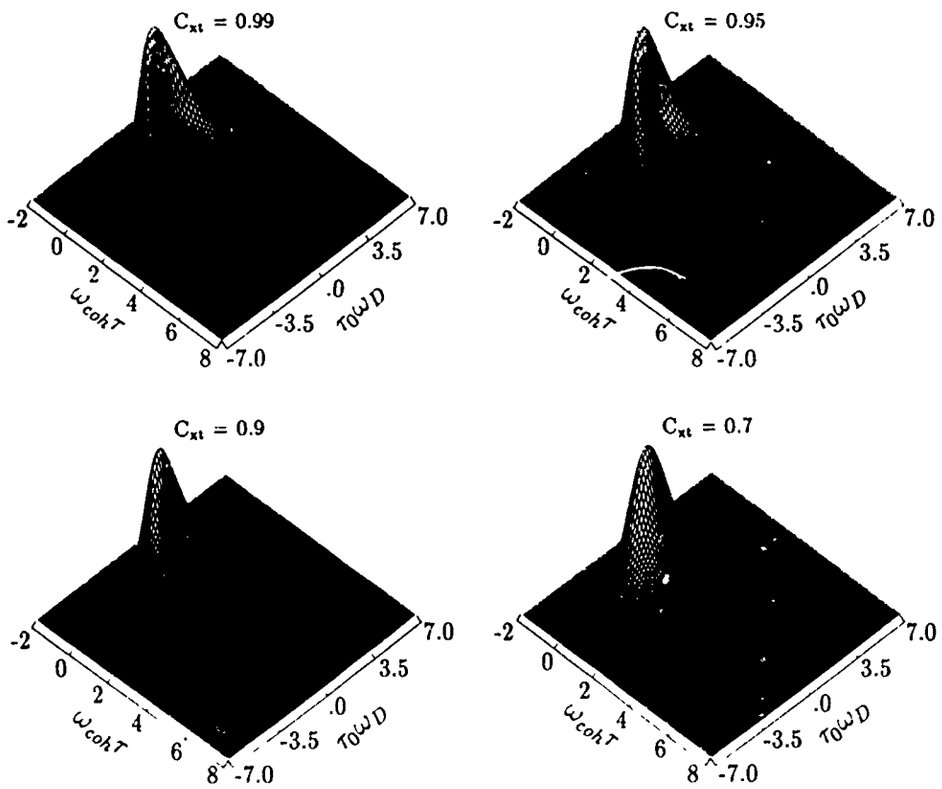


Figure 4. Scattering functions for the general model.

MATCHED FILTER EXAMPLE

Signal scintillation can be either flat fading (nonselective) or frequency selective, depending on the signal bandwidth and the degree ionospheric disturbance. If all frequency components of the received signal vary essentially identically with time, the propagation channel is referred to as nonselective or flat fading. When the scintillations exhibit statistical decorrelation at different frequencies within the signal bandwidth, the channel is referred to as frequency selective. The level of frequency selective scintillation disturbance is measured by the frequency selective bandwidth, f_0. This parameter is defined as

$$f_0 = \frac{1}{2\pi\sigma_\tau} = \frac{\omega_{coh}}{2\pi} \tag{21}$$

where σ_τ is the standard deviation of the time-of-arrival jitter.

Techniques have been developed to generate representative signal structures using statistical channel simulators [Wittwer 1980 and Dana 1986]. The latter reference describes techniques for generating realizations of the channel impulse response function, and describes how these realizations can be used to generate the output of a matched filter. A report is currently in preparation [Dana 1990] that extends these techniques to inlcude the general model.

A comparison of the matched-filter output amplitude generated with the frozen-in, general, and turbulent models is shown in Figure 5 for one level of frequency selective propagation disturbances, characterized by the ratio of the frequency selective bandwidth f_0 to the transmitted chip rate R_c (f_0/R_c = 0.1 for these examples). Each frame provides a three dimensional picture of the matched filter output for a single transmitted pulse as a function of time delay (horizontal scale) and time (scale directed into the figure). The total duration of each of the frames is 10 decorrelation times.

The top frame shows an example generated using the frozen-in model (C_{xt} = 1 and C_{yt} = 0). An effect of the frozen-in model that is evident in the top frame is that the signal arriving at long delays varies more rapidly in time than the signal arriving at shorter delays. The middle frame is a general model realization (C_{xt} = 0.9 and C_{yt} = 0), and the bottom frame is for the turbulent model (C_{xt} = C_{yt} = 0). The difference between the top and bottom frames is that the turbulent model amplitude has the same fading rate at all delays. It can be seen that the general model realization falls somewhere between these two limiting cases.

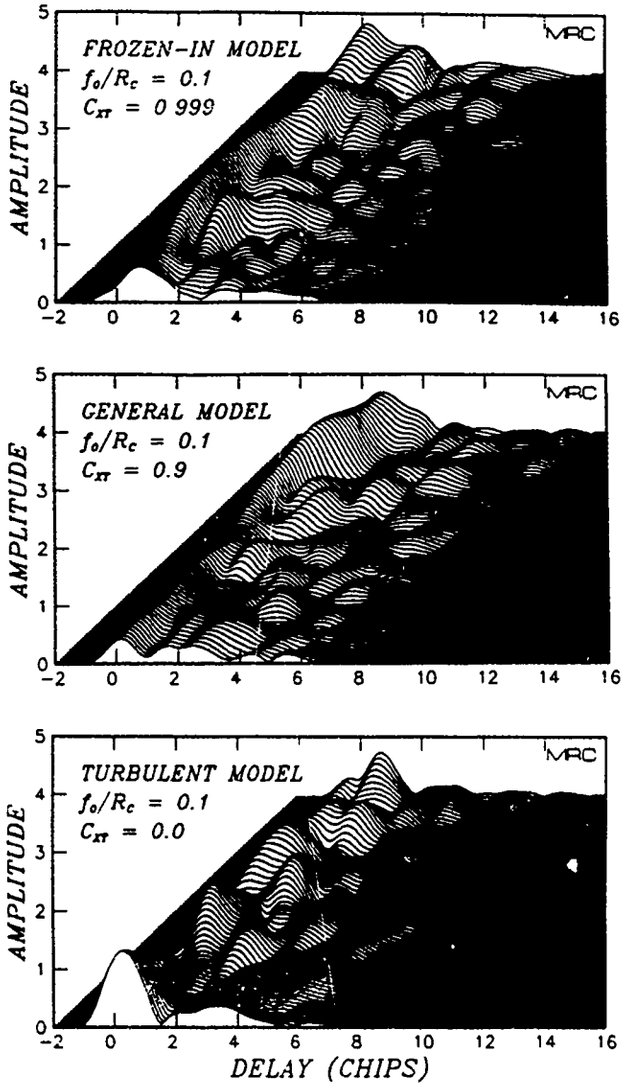


Figure 5. Matched filter output amplitude time histories for frequency selective channels.

ACKNOWLEDGEMENT

This work supported by the Defense Nuclear Agency under contract DNA 001-87-C-0169.

REFERENCES

Dana, R. A., *Propagation of RF Signals Through Structured Ionization. Theory and Antenna Aperture Effect Applications*, DNA-TR-86-158, MRC-R-976, Mission Research Corporation, May 1986.

Dana, R. A., *Propagation of RF Signals Through Structured Ionization. The General Model*, MRC-R-1262, Mission Research Corporation, (in preparation) 1990.

Knepp, D. L., "Analytic Solution for the Two-frequency Mutual Coherence Function for Spherical Wave Propagation," *Radio Science*, Vol. 18, No. 4, pp. 535-549, July 1983.

Sreenivasiah, I., A. Ishimaru, and S. T. Hong, "Two-Frequency Mutual Coherence Function and Pulse Propagation in a Randon Medium: An Analytic Solution to the Plane Wave Case," *Radio Science*, Vol. 11, No. 10, pp. 775-778, October 1976.

Tatarskii, V. I., *The Effects of the Turbulent Atmosphere on Wave Propagation*, translated by Isreal Program for Scientific Translations, National Technical Information Service, U.S. Department of Commerce, 1971.

Wittwer, L. A., *Radio Wave Propagation in Structured Ionization for Satellite Applications*, DNA 5304D, Defense Nuclear Agency, January 1979.

Wittwer, L. A., *A Trans-Ionospheric Signal Specification for Satellite C^3 Applications*, DNA 5662D, Defense Nuclear Agency, December 1980.

AD-P006 281

SYSTEM ISSUES RELATED TO SATELLITE COMMUNICATIONS IN A NUCLEAR ENVIRONMENT

Per A. Kullstam
PAIRCOM, Inc.
8508 Shadeway Place
Springfield, Va. 22153.

ABSTRACT

Nuclear induced signal scintillation effects are of great importance in design and deployment of military satellite systems that must provide survivable and enduring communications service. The induced scintillation will result in Rayleigh signal fading with limited signal decorrelation time and coherent bandwidth of the transmission channel as well as reduced signal power due to terminal antenna scattering loss.

In this environment the coherent bandwidth and signal decorrelation time are most important design parameters for modulation subsystem design. The antenna scattering loss is important for link power budgets and satellite network loading.

1. BACKGROUND

A high altitude nuclear detonation emits atomic particles and high energy photons, x-rays and γ-rays, that cause strong ionization of the gas in atmosphere and ionosphere. At lower altitudes the gas pressure is sufficiently high to cause very rapid recombining and annihilation of the charged particles while in the ionosphere, at the altitudes of about 100 to 1000 Km, the ionization will persist for substantial periods of time and affect signal propagation from minutes to hours and even days.

A hostile adversary may therefore elect to greatly increase the number of electrons and charged particles (ions) in the ionosphere to such a degree that it will cause multipath scattering effects of satellite communications signals passing through it. Strong ionization levels will affect and may even completely disrupt the satellite communications signal if this situation is not guarded against by proper design of the modulation subsystems and sufficient link power budget margins are used in engineering the communications networks.

The net result of the induced signal multipath propagation effects is that the signal power does not reach the terminal from the axial direction of the satellite but within an angular distribution about this "line of sight" direction. The angular distribution of signal components making up the composite signal has been determined to be approximately Gaussian [1]. That is, the probability density

$$p(\theta_x,\theta_y) = [1/(2\pi\sigma_x\sigma_y)]\exp[-(\theta_x/\sigma_x)^2/2-(\theta_y/\sigma_y)^2/2], \qquad 1\text{-}1$$

where θ_x, θ_y represent the angular deviation from the direct path and σ_x^2, σ_y^2

91-09673

are the variances of the angular deviations. The probability density is normalized to unit "mass" which is very appropriate to a direct description of the forward scattering situation since there is no loss of average signal power through the scintillation medium but rather an angular re-distribution of the total signal power into components by the multipath scattering process. Simply, forward scattering implies that the transmitted signal from a satellite will "sooner or later" reach the earth but not necessary through the direct path. Therefore, we may determine the average signal power at an earth terminal by first determining the power level in the case of no signal scintillation (the normal case) and apply the angular distribution of (1-1) to this power and determine the signal power level received at the earth terminal when signal scintillation is present. Note, even though we have referred to a satellite downlink in developing the rationale for the probabilistic description of the nuclear induced signal scintillation channel, it applies equally well to an uplink signal by the virtue of the general law of transmission reciprocity between transmit and receive terminals.

The angular distribution of signal components from the scintillation medium will generally not be rotationally symmetric. That is, the variances σ_x^2, σ_y^2 may not be equal. It is due to the influence of the earth's magnetic field on the moving charged particles which cause them "gyrate" about the magnetic field lines. Thus, the scattering effect will be accentuated in the direction along these lines and less so perpendicular to them. The x and y axes in (1-1) are chosen in such a way that they line up with the major and minor axes of the angular signal distribution. From theoretical considerations the two variances are related to the maximum medium angular scattering variance σ_θ^2 as

$$\sigma_x^2 = \sigma_\theta^2/K(\Phi), \qquad \sigma_y^2 = \sigma_\theta^2/K^3(\Phi), \tag{1-2}$$

where $K^2(\Phi) = \cos^2(\Phi)+225\sin^2(\Phi) = 1+224\sin^2(\Phi)$, and thus

$$\sigma_y^2 = \sigma_x^2/K^2(\Phi), \tag{1-3}$$

where Φ is the signal penetration angle through the scintillation medium relative to the magnetic field. [The angle $\Phi = 0$ if the propagation path is parallel to the magnetic field lines.] As $1 \leq K(\Phi) < 15$ with $K(\Phi) = 1$ for $\Phi = 0$, we have in this case $\sigma_y^2 = \sigma_x^2 = \sigma_\theta^2$ and a symmetric angular scattering distribution. This case is almost possible for a terminal located due north or south of a geosynchronous satellite. For a terminal at the equator the penetration angle $\Phi = \pi/2$ (90 degrees) and we have the most nonsymmetric case with $\sigma_y^2 = \sigma_x^2/225$ and $\sigma_x^2 = \sigma_\theta^2/15$, which shows that the angular scattering dispersion is strongly reduced, by a factor of 15, from the maximum medium angular scattering in the scintillating volume.

2. RECEIVED SIGNAL FADE CHARACTERISTICS

There are several important factors to consider for satellite communications network design and link power budget determination to achieve the required communications performance under nuclear stressed situations. Specifically, the effects of Rayleigh signal fading, signal decorrelation time, channel coherent bandwidth and terminal antenna scattering loss will be considered.

The well developed multipath situation caused by nuclear induced signal scintillation implies that both the inphase [$\sqrt{S}x(t)$] and quadrature phase [$\sqrt{S}y(t)$] components of a signal will be uncorrelated Gaussian distributed random processes with zero mean and equal variance. This situation is referred to as Rayleigh fading which is alternatively characterized by a uniform

phase distribution and an exponential received power probability density

$$p(S(t)|S) = (1/S)\exp[-S(t)/S], \qquad 2\text{-}1$$

where $S(t) = S[x^2(t)+y^2(t)]$ is the instantaneous received power level and S denotes the average received signal power level over time. [Generally, the amplitude probability density $p(a|S) = (2a/S)\exp(-a^2/S)$ is referred to as the Rayleigh distribution which is equivalent to (2-1) with $a = \sqrt{S(t)}$.] In other words, instead of receiving the power level $S(t) = S$ at all times, the receive power level will be random at any given time instant governed by (2-1) and also varying with time. The time varying nature of the signal is fully defined statistically by the signal auto-correlation function which from the Gaussian nature of the angular scattering distribution as well as from the gain pattern of a circular antenna can be expressed as

$$\begin{aligned} R(\tau) &= E[x(t)x(t')]+E[y(t)y(t')] \\ &= \exp[-(\tau/\tau_o)^2], \quad \tau = t-t', \end{aligned} \qquad 2\text{-}2$$

where also $E[x(t)x(t')] = E[y(t)y(t')]$. The parameter τ_o is referred to as the decorrelation time and signifies that the signal values separated in time by $\tau = \tau_o$ for each of the quadrature components have a correlation equal to exp(-1). The Rayleigh fading situation is completely characterized by the average received power level, the exponential distribution of the instantaneous power and its covariance function.

Associated with a wide sense stationary process, whose autocorrelation function $R(\tau)$ only depends time displacement $\tau = t-t'$, is its power spectral density

$$S'(f) = \sqrt{\pi}\tau_o\exp[-(\pi\tau_o f)^2] \qquad 2\text{-}3$$

obtained as the Fourier transform of (2-2), that physically represents the power spectral density of a received CW signal undergoing signal fading. If we define the signal bandwidth $B = 1/\sqrt{\pi}\tau_o$, then only $\mathrm{erf}(\pi\tau_o B/2) = 0.79$ or about 80 percent of the power will be received within this bandwidth for a transmitted CW tone. This tone spectral spreading places a lower bound on the minimum detection bandwidth and the maximum modulation symbol period for which the received signal can be efficiently received.

The effects of angular scattering defined by σ_x^2 and σ_y^2, or alternatively by the corresponding equivalent "antenna gains" $G_{\sigma x} = 2/\sigma_x^2$ and $G_{\sigma y} = 2/\sigma_y^2$, adjusted by the influence of terminal antenna gain G_T, will affect the decorrelation, the channel coherent bandwidth and the antenna scattering loss. Specifically, we have the result for the signal decorrelation time

$$\tau_o = 1/[(v_x/\ell_{ox})^2+(v_y/\ell_{oy})^2]^{1/2}, \qquad 2\text{-}4$$

where $\ell_{ox} = (\lambda/2\pi)\sqrt{G_x}$, $\ell_{oy} = (\lambda/2\pi)\sqrt{G_y}$ with $G_x = G_T+G_{\sigma x}$, $G_y = G_T+ G_{\sigma y}$, and where λ is the free space wavelength of the transmitted signal and v_x and v_y are the velocity components associated with the scattering medium motion relative to the terminal. Under strong nuclear scintillation conditions the terminal antenna gain G_T is typically much larger than $G_{\sigma x}$ and $G_{\sigma y}$ and then ℓ_{ox} and ℓ_{oy} approach their common minimum value $(\lambda/2\pi)\sqrt{G_T}$ that leads to the minimum decorrelation time $\tau_o = (\lambda/2\pi v)\sqrt{G_T} = D/2v$, where $v = \sqrt{(v_x^2+v_y^2)}$ is the relative medium velocity. The last equality holds for a circular dish antenna having the gain $G_T = (\pi D/\lambda)^2$, where D is the antenna diameter. Thus, the minimum decorrelation time of the signal is governed by the antenna size and not by the antenna gain and angular scattering of the medium. In Table 2-1 the minimum signal decorrelation time is given for the various antenna gains.

Table 2-1 Minimum Signal Decorrelation Time in a Strong Nuclear Induced Scintillation Environment.

Terminal Antenna Gain	(Size at 7/8 GHz)	Relative Velocity, v 100 m/s	300 m/s	1000 m/s
G_T = 62 dB	(60')	τ_o = 91.4 ms	= 30.5 ms	= 9.1 ms
= 58	(40')	= 61.0	= 20.3	= 6.1
= 52	(20')	= 30.5	= 10.2	= 3.1
= 44	(8')	= 12.2	= 4.1	= 1.2
= 35	(33")	= 4.2	= 1.4	= 0.4

The signal decorrelation time is directly related to the fade rate of the signal defined as the average number of times per second the signal power is less than a certain threshold level S_T. Specifically, the fade rate and the average fade duration [2]

$$R_f = (1/\tau_o)\sqrt{(2S_T/\pi S)}\exp(-S_T/S), \quad \text{Hz} \qquad 2\text{-}5$$
$$T_{fade} = \tau_o\sqrt{(\pi S/2S_T)}[\exp(S_T/S)-1], \quad \text{seconds.}$$

If we consider the threshold $S_T/S = 1/4$ -- that is, S_T is 6 dB below the average power level S -- the fade rate becomes $0.311/\tau_o$ and the average fade duration is $0.712\tau_o$. The maximum fade rate is $1/\tau_o\sqrt{(\pi e)} = 0.342/\tau o$, which occurs for $S_T/S = 1/2$, or for a 3 dB margin threshold. We see that the fade rate is about 1/3 of $1/\tau_o$ and not equal to $1/\tau_o$ as commonly thought. Between the deep fades the signal is above this threshold and we can determine the average "strong" signal period

$$T_{signal} = 1/R_f - T_{fade} \qquad 2\text{-}6$$
$$= \tau_o\sqrt{(\pi S/2S_T)}, \quad \text{seconds,}$$

associated with the fading signal process. In Table 2-2 we have determined the fade rate and average duration for various signal decorrelation times.

In the technical development above we tacitly assumed single link fading, i.e., uplink or downlink fading. However, if we do experience simultaneous uplink and downlink fading and we use a transponder satellite, the combined signal variation will not constitute Rayleigh fading. It can be shown that the received signal has a uniform phase ditribution and a "Bessel" power probability density [2]

$$p(S(t)|S) = (2/S)K_o[2\sqrt{(S(t)/S)}], \qquad 2\text{-}7$$

where $K_o(\ )$ is the modified Bessel function of second kind and $S(t) = S[x_1^2(t)+y_1^2(t)][x_2^2(t)+y_2^2(t)]$ with subscripts 1 and 2 refering to the uplink and downlink, respectively. This distribution leads to Laplace distributed inphase and quadrature components

$$p[x(t)] = \exp[-2|x(t)|] \qquad 2\text{-}8$$
$$p[y(t)] = \exp[-2|y(t)|],$$

where $x(t)+iy(t) = [x_1(t)+iy_1(t)][x_2(t)+iy_2(t)]$ is the normalized received signal. [It should be noted that since the two quadrature components x(t) and y(t) are not Gaussian distributed, they are not uncorrelated and statistically independent even if the signal phase $\phi = \arg[x(t)+iy(t)]$ is still

Table 2-2 Signal Fade Rate, Average Fade Duration and Signal Period; Rayleigh Fading, Fade Threshold $S_T/S = -6$ dB.

Signal Decorrelation Time τ_o	Fade Rate R_f	Average Fade Duration T_{fade}	Signal Period T_{signal}
1 ms	311.0 Hz	0.7 ms	2.5 ms
10	31.1	7.1	25.1
100	3.1	71.2	251
1,000	0.3	712.0	2,510

uniformly distributed $[0,2\pi)$.] The signal correlation function

$$R(\tau) = R_1(\tau)R_2(\tau) = \exp[-(\tau/\tau_o)^2], \quad \tau = t-t', \qquad 2\text{-}9$$

where $(1/\tau_o)^2 = (1/\tau_{o1})^2+(1/\tau_{o2})^2$ relates the resulting signal decorrelation time τ_o with those of the uplink (τ_{o1}) and downlink (τ_{o2}), respectively.

3. RECEIVED SIGNAL COHERENT BANDWIDTH

Closely associated with the decorrelation time of the signal is the coherent bandwidth (f_o) of the transmission channel. This performance parameter determines the maximum (approx.) bandwidth that allows for non distorted digital signal transmissions and is directly tied to the multi-path delay spread in the channel. The effective coherent bandwidth is also governed by σ_x^2 and σ_y^2 and the terminal antenna gain G_T, or equivalently $G_x = G_T+G_{\sigma x}$ and $G_y = G_T+ G_{\sigma y}$. From geomertic considerations [2], we have the result

$$f_o = (c/4\pi h)[(1/G_x^2+1/G_y^2)/2]^{1/2}, \quad \text{Hz}, \qquad 3\text{-}1$$

where c denotes the speed of propagation and h is the effective height (distance) to the scattering volume from the terminal. For strong scintillation conditions both G_x and G_y approach G_T and $f_o = (c/4\pi h)G_T$. Thus the minimum coherent bandwidth is determined by the terminal antenna gain and the effective scattering volume height. This observation is most important from a systems engineering point of view; a satellite network with a certain minimum size of earth terminals can be ensured a significantly larger coherent bandwidth than that which the scintillation medium itself defines to a very small (omni directional) antenna. [That is, if we let G_T approach zero above]. In Table 3-1 the minimum coherent bandwidth has been determined for various size terminal antennas.

The important observation to be made is that the minimum signal decorrelation time and coherent bandwidth are lower bounded by the terminal antenna gain, the terminal-satelite geometry and the velocity of the scattering medium. The results are not dependent on the angular sattering variances which are strongly dependent on the nuclear senario. This implies that if the minimum decorrelation and coherent bandwidth results are used as design criteria for the modulation subsystem developments the resulting design will meet operational requirements for all nuclear senarios.

Again we have tacitly only considereed a single link. If we consider both uplink and downlink fading, the resultant channel coherent bandwidth f_o

Table 3-1 Minimum Signal Coherent Bandwidth in a Strong Nuclear Induced Scintillation Environment.

Terminal Antenna Gain	(Size at 7/8 GHz)	Effective Height, h 300 km	1000 km	3000 km
G_T = 62 dB	(60')	f_o = 126 MHz	= 38 MHz	= 13 MHz
= 58	(40')	= 50	= 15	= 5.0
= 52	(20')	= 13	= 3.8	= 1.3
= 44	(8')	= 2.0	= 600 kHz	= 200 kHz
= 35	(33")	= 250 kHz	= 75	= 25

is given by

$$(1/f_o)^2 = (1/f_{o1})^2+(1/f_{o2})^2 \qquad 3\text{-}2$$

in terms of the uplink (f_{o1}) and downlink (f_{o2}) coherent bandwidths, respectively [2].

4. ANTENNA SCATTERING LOSS

Even if the antenna scattering loss does not affect the modulation subsystem design we will briefly mention its effect. The average receive power level is also affected by the angular scattering of the medium and the terminal gain as the limited beam width of an antenna will reject signal component with large scattering angles. Specifically, the net received (average) signal power level

$$S = S_o/[(1+G_T/G_{\sigma x})(1+G_T/G_{\sigma y})]^{1/2}, \qquad 4\text{-}1$$

where S_o is the received power level before the nuclear event and as before $G_{\sigma x} = 2/\sigma_x^2$ and $G_{\sigma y} = 2/\sigma_y^2$ are the equivalent terminal gains associated with the angular scattering process [2]. Thus, the terminal antenna scattering loss

$$\begin{aligned} L_s &= 10\log(S_o/S) \quad \text{dB} \\ &= 5\log(1+G_T/G_{\sigma x})+5\log(1+G_T/G_{\sigma y}) \end{aligned} \qquad 4\text{-}2$$

expressed in decibels. It is important to recognize that since typically $\sigma_y^2 = \sigma_x^2/K^2(\Phi) << \sigma_x^2$ or equivalently $G_{\sigma y} = G_{\sigma x}K^2(\Phi) >> G_{\sigma x}$, where $K^2(\Phi) = 1+224\sin^2(\Phi)$ with Φ being the signal penetration angle through the scattering volume, the antenna scattering loss will be dominated by the first term of (4-2). In Table 4-1 we have illustrated the antenna scattering loss as it depends on the various terminal sizes.

5. UPLINK AND DOWNLINK FADING

With respect to simultaneous uplink and downlink fading using a transponder satellite the resulting signal fading characteristics may be viewed as "the product" of two Rayleigh fading variables resulting in the Bessel type received power or signal amplitude probability distributions. This model assumes that the downlink is received by a terminal that it thermal

Table 4-1 Terminal Antenna Scattering Loss
Illustrative Example.

Penetration Angle: Φ = 27.9 Degrees, [$K^2(\Phi)$ = 15 dB.]
Equivalent Antenna Gains: $G_{\sigma x}$ = 40 dB, $G_{\sigma y}$ = 55 dB.

Terminal Antenna Gain	(Size at 7/8 GHz)	Terminal Antenna Scattering Loss
G_T = 62 dB	(60')	L_s = 11.0+3.9 = 14.9 dB
= 58	(40')	= 9.0+2.4 = 11.4
= 52	(20')	= 6.1+0.9 = 7.0
= 44	(8')	= 2.7+0.2 = 2.9
= 35	(33")	= 0.6+0.0 = 0.6

front-end noise limited. That is by a small, low gain, earth terminal. Now if the signal is received by a large earth terminal for which the thermal front-end "noise floor" is insignificant relative to the noise level established by all other signals, the multiple access or interference signals, received from the satellite this noise level will also experience the same downlink fading as the desired signal. This implies that the effective received signal-to-noise ratio will not vary due to downlink fading and generally the automatic gain control (AGC) in the receiver will maintain an almost constant detection signal level with respect to downlink fading. However, downlink fading will result in phase modulation which can not be removed by the AGC and therefore downlink fading will affect the receiver demodulation performance.

Even if the receiver performance degradation from phase variations alone does not lend itself to analytical solutions we may be justified by using the small terminal model to bound the problem while we recognize that we need not account for antenna scattering loss for the downlink as long as the receive terminal noise level is set by the "other" received signals from the satellite. We know that the average fade rate is about $1/3\tau_o$ and thus the "deep" fades are separated in average by $\Delta t = 3\tau_o$ for which the normalized signal autocorrelation function $R(\Delta t) = \exp[-(\Delta t/\tau_o)^2] \cong 1\times10^{-4}$. In other words, the signal is virtually uncorrelated within a fraction of the average fade separation in time. Thus, we may attribute the signal decorrelation to the random phase process associated with the fading process and base the modulation subsystem designs not only on Rayleigh signal fading but on the Bessel type fading as well.

6. SYSTEMS ENGINEERING IMPLICATIONS

More importantly than the actual expressions for the signal decorrelation time (τ_o), coherent bandwidth (f_o) and the terminal antenna scattering loss (L_s) is the fact that these characteristics are all governed by the angular scattering distribution and in particular the angular scattering variances σ_x^2 and σ_y^2, or equivalently the equivalent antenna gains $G_{\sigma x}$ and $G_{\sigma y}$. This implies that as the angular scattering distribution changes over time the three characteristics also change together in time. Specifically, the minimum correlation time (fast signal fading), the minimum coherent bandwidth and the maximum antenna scattering loss are tied together and when we experience long signal decorrelation times (slow signal fading) the coherent bandwidth

is generally extremely large and the antenna scattering loss is negligible.

The question often raised is: how is it possible to mitigate against nuclear induced signal scintillations that will cause Rayleigh fading and associated signal decorrelation time and coherent bandwidth limitations? First, it is necessary not to employ a larger instantaneous transmission bandwidth than the minimum channel coherent bandwidth for an ECCM modulation subsystem whenever the digital transmission rates are comparable and lower than the maximum received signal fade rate. However, the total transmission bandwidth, or spread spectrum modulation bandwidth, may be made much larger than the coherent bandwidth by using frequency hopping since the bandwidth associated with each hop is the instantaneous bandwidth of the signal and much smaller. At the receiver it is necessary to restrict the demodulation coherence (integration) time to a value comparable or less than the minimum signal decorrelation time. This constraint still allows for non-coherent combining of many signal elements (chips) corresponding to a bit or transmission symbol.

7. REFERENCES

1. Wittwer, Leon A Radio Wave Propagation in Structured Ionization for Satellite Applications II, Report DNA-IR-82-02, 1 August 1982, Defense Nuclear Agency, Washington D.C. 20305.

2. Kulistam, Per A. The Rationale for the SHF AJ/AS Government Waveform Standard (GWS) PAIRCOM Report for the Defense Communications Agency, Code R420, Contract DCA100-88-C-0012; 22 December 1988. Washington D.C. 22090.

AD-P006 282

A REVIEW OF THE HISTORY OF FIELD STRENGTH MODELS IN PROPHET

David B. Sailors
Ocean and Atmospheric Sciences Division
Naval Ocean System Center
San Diego, CA 92152-5000

ABSTRACT

Changes in the PROPHET field strength model are reviewed. Until recently, the model was based on that developed by the Deutsche Bundespost (Damboldt, 1975). In this model the field strength depends on path distance, the transmitted power, a lower frequency limit fl, and a upper frequency limit fo. The different PROPHET versions differ on how fl and fo were formulated. PROPHET implementations of fl and fo are reviewed. The most accurate PROPHET model using the Deutsche Bundespost model is a version called HFTDA. Recently, a new model was developed for paths less than 7000 km which was a significant departure from prior versions and calculates modes propagating via the E-layer as well as by the F-layer. The ionospheric absorption and over-the-MUF loss are estimated as in IONCAP. As this model made a significant improvement in the accuracy of the estimates of field strength, it was incorporated into HFTDA.

INTRODUCTION

The effective operation of long distance high frequency (HF) communications has increased in proportion to the ability to predict variations in the ionosphere. These variations are affected in a complex manner by solar activity, seasonal and diurnal changes, as well as latitude and longitude. Such a predictive capability has permitted communications to optimize frequencies, antennas and other circuit parameters.

Initially, manual methods were developed for analyzing ionospheric variations on HF circuits of short, intermediate, and long distances. Because the manual methods were laborious and time consuming, various organizations developed computer programs to analyze HF circuit performance.

The advent of micro-computers allowed the development of real time electromagnetic/electrooptic propagation assessment systems (Richter, 1989). For HF radio propagation, a family of assessment systems was developed which is known under the name of PROPHET, acronym for Propagation Forecasting Terminal (Richter et al., 1977; Argo and Rothmuller, 1979; Rose, 1981; Rose, 1989). The systems consider solar effects on the ionosphere (both diurnal and seasonal as well as effects caused by solar disturbances) and provide path-specific system performance data. PROPHET was primarily developed using simple, empirical models and has proliferated into many different versions.

In addition to estimating the path MUF and LUF, one of the main parameters estimated in PROPHET is the skywave field strength. There have been many changes in the field strength model since it was first developed by Argo (private communication, 1981). The model was based on that developed by the Deutsche Bundespost (Damboldt, 1975). Recently, a new field strength model was developed that was a significant departure from prior versions. In this model the Deutsche Bundespost model was replaced for paths shorter than 7000 km by a model relying on propagation by both the E-layer and F-layer (Gnession et al., 1988). The ionospheric absorption equation is the one used in IONCAP for D-region absorption. The over-the-MUF losses are calculated using the Phillips (Wheeler, 1966) method for values of frequency/MUF ratio up to about 1.5 depending on ground distance and with values greater being considered in the "scatter" region. The value of the absorption index at night is calculated according to Wakai et al. (1971).

After describing how the field strength is calculated by both the Deutsche Bundespost and newer methods, different implementations of these methods in PROPHET will be reviewed. The accuracies of the latest two versions are presented. Samples are given of PROPHET field strength displays for different PROPHET versions. Possible improvements in the field strength model are suggested.

DEUTSCHE BUNDESPOST FIELD STRENGTH MODEL

The Deutsche Bundespost method of field strength calculation was developed by Beckman (1967) and has been continually refined by the work of Damboldt (1975) and Damboldt and Suessmann (1989). It was derived from an extensive field strength measurement program. The measurements revealed a steady increase of field strength from the lowest usable frequency (fl) for the path to higher frequencies, following approximately an inverse frequency dependence. Beckman defines fl as that frequency for which the field strength is some low limiting value. After a maximum value is reached, the field strength decreases steadily until it reaches the operational MUF. On the diminishing side of the curve, after the maximum, is the basic MUF, up to which frequency the usual propagation by ionospheric refraction takes place. Above the basic MUF, propagation is possible as a consequence of several mechanisms (e.g., ionospheric scatter by irregularities in the F-layer) which are not taken into account by prediction techniques based on theoretical considerations of effective propagation modes. The operational MUF (fo) is obtained as the product of a numerical factor K, determined from geographical and diurnal characteristics of the path, times the MUF.

91-09675

Given the above behavior of the field strength with frequency, Beckman (1967) developed an empirical formula that fit the measured field strength F, which is valid for non-deviative absorption:

$$F = Fo\{1 - fo^2[(fl/f)^2 + (f/fo)^2]/[fo^2 + ft^2]\} \tag{1}$$

for 1000 kW effective radiated power. F is the received field strength in dB relative to 1 μV/m; f is the operating frequency. Here, Fo is the free-space field strength given by

$$Fo = 20log_{10}\frac{3 \times 10^5 (Perp)^{1/2}}{d} \tag{2}$$

where Perp is the effective radiated power in kilowatts and d is the slant path range in kilometers for mirror reflection from a height of 300 km. For 1000 kW radiated Fo is

$$Fo = 139.6 - 20log_{10}d \tag{3}$$

Equation (1) can be adjusted for arbitrary power levels and antenna gains by including the actual power level used and by the addition of transmitter antenna gain. Since Beckman originally used a reference field due to an isotropic antenna over perfect ground, which has a gain of 4.8 dB with respect to an isotropic antenna in free space, 4.8 dB is subtracted to go from the effective radiated power to radiated power. Equation (1) then becomes

$$F = Fo\{1 - \frac{fo^2[(fl/f)^2 + (f/fo)^2]}{fo^2 + ft^2}\} - 4.8 + Gt + 10log\ [\frac{Pt}{1000}] \tag{4}$$

where Pt is the transmitted power in kilowatts and Gt is the transmitter gain in dB with respect to an isotropic antenna in free space. For a more exact computation of the field strength, the influence of the Earth's magnetic field has to be allowed for. In practice, it has proved to be sufficient to add the value of the gyro-frequency (about 1 MHz) to the frequencies f, fl, and fo.

At low frequencies near fl, the D-region absorption is greatest and is the dominant factor in determining operational frequencies. The field strength can be approximated by F=Fo-Li where Li is the ionospheric absorption. Ionospheric absorption has the approximate form for oblique propagation

$$Li = \frac{Ai \sec io}{(f + fh)^2} \tag{5}$$

where fh is the gyro-frequency and io is the D-region incidence angle. If fl is defined as the frequency that produces 0 dB (μV/m) for 1000 kW effective radiated power, one can obtain fl by substituting equations (2) and (5) into the expression for the field strength:

$$fl = (\frac{Ai \sec io}{20log_{10}[9.5 \times 10^4/d]})^{1/2} - fh \tag{6}$$

For the absorption expression currently utilized by Damboldt and Suessmann (1989), fl is given by

$$fl = 5.3[\frac{\sum_{2N} cos^{1/2}\chi(1 + 0.009R12)}{cos\ io\ log_{10}((9.5 \times 10^6)/d)}]^{1/2} - fh \tag{7}$$

where χ is the sun's zenith angle, R12 is the twelve month running mean of the sunspot number, and N is the number of hops.

The current expression for the operational MUF used by the Deutsche Bundespost (Damboldt and Suessmann, 1989) is

$$fc = K * Cr * fg \tag{8}$$

where, for E-layer modes, K=1. For F2-layer modes

$$K = 1.2 + W\frac{fg}{fg,noon} + X((\frac{fg,noon}{fg})^{1/3} - 1) + Y(\frac{fg,min}{fg,noon})^2 \tag{9}$$

In equations (8) and (9); fg; fg,noon; and fg,min are, respectively, the MUF for the hour of interest and the maximum and minimum MUF for the 24-hour day. W, X, and Y are empirical constants, which are functions of raypath azimuth at the midpath point derived by linear interpolation with angle between W=0.1, X=1.2, Y=0.6 for east-west paths and W=0.2, X=0.2, and Y=0.4 for north-south paths. For paths of length less than or equal to 4000 km, Cr = 2-(D/4000) where D is the path length in km. For paths greater than 4000 km, Cr = 1.

Of particular interest is the values of field strength given by equation (4) at a frequency for which the field strength is a maximum for specified path and time. In particular, F=-34.8 dB (μV/m) for f=fl and f=fo, for Pt= 1 kW radiated and for Gt=0. F attains a maximum at the geometric mean of the limits of the transmission frequency range given by

$$fmax - fh = ((fl + fh)(fo + fh))^{1/2} \tag{10}$$

The maximum field strength is given as (Bradley and Liu, 1982)

$$Fmax = (139.6 - 20\log d)\left(1 - \frac{2b}{1+b^2}\right) - 34.8 \tag{11}$$

where

$$b = \frac{fl + fh}{fo + fh} \tag{12}$$

Figure 1 shows the dependence of Fmax on ground range for various values of b. Note that the curve for b=0 corresponds to a free-space distance dependence. The field is reduced by some 10 dB from the free-space value for b=0.1 and by 20 dB for b=0.2. For an operational MUF given by fo=K*fb where fb is the basic MUF, equations (10) and (12) give (Bradley and Liu, 1982)

$$fmax/fb = K[\sqrt{b} - (1 - \sqrt{b})(fh/fo)] = K*b \tag{13}$$

Figure 2 shows fmax/fb versus K for selected values of b. With typical values of b in the range 0.1-0.5 and K between 1.2 and 2.4, it is seen that the frequency with maximum field strength lies in the range 0.4-1.7 times the basic MUF.

MODIFICATIONS TO THE DEUTSCHE BUNDESPOST MODEL

The Deutsche Bundespost model was first implemented in PROPHET in 1980 (Argo, private communication, 1981). This version adhered to Damboldt's expression for operational MUF but used the PROPHET LUF algorithm for estimates of fl.

By 1981, the versions of PROPHET running on the Tektronix 4050 series of computers, known as the Tactical Prediction Modules (e.g., The Army PROPHET Evaluation System (APES)), were using the MUF as the operational MUF fo. Also, the version of the field strength model for use on the Apple computer reported by Li et al. (1983) used the MUF for fo.

One should note that the shape of the field strength curve does not change with changing radiated power; rather, the received field strength simply increases or decreases. The operational MUF will increase and decrease with radiated power. The parameter fl will decrease with increasing power, and vice versa.

This nonchanging of the field strength curve shape is important. If the parameters fo or fl are held constant (the basic MUF used for fo) and the transmitter power varied, then the field strength versus frequency curve will change. This is contrary to the original assumptions in its derivation. Hence, fl and fo must be specified at a given power.

Figure 3 illustrates these effects and shows field strength versus frequency determined using equation (4). It is plotted for two different power levels (1 kW and 10 kW) and for use of two different MUF values (the MUF itself and the operational MUF). Also included is a curve for 10 kW APES Version 3.0. The curves "10000 W AP OM" and "1000 W AP OM" illustrate the desired effects of changing power, namely the shape of the curve does not change with changing radiated power. The curves stop at the LUF and operational MUF labeled OM. Note that OM increased with increase of power. In the other two cases, "1000 W AP" and "10000 W AP", MUF is used instead of OM for fo. This results in three effects. The first is that the shape of the field-strength curve is distorted as the power is increased from 1 kW to 10 kW. The second effect is that the peak field strength is predicted low. The third is that the peak value occurs at a lower frequency than it should. The fifth curve is for "10000 W APES 3.0". Apes Version 3.0 used a formula for equation (3) which included the power and the power term in equation (4) was not used. At 10 kW this curve and "10000 W AP OM" are identical at the low frequency end. At 1 kW radiated "1000 W AP OM" and "1000 W APES 3.0" are identical. The use of this of this formula results in extreme distortion in the field strength at the MUF when the radiated power is varied.

It was because of the problems in the field strength algorithms that an effort was begun in 1984 to modify the Deutsche Bundespost model to eliminate these sources of error. In 1985 the modified field strength model was incorporated into APES. In 1986 it was incorporated into a PROPHET version called HFTDA (High Frequency Tactical Decision Aids). The improved model and its accuracy have been reported by Sprague (1987).

Two important modifications were made to the field strength models previously used. First, the use of a LUF model was replaced by an expression for fl based on equation (6). The absorption index used was based on Schulz and Gallet (1970). For paths less than 4000 km, all calculations are done at a single control point, the path midpoint. At greater ranges control points are established at the midpoint and at points 1000 km from each end point. Ai is then determined at each control point, and the average of these values determines the absorption index for the path. A simplified method of estimating the operational MUF was implemented. It was assumed that the MUF value at a given time, season, geographic position and and sunspot number is a statistical parameter and has Gaussian distribution around a mean. Given this assumption the operational MUF in PROPHET was determined by the product of the MUF and the 99.1 percentile value of the MUF distribution

$$fo = 1.85*(HPF)*MUF \tag{14}$$

where HPF is the 90th-percentile value factor for an assumed Gaussian MUF distribution (Barghausen et al., 1969). The reason this was used rather than equation (9) was that equation (9) was based mainly on data measured on paths terminating in one site and might contain aspects specific to that site.

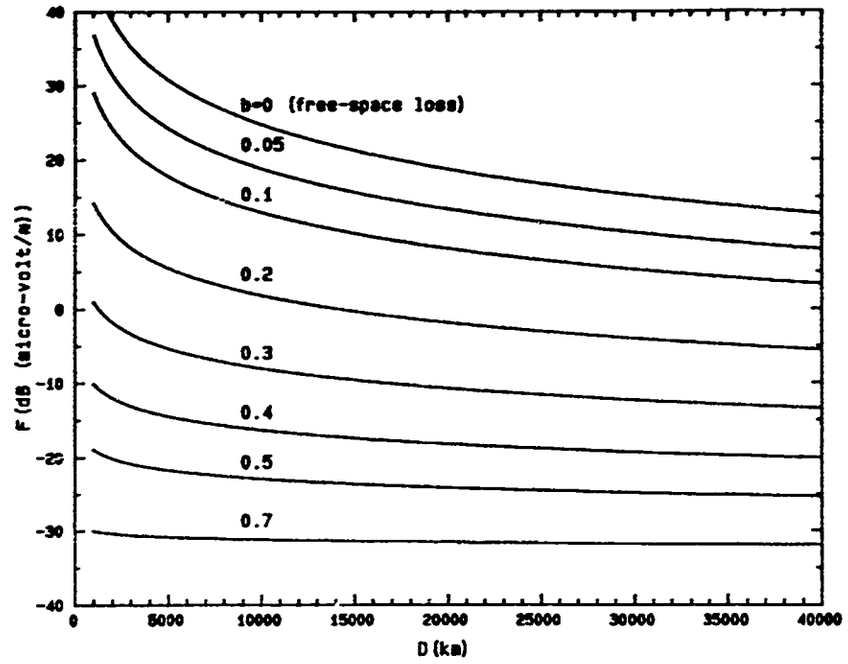


Figure 1. Peak field strength versus ground range for various values of the parameter b.

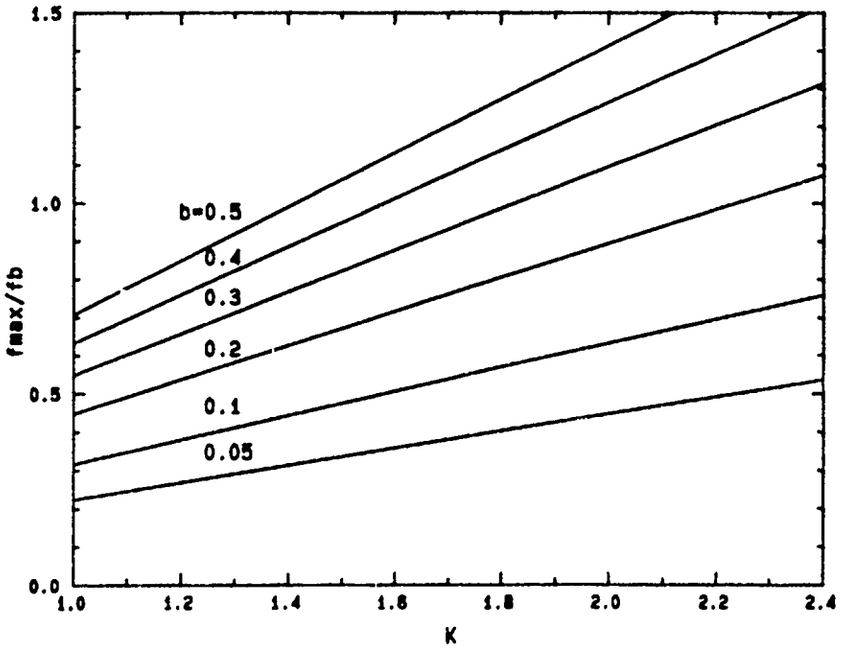


Figure 2. Variation of the ratio of the frequency with maximum field strength to the basic MUF versus K for various values of the parameter b.

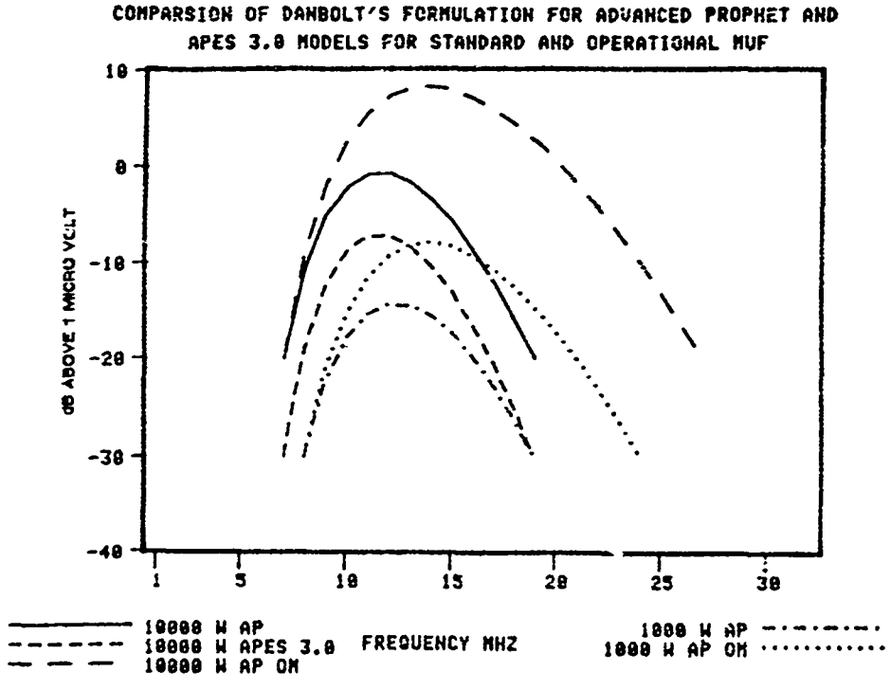


Figure 3. Field strength curves from the Deutsche Bundespost model as a function of power and frequency and implementation method.

PROPHET VERSION 3.2

In PROPHET Versions 3.1 and later, fl is given by the LUF; field strength is calculated for 1 kW transmitted only; and above-the-MUF loss term is increased by 12 dB per MHz until a maximum increase in loss of 60 dB is obtained. In PROPHET Versions 3.2 and earlier, the LUF model is QLOF Version 1.0 (Argo and Sailors, 1979; Sailors and Muision, 1987). The field strength model in CHIMPS (Chirpsounder Integrated MUF Prediction System) is the same as that used in PROPHET Version 3.2.

NEW FIELD STRENGTH MODEL

Recently, a new field strength model was developed which was a significant departure from prior versions. In this model the Deutsche Bundespost model was replaced for paths shorter than 7000 km by a model relying on propagation via the regular E-layer and the F2-layer (Gnession et al., 1988). E-layer, F2-layer, and mixed-mode hops are calculated. The MUF model used is MINIMUF-85 (Sailors et al., 1986) for the F2-layer. The MUF for the E-layer is calculated as in ESSA-ITSA-1 (Lucas and Hayden, 1966) The ionospheric absorption equation is the one used in IONCAP (Teters et al., 1983). The absorption index at night has a value which is a function of solar activity, which was taken from Wakai (1971). The over-the MUF losses are calculated using the Phillips method (Wheeler, 1966) for values of frequency/MUF ratio up to about 1.5 MHz The losses at ratios above about 1 5 MHz are estimated as in IONCAP for the "scatter" region. Above about 7000 km, the field strength method of HFTDA is retained.

The minimum hop predicted for the E-layer is calculated for an h'E of 110 km and the path length. The minimum radiation angle associated with this mode is assumed to be at the horizon. If the frequency is greater than 1.4 * EMUF, no E-mode reflection is computed.

For F2-layer hops the minimum radiation angle is assumed to be at the horizon and the mode must also penetrate the E-layer at the F2-layer angle calculated using h'F2 as calculated in MINIMUF-85 for ray-trace displays in HFTDA.

The mixed mode is calculated using the E-layer height (110 km) and the h'F2 to determine a take-off angle. Only one E-layer hop is considered in the mixed mode. The frequency must be supported by the E-layer at one end of the path and penetrate at the other end for this mode to be possible.

Only one minimum hop is chosen for each mode of propagation. The value of the transmission loss used to predict the fields is the loss associated with the mode of least loss.

ACCURACY OF METHODS

Both the field strength models in HFTDA have been compared with CCIR Data Base C. There are 81 paths in the data base. In the results that follow LtFld refers to the new field strength model and DmBlt refers to the modified Deutsche Bundespost model.

Table 1 shows the overall results for all 81 paths at all frequencies recorded in the database. The results are satisfying and show the new field strength method to have much lower average residuals than the old method. The RMS residual is also better by at least a factor of two to one. The correlation coefficient is better but the improvement is not pronounced. This could be caused by a time lag in the prediction relative to the observation data; however this has not been verified. The most pronounced improvement seems to be in the over-the-MUF calculations which show a much lower average residual of approximately one-fifth and a RMS residual of less than one-half. Table 2 shows the accuracy as function of the ratio of frequency to MUF. It can be seen that the shape of the distribution of losses above the MUF is not that of a normal curve, but possibly an exponential curve as suggested by Damboldt and Suessmann. The scatter fields (well above 1.4 time the MUF) are stronger than predicted.

Table 1. Summary of overall results for entire CCIR Data Base C (dB)

Condition f/MUF	Avg Residual		RMS Residual		Correlation		% of Total	
	LtFld	DmBlt	LtFld	DmBlt	LtFld	DmBlt	LtFld	DmBlt
over all	0.0	4.7	14.2	23.1	0.70	0.60	100.0	100.0
<= 1.00	-1.1	1.5	13.5	16.5	0.69	0.65	70.3	71.1
> 1.00	2.7	12.6	15.8	34.5	0.60	0.48	29.7	28.9

Inspection of tables of comparison results for path length show an improvement for very short paths (less than 1000 km) with an average residual for the new method of about one-fifth the older value. The RMS residuals also show good improvement. However, the correlation coefficient looks very bad for paths between 2000 and 3000 km, but shows good improvement for other ranges. It is suspected that paths of this length should be easily predicted. However, examination of the results shows that only 1.2% of the data is in this path length range. As this distance is the crossover region for the one-hop E-layer mode and the one-hop F2-layer or 2-hop E-layer mode, changing day-to-day ionospheric heights and critical frequencies cause the least mode to change. A similar trend is shown for the 4000-5000 km range which is the crossover range for the one-hop and two-hop F2 layer modes. Antenna patterns are sensitive to these changing modes due to vertical takeoff angles associated with the modes' availability day-to-day. Only 0.9% of the data was in the 4000-5000 km range.

Table 2. Summary of accuracy results as a function of frequency to MUF ratio for the entire data base.

f/MUF Condition	Avg Residual LtFld	Avg Residual DmBlt	Rms Residual LtFld	Rms Residual DmBlt	Avg Rel Res LtFld	Avg Rel Res DmBlt	Rms Rel Res LtFld	Rms Rel Res DmBlt	Avg Abs Rel Res LtFld	Avg Abs Rel Res DmBlt	Correlation LtFld	Correlation DBlt	% of Total LtFld	% of Total DBlt
.40	31.7	55..	44.3	63.6	-1.6	8.5	2.2	71.2	-1.6	8.6	90	88	.1	3
.60	6.1	15.5	21.4	31.6	-2.4	1.6	17.2	65.2	1.6	10.6	.64	67	2.3	2.5
.80	3	6	11.5	17.7	0	.3	16.2	19.9	1.7	1.9	.76	.76	9.6	9.4
.90	2.1	3.6	11.1	13.5	0	.3	17.2	22.5	2.6	2.9	.74	.75	12.3	12.4
1.00	1.9	2.9	11.8	13.1	1.2	1.7	29.3	32.9	2.3	3.3	.71	.73	11.7	11.9
1.10	-.1	1.0	12.2	13.5	-.9	-.3	14.7	15.7	1.5	2.0	.72	.69	10.5	10.6
1.20	-3.5	-2.5	15.1	15.6	-1.1	-1.3	18.0	21.9	2.0	2.8	.67	.60	9.8	9.9
1.30	-6.9	-4.7	15.6	16.4	-2.1	-1.6	18.0	17.5	2.2	2.8	.54	.66	7.7	7.7
1.40	-8.3	-5.5	16.7	17.4	-1.2	-.1	15.8	12.7	1.5	1.1	.73	.66	6.4	6.4
1.50	-6..	-5.0	16.8	17.4	-1.8	-1.8	20.3	19.9	2.0	2.4	.70	.69	5.0	5.0
1.60	-4.7	-4.2	16.0	16.4	-.4	1	13.3	9.8	.9	.4	.67	.66	4.1	3.9
1.80	.0	-4.5	17.6	17.9	.2	-.8	12.3	14.6	1.4	1.1	.51	.50	2.9	2.8
2.00	4.2	-5.1	17.3	16.6	.9	.4	2.4	5.8	.5	.3	.37	.50	2.0	1.9
2.50	1.0	-4.7	15.3	16.8	1.2	.1	4.8	7.4	.5	.5	.48	.35	2.3	2.2
3.00	5.6	-.1	17.0	14.0	1.4	1.5	3.6	6.9	1.0	.7	.35	.43	1.6	1.5
3.50	6.9	3.5	14.3	12.5	2.9	2.1	14.2	10.9	.7	1.2	.51	.48	1.3	1.2
4.00	5.0	4.9	14.9	13.5	4.4	3..	17.4	14.5	4.0	3.6	.31	.16	1.1	1.0
4.50	4.9	9.0	12.6	15.0	3.2	6.	18.6	28.4	3.	6.4	.40	.09	.9	1.0
5.00	3.8	14.3	13.7	25.9	5.6	10.7	18.4	48.1	3.4	10.7	.50	.06	1.1	1.0
5.50	8.5	31.6	13.4	44.8	6.4	14.9	30.9	81.8	6.4	14.0	.36	.18	3.2	3.1
6.00	11.4	66.7	14.6	88.5	4.5	14.9	29.9	120.7	4.5	14.9	.26	.12	1.8	1.7
6.50	14.1	83.1	16.1	91.5	8.4	10.1	50.2	131.3	8.4	10.1	.43	.09	1.0	1.0
7.00	15.5	82.3	17.4	82.6	3.5	-.7	11.3	14.5	3.5	-.7	.18	.00	.7	.7
7.50	17.5	90.3	20.2	85.5	-1.1	-5.0	6.9	5.6	-1.1	-5.0	.18	.00	.4	.4
8.00	19.0	75.9	21.1	76.2	-.8	-3.8	1.0	4.	-.8	3.8	.64	.00	.2	.2
9.00	21.4	71.8	22.	2..	-.7	-3.1	.0	3.3	-.7	-3.1	.18	.00	.2	.2

Over-the-MUF calculations shown as a function of path range shows large average residuals for these same ranges. This suggests an error in calculating the MUF for paths in these ranges. If over-the MUF losses were calculated when in fact the path observed was operating below the MUF, an error is introduced in both the estimates of field strength and the comparison itself. This has not been verified, and it is doubtful that data is available to verify this assumption.

Large errors occurring for cases when very few samples were recorded may be due to an insufficiently large sample size to produce meaningful statistics. Although included for completeness, it is bothersome. Another possibility is that data were recorded only on those days of the month that the signal was readable above the noise; hence, below the MUF for the path. If not all 30 days of a month at a given hour are not present in the data, it is difficult tell what part of the distribution the data represents.

SAMPLE DISPLAYS

Figures 4 through 9 illustrate some of the differences between more recent versions of PROPHET. Examples are given for the new HFTDA field method (LtFld), the old HFTDA field strength method (DmBlt), and PROPHET Version 3.2. Two types of displays are shown. Figures 4 through 6 show a diurnal display of signal-to-noise ratio for the path from Sicily to a mid-point on the radial to Beirut,Lebanon—a range of about 1000 km. This is beyond the skip zone and is not so far as to have excessive losses so that communications should be possible at some frequency. Figures 7 through 9 show a map display of 10.0 MHz signal-to-noise at 9:00Z transmitted from Sicily. In these figures the 10.7 cm flux was 234.2 and the date was 15 January 1990. The new version shows the strongest signal-to-noise at mid-day. PROPHET Version 3.2 signal-to-noise at mid-day is about 20 dB below that of the new model. On comparing figures 7 and 8, the effect of the E-layer is quite noticeable. In figure 7 the signal-to-noise are shown to be much stronger near the transmitter than in figure 8 for the modified Deutsche Bundepost method. In the case of PROPHET Version 3.2, the results are disappointing. Figure 9 shows little usable signals.

DISCUSSION

The importance of the correct choice of fl and fo was illustrated in the previous section where the PROPHET versions, Version 3.2 and HFTDA, estimate quite different values of field strength. Yet both versions have the same fo. Using figures 1 and 2 or equations (11) and (13), it is possible to understand the source of this difference. For the path and conditions for which figures 7 through 9 apply, the LUF, fl, FOT, MUF, and fo were printed. Assuming fh equal 1 MHz, b and b' where obtained for fl and LUF, respectively. K was found from K=fo/MUF. The peak field strengths, Fmax and Fmax', and the ratio of the frequency of maximum field strength to the basic MUF fb, fmax/fb and fmax'/fb, were determined for HFTDA and PROPHET Version 3.2, respectively. The results are given in table 3 for six different times of day. Table 3 shows that PROPHET Version 3.2 predicts the peak field strength from 7.4 to 21.9 dB below that of HFTDA for the case examined. At night the peak field strength estimated by PROPHET Version 3.2 is at the MUF or greater. Using figures 8 and 9 as additional information, at a frequency below the peak field strength (10.0 MHz, MUF=18.3, FOT=16.1) at 09:00, PROPHET Version 3.2 estimates field strengths even farther below that of HFTDA. However, at hour 05:00 HFTDA predicts a very low fl; this is probably due to an inadequate nighttime absorption index model in HFTDA.

Table 3. Differences between HFTDA and PROPHET Version 3.2

TIME	LUF (MHz)	fl (MHz)	fo	b	b' (MHz)	K	Fmax	Fmax'	fmax/fb (dB)	fmax'/fb (dB)
00:00	2.0	0.4	12.0	0.11	0.23	2.13	27.3	9.9	0.71	1.02
04:00	2.0	0.4	9.0	0.14	0.30	2.11	22.8	0.9	0.79	1.16
05:00	2.0	0.0014	22.9	0.04	0.13	2.10	38.3	24.3	0.42	0.76
09:00	5.6	3.6	37.9	0.12	0.17	2.07	25.8	18.4	0.72	0.85
11:00	6.6	3.9	37.6	0.13	0.18	2.03	24.3	16.9	0.73	0.86
17:00	2.0	0.4	17.5	0.07	0.16	2.04	33.5	19.8	0.54	0.82

The progress in the development of the field strength model for PROPHET is encouraging. Worthwhile items to be revised in the future and checked with observations are:

1) Use of Schulz-Gallet absorption equation
2) Damboidt over-the-MUF exponential distribution
3) Improved model for scatter fields above about 1.4 * MUF
4) Daytime low-frequency specular reflection algorithm.
5) Improved F2-layer height calculation
6) High-latitude description of losses
7) Improved MUF model
8) E-layer algorithm for MUF
9) Winter anomaly model for absorption
10) Improved absorption index for nighttime long paths

These are not necessarily in the order of priority. They represent potential areas of model enhancement which could lead to expected improvements in the predictions. These changes should be made to see which algorithm shows the best fit with the overall data and to make PROPHET a true predictor field strengths.

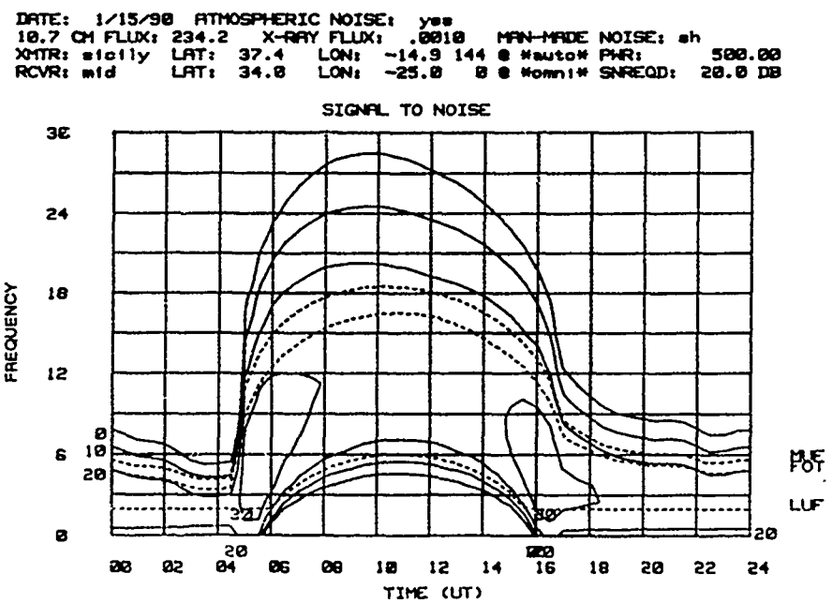


Figure 4. Signal-to-noise ratio for a path from Sicily to the mid-point of a radial to Beirut, Lebanon for the new field strength model in HFTDA.

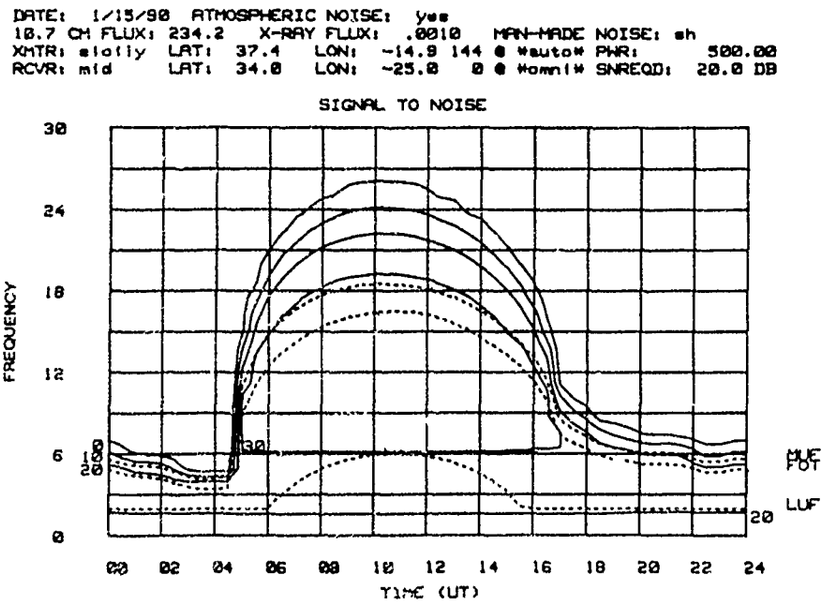


Figure 5. Signal-to-noise ratio for a path from Sicily to the mid-point of a radial to Beirut, Lebanon for the old field strength model in HFTDA.

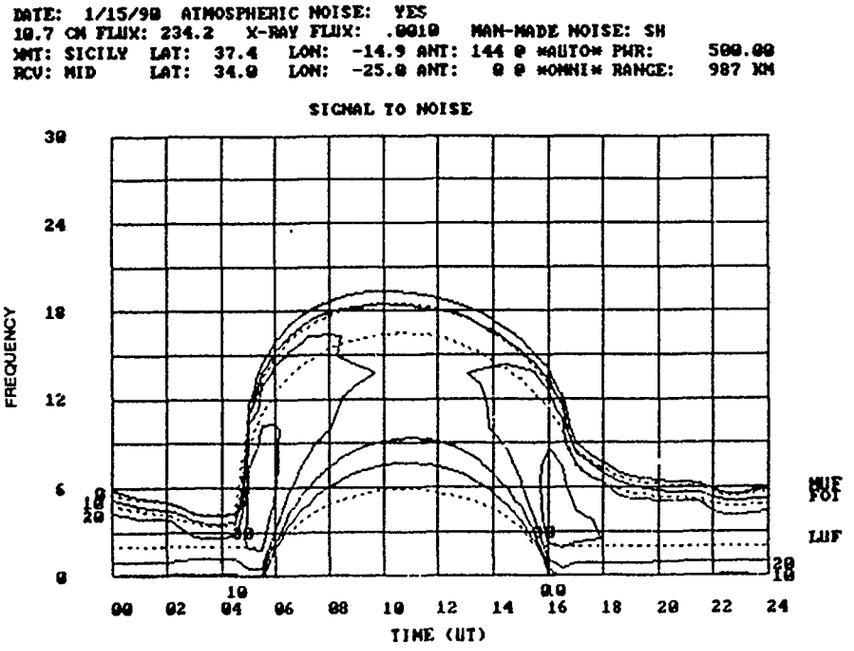


Figure 5. Signal-to-noise ratio for a path from Sicily to the mid-point of a radial to Beirut, Lebanon for PROPHET Version 3.2.

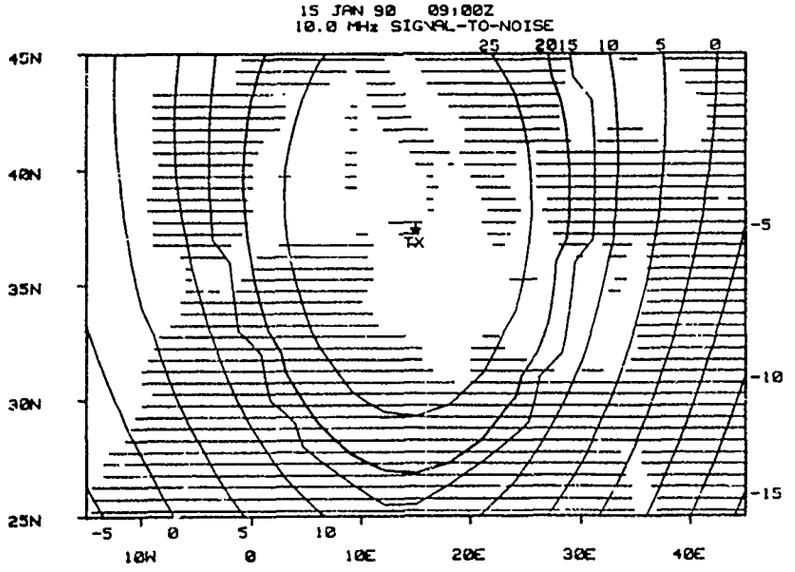


Figure 7. Map of signal-to-noise for a transmitter at Sicily for 15 January 1990 at 09:00Z at 10 MHz for the new field strength model in HFTDA.

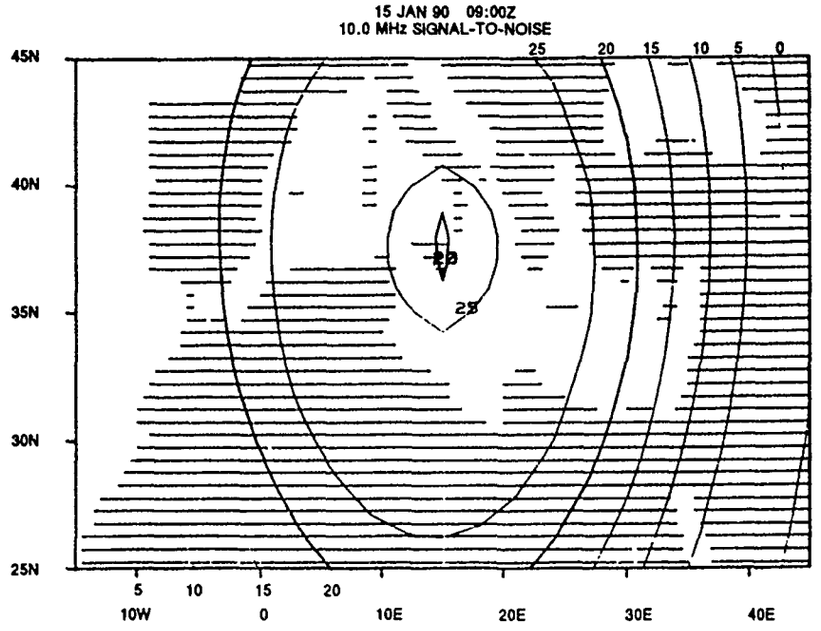


Figure 8. Map of signal-to-noise for a transmitter at Sicily for 15 January 1990 at 09:00Z at 10 MHz for the old field strength model in HFTDA.

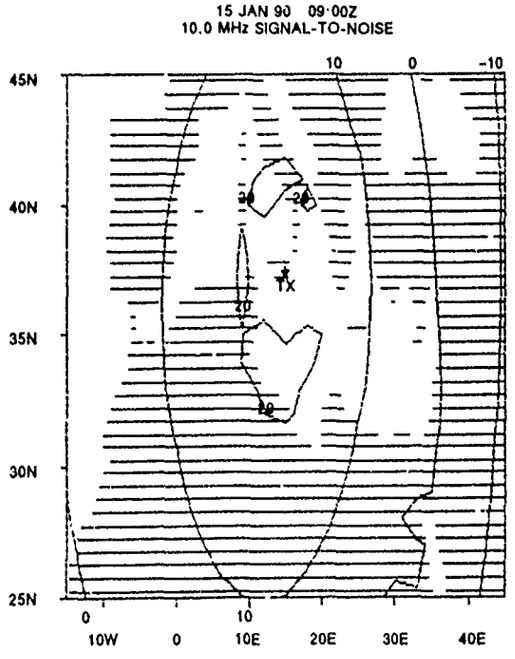


Figure 9. Map of signal-to-noise for a transmitter at Sicily for 15 January 1990 at 09.00Z at 10 MHz for PROPHET Version 3.2.

REFERENCES

Argo, P. E. and I. J. Rothmuller, (1979), "PROPHET: an application of propagation forecasting principles," Solar-Terrestrial Predictions Proc., Vol. 1, pp. 312- 321.

Barghausen, A. F., J. W. Finney, L. L. Proctor, and L. D. Schultz, (1969), Predicting long-term operational parameters of high frequency sky-wave telecommunications systems, Environmental Sciences Services Administration TR ERL-110-ITS-78.

Beckmann, B., (1967), Notes on the relationship between the receiving-end field strength and the limits of the transmission frequency range MUF-LUF, NTZ-Comm. J., Vol. 6, pp. 37-47.

Bradley, P. A. and R. Y. Liu, (1982), An evaluation of the FTZ sky-wave field-strength prediction method, Submission to the CCIR XVth Plenary Assembly, Study Group 6, IWP 6/1 Doc. 205, 26 July.

Damboldt, T., (1975), "A comparison between the Deutsche Bundespost ionospheric HF radio propagation predictions and measured field strength," Radio Systems and the Ionosphere, AGARD Conf. Proc. No. 173, Athens, Greece, 26-30 May. Propagation Effects, AGARD Conf. Proc. No. 453, San Diego, CA, 15-19 May.

Damboldt, T. and P. Suessmann, (1989), "The FTZ HF propagation model for use on small computers and its accuracy," Operational Decision Aids for Exploiting or Mitigating Electromagnetic Propagation Effects, AGARD Conf. Proc. No. 453, San Diego, CA, 15-19 May.

Gnession, J. R., D. J.Brandon, D. L.Lucas, (1988), Sounder update and field strength modifications for special operations radio frequency management system (SORFMS), Contract N66001-86-D-0079, 31 March.

Li, S. T., J. W. Rockway, J. C.Logan, and D. W. S.Tam, Microcomputer Tools for Communications Engineering, Artech House, Inc., Dedham, MA,1983.

Lucas, D. L. and G. W. Haydon, (1966), Predicting the statistical performance indexes for high frequency ionospheric telecommunication systems, Environmental Sciences Services Administration TR IER-1-ITSA-1,A Aug.

Richter, J. H., (1989), "Propagation assessment and tactical decision aids," Operational Decision Aids for Exploiting or Mitigating Electromagnetic Propagation Effects, AGARD Conf. Proc. No. 453, San Diego, CA, 15-19 May.

Richter, J. H., I. J. Rothmuller and R. B.Rose, (1977), "PROPHET. real time propagation forecasting terminal," Proc. 7th Technical Exchange Conference, Published by ASL, WSMR NM 88002, pp. 77-81, April.

Rose, R. B., (1981), "PROPHET—an emerging HF prediction technology," Symposium on the Effect of the Ionosphere on Radiowave Systems, Alexandria, VA, 14-16 April, pp. 534-542.

Rose, R. B., (1989), "PROPHET and future signal warfare decision aids," Operational Decision Aids for Exploiting or Mitigating Electromagnetic Propagation Effects, AGARD Conf. Proc. No. 453, San Diego, CA, 15-19 May.

Sailors, D. B. and W. K. Moision, (1987), Quiet time lowest observable frequency model uncertainty assessment, Naval Ocean Systems Center TR 1189, August.

Schulz, L. D. and R. M. Gallet, (1970), A survey and analysis of normal ionospheric absorption measurements obtained from radio pulse reflection, Environmental Sciences Services Administration Paper 4.

Sprague, R. A., (1987), Accuracy of field strength models: PROPHET and HFBC84, Naval Ocean Systems Center TR 1147,July.

Teters, L. R., J. L., Lloyd, G. W. Haydon, and D. L. . ucas, (1983), Estimating the performance of telecommunication systems using the ionospheric transmission channel-Ionospheric Communications Analysis and Prediction program, National Telecommunications and Information Administration Report 83-127.

Wakai, N., M. Ose, and K. Tanohata, (1971), Solar control of HF radio wave absorption in the nighttime, J. Radio Res. Labs., Japan, Vol. 18, pp. 1-17.

Wheeler, J. L., (1966), Transmission loss for ionospheric propagation above the standard MUF, Radio Sci., Vol. 1, pp. 1303-1308.

AD-P006 283

HF NOISE ENVIRONMENT MODELS

Catherine M. Keller
Lincoln Laboratory, Massachusetts Institute of Technology
Lexington, Massachusetts 02173-9108

INTRODUCTION

An L-shaped antenna array with a large aperture is used for characterizing the spatial distribution and temporal variability of the noise environment for several frequencies in the HF band. *Spatial distribution* of the noise environment refers to the distribution of noise power over all array azimuth and array elevation coordinates; that is, noise spatial distribution is the noise power distribution in space as seen at the array. *Temporal variability* indicates how the spatial distribution of the noise environment changes as a function of time. All of the sources of HF noise are likely to change in power and spatial distribution as a function of time of day and time of year. The noise may have different temporal characteristics for different frequencies. The noise sources are also likely to have different spatial distributions as a function of frequency.

In the temporal variability analysis of the noise environment, the task is to determine if the spatial distribution of the noise is ***stationary*** for a fixed frequency. The spatial distribution of the noise can be considered as a statistical random process, and it is deemed stationary if the statistics of the process are not dependent on time. One expects that the spatial distribution of the noise environment is not a strictly stationary process, and so an estimate of the rate of change is of interest. This estimate is obtained by forming a model from data samples and seeing how long the model remains a good representation of the noise spatial distribution. Stationarity is evaluated using several data records of noise samples taken at different times of day and on different days.

The analysis of the spatial distribution of the noise environment involves the use of simple symmetrical models that are discussed in the literature. These models are quantitatively compared to the data. The spatial distribution of noise power can be modeled via a spatial spectrum such as the *maximum likelihood method* (MLM) spatial spectrum. The flatness of the spatial spectrum can provide some indication of the spatial uniformity of the noise sources. There may be directions which contain a larger average noise power than others; perhaps knowledge of the sources of asymmetry in the noise spatial spectra can be used to improve the model for the noise as seen by the array.

There are several sources of HF noise that should be incorporated in the models. *Sky noise* is one source of noise at the antenna array. Sky noise includes solar, atmospheric, and cosmic noise, and is usually modeled as spherically or hemispherically isotropic noise; i. e., the model assumes far field uncorrelated sources of equal power placed everywhere on a sphere or hemisphere above the array. This type of noise produces correlation between any pair of array antennas with a value that depends on the distance (in wavelengths) between those antennas.

Another category of HF noise is *man-made* noise due to electric power generators, transformers, and transmission lines, and due to industrial equipment. The effect of man-made noise on the array depends on the location of the noise sources relative to the array—that is, the interference environment is locally unique. A simple model for the spatial distribution of man-made noise that belies the dependence on noise-source locations is a ring (of far field uncorrelated sources of equal power) in the plane of the array (on the horizon). This noise also produces antenna to antenna correlations.

Thermal noise in the antenna preamplifiers is another source of noise. This noise is usually modeled as uncorrelated between array antennas; this uncorrelated noise assumption is good if there is no coupling between the preamplifiers of pairs of antennas.

There may be many *discrete HF transmissions* that are always present or are present for long periods of time, such as coherent transmissions used in broadcast services, mobile radio services, radio navigation, control signals, and meteorological monitoring services. Such signals could be included in the HF noise model. The discrete sources produce correlated noise between each pair of array antennas. There may be other HF transmissions that are signals of interest, interferers of interest, or intermittent coherent transmissions, that should be excluded from the noise model. [1][2][3]

The array used in this study is an L-shaped array with 15 omnidirectional antennas located in Massachusetts. The aperture of the array is about 13.7 wavelengths at 8 MHz and is about 27.3 wavelengths at 16 MHz. When modeling the array, the

91-09676

antennas are assumed to be *isotropic*; that is, all the modeled array antennas are identical and have the same complex gain factor (assumed to be unity). Calibration measurements suggest that the antennas in the array are nearly isotropic.

The tools employed to study the noise environment all use data or models in the form of *spatial covariance matrices. Sampled covariance matrices* are obtained from samples, or *looks*, collected on the antenna array. The sampled spatial covariance matrices are based on a finite number of looks, such that the number of looks is greater than the number of array antennas.

Two types of *modeled covariance matrices* are considered. One type of modeled spatial covariance matrix is calculated from a model of the spatial distribution of the noise environment [4]. Such models are based on a set of independent far field narrowband discrete emitters of the same power covering the entire hemisphere in the *hemispherically isotropic model,* the horizon only in the *zero elevation model,* or a strip from the horizon up to a specified elevation in the *elevation spread model.* The spatial covariance matrices are easily derived from the symmetrical spatial distributions. This type of modeled spatial covariance matrix may also include contributions from thermal noise and discrete (in space) emitters. Comparisons between this type of modeled covariance matrix and the sampled covariance matrices derived from the data, quantify the validity of the spatial distribution model.

The other type of modeled spatial covariance matrix comes from a long-term average of the noise data. This type of model is useful for testing the noise stationarity.

AVAILABLE TOOLS

Several tools are available for analyzing the underlying noise environment from the information contained in the spatial covariance matrices. One tool is based on the eigenvalues and eigenvectors of the spatial covariance matrices. Let $\mathbf{R}$ denote the sampled spatial covariance matrix for the noise. If $\lambda_{\max}$ denotes the largest eigenvalue of $\mathbf{R}$ and $\lambda_{\min}$ denotes the smallest, then the *eigenvalue spread* of $\mathbf{R}$ is defined as

$$s \triangleq \frac{\lambda_{\max}}{\lambda_{\min}}. \tag{1}$$

The number s is also known as the *condition number* of the matrix $\mathbf{R}$ [5, page 364]. The eigenvalue spread is often expressed in decibel units $s_{\mathrm{dB}} \triangleq 10 \log_{10} s$ dB; the eigenvalue spread in dB is the difference $\lambda_{\max}$ (dB) $- \lambda_{\min}$ (dB).

To explain how the eigenvalue spread is related to antenna to antenna correlation, consider first a covariance matrix with all eigenvalues equal, so the eigenvalue spread is unity. It can be shown by using the intertwining eigenvalue property [5, page 350] that a Hermitian matrix has unity eigenvalue spread if and only if it is proportional to the identity matrix. Thus, if a covariance matrix has unity eigenvalue spread, then there is no antenna to antenna correlation. In many array processing algorithms, it is assumed that the noise background can be modeled by a matrix proportional to the identity matrix.

As an example, consider a hemispherically isotropic noise model. This model consists of a set of far field emitters that are narrowband, independent, and of equal power coming in at every direction on the hemisphere. When the covariance matrix for this model is derived, even assuming a perfectly calibrated array, it is not proportional to the identity. There is some degree of correlation between array antennas, and this depends on the sampled aperture used. Some of the antennas of an over-sampled aperture receive redundant information leading to higher antenna to antenna correlation than that between antennas receiving unique pieces of information. High antenna to antenna correlation gives a covariance matrix with large off-diagonal terms and such a covariance matrix has a large eigenvalue spread. If the covariance matrix has small eigenvalue spread, then all of the antenna to antenna correlations must be small. Thus, eigenvalue spread can be used to test whether the assumption of uncorrelated noise made in the analyses of most algorithms is valid.

Another tool is the process of *whitening.* One way to compare two matrices is to check the effect of whitening one with the other. Let $\mathbf{R}_M$ denote the modeled spatial covariance matrix. The whitening operation on the covariance matrix $\mathbf{R}$ with the modeled covariance matrix $\mathbf{R}_M$ leads to a whitened matrix $\mathbf{R}_W$ defined as

$$\mathbf{R}_W \triangleq \mathbf{R}_M^{-1/2}\mathbf{R}(\mathbf{R}_M^{-1/2})^H = \mathbf{R}_M^{-1/2}\mathbf{R}\mathbf{R}_M^{-1/2}, \tag{2}$$

since $\mathbf{R}_M$ is an Hermitian matrix. Of course, if $\mathbf{R} = \mathbf{R}_M$, then the result of the whitening is the identity matrix, and the eigenvalue spread is 0 dB. The whitening of $\mathbf{R}$ by $\mathbf{R}_M$ is considered effective if, as the number of looks used to construct $\mathbf{R}$ goes to infinity, the whitened covariance matrix becomes proportional to the identity and the eigenvalue spread of the whitened matrix goes to unity.

When a finite number of looks are available to construct $\mathbf{R}$, this convergence is not complete and the whitened matrix is not an identity matrix. It is appropriate to estimate the expected deviation of the whitened matrix from an identity matrix, and thus, the eigenvalue spread from unity, as a function of the number of looks. There is an approximation for the expected eigenvalue spread of a sampled covariance matrix that depends on the number of antennas and on the number of samples used in its construction. This result is based on a lemma given by Edelman [6]. Let $\mathbf{R} = \sum_{k=1}^{L} \mathbf{z}_k \mathbf{z}_k^H$ be a complex N by N matrix with a Wishart distribution, so that the $\mathbf{z}_k$ are independent sample vectors with zero mean, jointly complex Gaussian components, and with $E[\mathbf{z}_k \mathbf{z}_k^H] = \mathbf{I}_N$, the N by N identity matrix. N is the number of antennas in the array, and L is the number of looks used to construct $\mathbf{R}$. Assume that $L \geq N$. Then the approximate values of the minimum and maximum eigenvalues of $\mathbf{R}$ are

given by

$$\frac{\lambda_{max}}{N} \approx (1+\sqrt{\frac{N}{L}})^2 \tag{3}$$

$$\frac{\lambda_{min}}{N} \approx (1-\sqrt{\frac{N}{L}})^2. \tag{4}$$

Thus, the eigenvalue spread approximation is

$$s_E = \frac{(1+\sqrt{\frac{N}{L}})^2}{(1-\sqrt{\frac{N}{L}})^2}. \tag{5}$$

For a fixed number of array antennas, as the number of looks increases, the eigenvalue spread approximation decreases to unity (0 dB). The eigenvalue spread approximation is 6.56 dB for a covariance matrix constructed for an isotropic array with 13 antennas and using 100 looks; it is 7.10 dB for an array with 15 antennas. If $\mathbf{R} \neq \mathbf{R}_M$ and yet the whitening decreases the eigenvalue spread of the sampled covariance matrix to the estimated eigenvalue spread, then one can conclude that the model fits the data well.

The eigenvalue spread approximation in (5) only accounts for the finite number of looks used to construct R; if whitening is performed by a matrix $\mathbf{R}_M$ constructed from an infinite number of looks, the approximation is expected for $\mathbf{R}_W$ as well. However, the expected eigenvalue spread of a whitened matrix can be affected if $\mathbf{R}_M$ is constructed from a finite number of looks. A set of simulations was run to show that it is reasonable to use the eigenvalue spread approximation when the number of looks used to construct $\mathbf{R}_M$ is above 600.

There are a few other guidelines for what to expect for the eigenvalue spread of a whitened covariance matrix. By using (2) and the extremal properties of eigenvalues of Hermitian matrices [7, pages 176-177], it is easy to show that

$$1 \leq s_W \leq s \cdot s_M \tag{6}$$

where s_W, s, and s_M are the eigenvalue spreads of $\mathbf{R}_W$, $\mathbf{R}$, and $\mathbf{R}_M$, respectively. Further, it can be shown that the inequality

$$s_W \geq \frac{s}{s_M} \tag{7}$$

holds. That is, whitening with the modeled covariance matrix $\mathbf{R}_M$ does not decrease the eigenvalue spread of R by more than the eigenvalue spread of $\mathbf{R}_M$, s_M. The inequalities in (6) and (7) are illustrated in the examples to appear later.

Another available tool is the continuous spatial spectrum of the spatial covariance matrix. The MLM spectrum of Capon is a reasonable choice, because it gives true power estimates. The MLM spectrum corresponding to a covariance matrix R is given by

$$P(\alpha,\Delta) = \frac{1}{\vec{v}^H(\alpha,\Delta)\mathbf{R}^{-1}\vec{v}(\alpha,\Delta)}, \tag{8}$$

where α is the azimuth coordinate and Δ is the elevation coordinate. The k-th element of the steering vector $\vec{v}(\alpha,\Delta)$ is given by

$$v_k = e^{j\frac{2\pi}{\lambda}\vec{d}_k^T\vec{u}} \tag{9}$$

where $\vec{d}_k^T = (x_k, y_k)$ is the transpose of the position vector for the k-th antenna, $\vec{u}^T = (u_x, u_y)$, $u_x = \cos\Delta\cos\alpha$, and $u_y = \cos\Delta\sin\alpha$. To obtain a spectral value for a particular direction, the MLM spectral algorithm finds the array weights that give unity gain in that direction while minimizing the array output power. The MLM spectrum can indicate to what extent the environmental noise varies over all azimuth and elevation angles—i.e., how flat the MLM spectrum is. The flatness of the spatial spectrum provides an indication of the spatial uniformity of the noise sources.

STATIONARITY TESTS

To determine how stationary the HF noise environment is at a particular frequency, a sampled covariance matrix is formed starting at a selected time in a chosen data record. This matrix is used to model the noise environment at the selected time and is called the modeled covariance matrix $\mathbf{R}_M$. It is then compared to sampled covariance matrices taken at different times within the same data record, sampled covariance matrices taken from a different time of day, and sampled covariance matrices taken from a different day at approximately the same time of day.

In some of the examples, $\mathbf{R}_M$ is constructed from 1200 consecutive looks, which corresponds to 1.56 seconds. In other examples 800 looks is used for $\mathbf{R}_M$, which corresponds to 1.04 seconds. Covariance matrices R, based on a smaller number of consecutive looks (e.g., 100) starting at various times within the data records, are formed.

Each R is then whitened with $\mathbf{R}_M$. The modeled covariance matrix $\mathbf{R}_M$ is a good representation of the noise over a certain period of time if the eigenvalue spread is reduced by the whitening process ($s_W < s$) and, according to (5), $s_W \approx 6.56$ dB (13 antenna array) or $s_W \approx 7.1$ dB (15 antenna array) for that period of time. The time for which the modeled covariance matrix reduces the eigenvalue spread quantifies the noise stationarity.

Examples at 16.1 MHz

Figure 1(a) shows a discrete sequence of eigenvalue spreads before and after whitening. The *sampled* covariance matrices are constructed from 100 consecutive looks of the noise data from the same data record as the modeled covariance matrix. The modeled covariance matrix is based on a data record at 16.1 MHz from February 26, 1988 at 16:46; it is constructed from 1200 consecutive looks starting at 1 second into the data record. The 100 look sampled covariance matrices to be whitened are taken starting at every half of a second into the data record; the 100 looks begin at the times shown on the abscissa below the curve symbols. 13 array antennas are used in the example.

It is clear from Figure 1(a) that the whitened covariance matrices have smaller eigenvalue spreads than the unwhitened covariance matrices, and in many cases, the eigenvalue spreads of the whitened covariance matrices are nearly down to the 6.56 dB eigenvalue spread approximation. Thus, the modeled covariance matrix is a good representation of the noise environment over much of the 16 second data record.

There are several peaks in the curves of Figure 1(a) indicating that the modeled covariance matrix does not represent the noise environment well; these peaks actually correspond to times when there is some activity on the channel due to intermittent emitters. At about 7.5 and 12.5 seconds into the data record, the signal level increases, indicating the presence of interference that is not well modeled by the fixed long-term average covariance matrix taken earlier in the data record.

Figure 1(b) shows the sequence of eigenvalue spread results for a similar whitening process when the sampled covariance matrices are based on a data record taken six hours earlier than the modeled covariance matrix. The whitening process no longer appears effective, as the eigenvalue spreads of the whitened matrices are not as low as expected. However, the eigenvalue spreads of the whitened matrices are lower than the eigenvalue spreads of the unwhitened matrices most of the time (up to 4 dB lower). Thus, some part of the noise structure seems to remain the same for the 6 hour period.

Note that the eigenvalue spreads of the unwhitened covariance matrices in Figure 1(b) are about 3 dB lower than those taken from the later data record in Figure 1(a). The increase of the power on the channel can be attributed to an increase in emitter activity and this causes an increase in the largest eigenvalue of the sampled covariance matrix. The smallest eigenvalue is not increased significantly; thus, the eigenvalue spread is larger. Also, the peak in the eigenvalue spread sequence in Figure 1(b) corresponds to activity on the channel at 7.5 to 8 seconds.

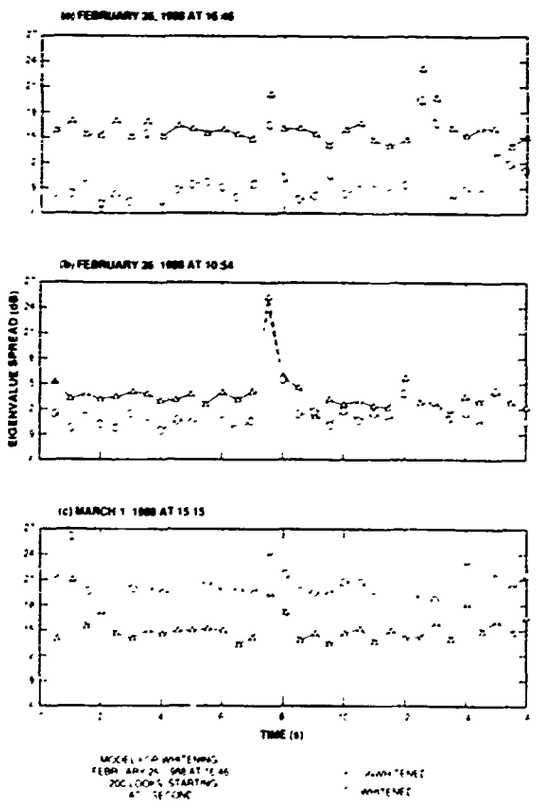


Figure 1: Eigenvalue spreads of unwhitened and whitened sampled covariance matrices constructed from 100 looks collected at 16.1 MHz using a 13 element array.

Figure 1(c) compares similar whitened and unwhitened results for a data record taken 4 days later, at approximately the same time of day. The modeled covariance matrix does not fit the data—the eigenvalue spreads of the whitened matrix are significantly larger than those of the unwhitened matrix (6-7 dB higher). This example indicates that the spatial distribution of the noise environment at 16.1 MHz can not be formulated as a function of the time of day. The spatial distribution of the noise environment may have continuously changing components that dominate the time-of-day structure.

Examples at 8.063 MHz

The examples in this section show that the noise at 8.063 MHz is fairly stationary. This conclusion is based on the noise samples taken within intervals without the intermittent emitters that frequently appeared. Figure 2 shows four sequences of eigenvalue spreads of unwhitened and whitened covariance matrices. The modeled covariance matrix is constructed from 800 looks in a data record from February 26, 1988. Figure 2(a) gives eigenvalue spreads of 100-look sampled covariance matrices from the same data record as the model. A 13 antenna array is used; thus, 6.56 dB eigenvalue spread is expected. Note that the whitening is very good throughout the 16 seconds except for times in which an intermittent emitter is on (8 seconds and 11 seconds). Figure 2(b) is a sequence of eigenvalue spreads of unwhitened and whitened sampled covariance matrices taken from a data record that is more than two hours later. The whitening process brings the eigenvalue spread down from 16-18 dB to 9-10 dB.

Figure 2(c) is a sequence of eigenvalue spreads of unwhitened and whitened sampled covariance matrices taken from a data record collected 5 days later at around the same time of day as the model. Again, the whitening brings the eigenvalue spreads down from 15-18 dB to 9-10 dB throughout most of the 16 second data record. Figure 2(d) has the sampled covariance matrices constructed from a data record taken nearly a year later than the model. Although the conclusion is that the model is no longer

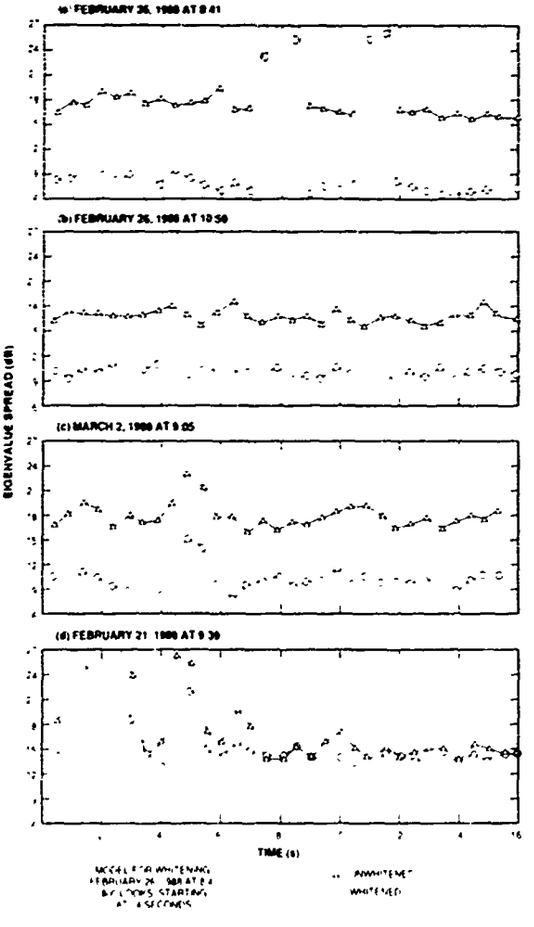


Figure 2: Eigenvalue spreads of unwhitened and whitened sampled covariance matrices constructed from 100 looks collected at 8.063 MHz using a 13 element array.

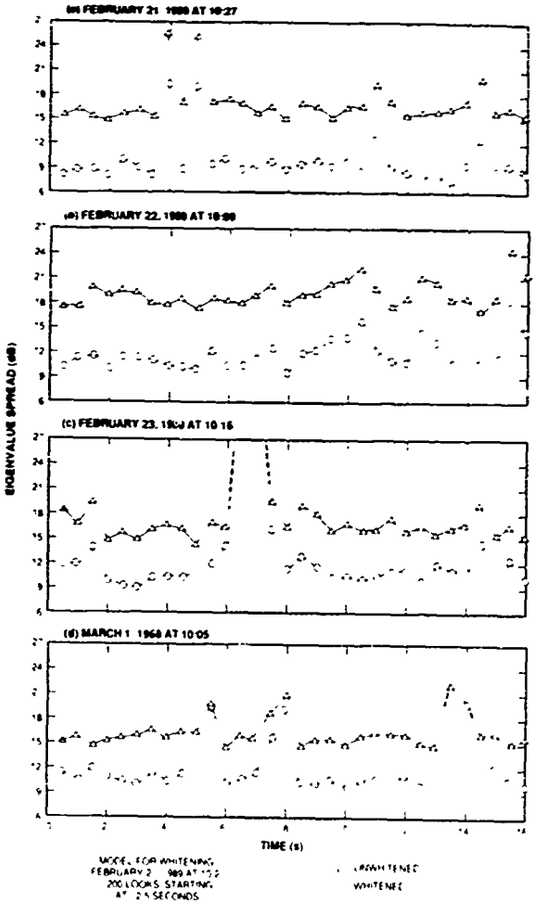


Figure 3: Eigenvalue spreads of unwhitened and whitened sampled covariance matrices constructed from 100 looks collected at 10.648 MHz using a 13 element array.

a good fit, it is encouraging that most of the time, the eigenvalue spreads of the whitened sampled covariance matrices are lower than the eigenvalue spreads of the unwhitened matrices!

It should be noted that the noise data collected is dominated by environmental noise, and not receiver thermal noise; the thermal noise is estimated to be about 7 dB below the environmental noise. A conjecture is that the noise environment at 8 MHz is more likely to be dominated by sky noise than the noise environment at 16 MHz, and it is the sky noise that remains fairly stationary.

Examples at 10.648 MHz

Figure 3 show stationarity tests for noise data at 10.648 MHz. Figure 3(a) shows a sequence of eigenvalue spreads of unwhitened and whitened covariance matrices for a model constructed from 1200 looks in the same data record. Figures 3(b) and 3(c) show sequences of eigenvalue spreads of unwhitened and whitened covariance matrices taken one and two days later than the model. The whitening is very good. Figure 3(d) shows a sequence of eigenvalue spreads for unwhitened and whitened matrices when the data set is nearly a year earlier than the model. Again, the whitening is good throughout most of the data record: although the eigenvalue spreads are not as low as expected, the whitening results in more than a 5 dB reduction in eigenvalue spread.

NOISE ENVIRONMENT MODELS

Because the noise environment at 8 and 10 MHz exhibits some stationarity, it is reasonable to try to find a model that fits the data. The conjecture that the noise environment at these frequencies is dominated by sky noise can be tested by comparing data to appropriate models. Modeled covariance matrices for the noise floor are found by adding together modeled covariance matrices for sky noise, man-made interference, and thermal noise. Symmetrical spatial distribution models for sky noise and man-made noise are discussed in the literature [1] [4]. The modeled covariance matrices are found by constructing noise fields from uncorrelated narrowband discrete emitters distributed in some way on a far-field hemisphere surface with the antenna array at the center. Four noise fields were investigated in this study:

- Hemispherically isotropic—the emitters are uniformly distributed on the surface of the hemisphere.
- Zero elevation—the emitters are all located at 0° elevation, uniformly spaced on the horizon.
- Elevation spread—the emitters are uniformly distributed in a strip of elevations 0° to 90°ϵ, where $0 < \epsilon < 1$, on the surface of the hemisphere. $\epsilon = 1/4$ for many of the tests.
- Asymmetric elevation spread—the hemisphere is partitioned into four azimuth segments going from 0° to 90°ϵ elevation in which emitters are uniformly distributed with different powers in different segments. MLM spectral plots indicate how to do the partitioning.

Each array configuration is assumed to be calibrated and the array antennas are assumed to be isotropic. Models for arrays with antennas that are not isotropic, but which have calibration data available, can also be employed. Such models are not included in this paper.

The form of the covariance matrices and their eigenvalue spreads for the noise fields listed above are of interest here. It is easy to derive the covariance between two antennas spaced $d_{i,i'}$ meters apart for each of the four noise fields. These pair-wise covariances are entries $R_{i,i'}$ in the normalized noise field spatial covariance matrix, where i and i' are the row and column of the normalized spatial covariance matrix $\mathbf{R}$. The parameter λ is the wavelength in meters. For the hemispherically isotropic noise field,

$$R_{i,i'} = \sin(\frac{2\pi d_{i,i'}}{\lambda})/(\frac{2\pi d_{i,i'}}{\lambda}). \tag{10}$$

The modeled spatial covariance matrix entries for the zero elevation environmental noise model are

$$R_{i,i'} = J_0(\frac{2\pi d_{i,i'}}{\lambda}). \tag{11}$$

For the elevation spread environmental noise model,

$$R_{i,i'} = \int_0^{\frac{\pi\epsilon}{2}} \cos\Delta J_0(\frac{2\pi d_{i,i'}}{\lambda}\cos\Delta)d\Delta. \tag{12}$$

The covariance matrix entries for the asymmetric elevation spread noise model is a sum of four double integrals over azimuth and elevation coordinates, one double integral per segment.

Table 1 lists eigenvalue spreads for modeled covariance matrices for arrays with 3 and 13 antennas, both with full apertures. In general, for a given aperture, the eigenvalue spread tends to increase as the number of antennas filling in the array increases. The aperture is oversampled so that there is some redundant information at the antenna outputs. The table shows that the

TABLE 1
Eigenvalue spread (dB) for three noise models.

Frequency (MHz)	Number of Antennas	Noise Model		
		Hemispherically Isotropic	Zero Elevation	Elevation Spread $\epsilon = 1/4$
8.063	3	0.31	0.96	1.01
10.648	3	0.23	1.35	0.78
16.1	3	0.11	1.00	0.20
8.063	13	19.65	20.55	21.04
10.648	13	10.82	9.40	9.74
16.1	13	3.90	9.04	8.42

sparse array at 8.063 MHz has a small eigenvalue spread—the array is are not over-sampled. The array with 13 antennas has a large eigenvalue spread; the antennas are closely spaced relative to the wavelength. This results in an over-sampled array and the noise is correlated between any pair of the closely spaced antennas. This behavior is similar for all three environmental noise models.

Table 1 also indicates that typically the eigenvalue spread decreases as the frequency increases. At higher frequencies and shorter wavelengths, the array configurations are not as over-sampled as they are for lower frequencies. Exceptions to the trend are possible for particular antenna spacings. For example, an array operating at a frequency so that the distance between adjacent antennas is a multiple of half a wavelength has 0 dB eigenvalue spread for the hemispherically isotropic environmental noise model. The same array operating at either a smaller or larger frequency has a larger eigenvalue spread.

It is easy to include thermal noise in the model for the noise. A new modeled covariance matrix that includes thermal noise that is the same at each antenna can be formed by taking $\mathbf{R} + \sigma^2\mathbf{I}$ where σ^2 is the thermal noise power relative to the environmental noise. Thus, since the diagonal entries of $\mathbf{R}$ are unity, adding thermal noise that is 7 dB down from the environmental noise requires $\sigma^2 = 0.1995$.

TABLE 2
Eigenvalue spreads (dB) for modeled covariance matrices constructed from the hemispherically isotropic noise field and thermal noise models.

Frequency (MHz)	No Thermal Noise	Thermal Noise		
		-10.0 dB	-7.0 dB	-3.0 dB
8.063	19.65	12.71	10.33	7.13
10.648	10.82	8.94	7.74	5.68
16.1	3.90	3.50	3.18	2.49

The eigenvalue spreads of modeled covariance matrices with thermal noise included are lower than the eigenvalue spreads of the corresponding modeled covariance matrices with σ^2 set to 0 ($= -\infty$ dB). The eigenvalues of $\mathbf{R} + \sigma^2\mathbf{I}$ are simply the eigenvalues of $\mathbf{R}$ plus σ^2. Since $\lambda_{max} > \lambda_{min}$, it is easy to see that

$$\frac{\lambda_{max}}{\lambda_{min}} \geq \frac{\lambda_{max} + \sigma^2}{\lambda_{min} + \sigma^2}. \tag{13}$$

Table 2 lists eigenvalue spreads for modeled covariance matrices constructed from the hemispherically isotropic noise field and thermal noise models for the array with 13 antennas, illustrating (13).

COMPARING HF DATA TO THE MODELS

Next some results on trying to whiten the HF data with the environmental noise field models are presented. Figures 4(a) and 4(b) show a sequence of eigenvalue spreads for sampled covariance matrices whitened by the hemispherically isotropic noise field modeled covariance matrices with several different levels of thermal noise included. Thermal noise at $-\infty$ dB corresponds to no thermal noise in the model. The model is not a good fit to the data; the eigenvalue spread increases after whitening is performed. A model with the thermal noise only 3.0 dB below the environmental noise does the best whitening most often in the example of Figure 4(a). Increasing the thermal noise level above -3.0 dB gives a model with a poorer fit to the data. A model with the thermal noise 6.0 dB below the environmental noise does the best whitening most often in the example of Figure 4(b).

Models consisting of the zero elevation or elevation spread noise fields and thermal noise show even poorer fits to the data than the hemispherically isotropic noise models. Thus, figures concerning these models are not shown.

Whitening with some of the models does reduce the eigenvalue spread. However, the eigenvalue spreads of the whitened matrices are expected to be near 7.1 dB. Figure 4(a) shows that there are differences of 4.0 to 6.0 dB between the eigenvalue

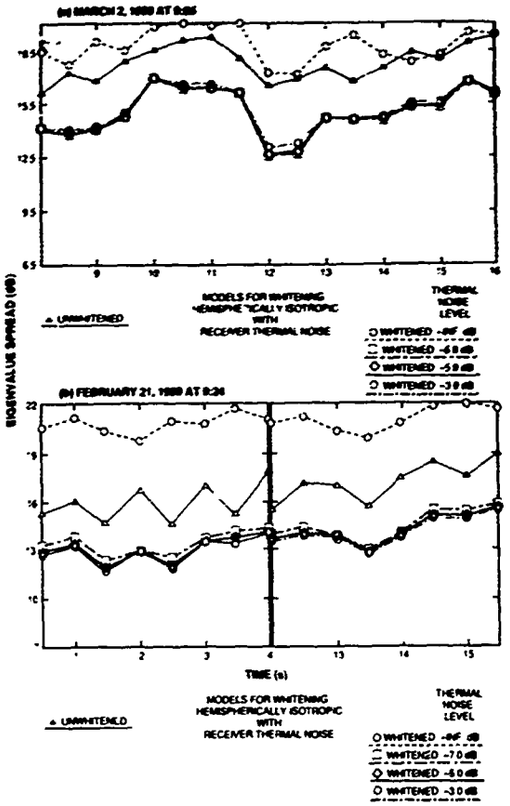


Figure 4: Eigenvalue spreads of unwhitened and whitened sampled covariance matrices constructed from 100 looks collected at 8.063 MHz using a 13 element array for (a) and a 15 element array for (b). Each model consists of both hemispherically isotropic noise and thermal noise.

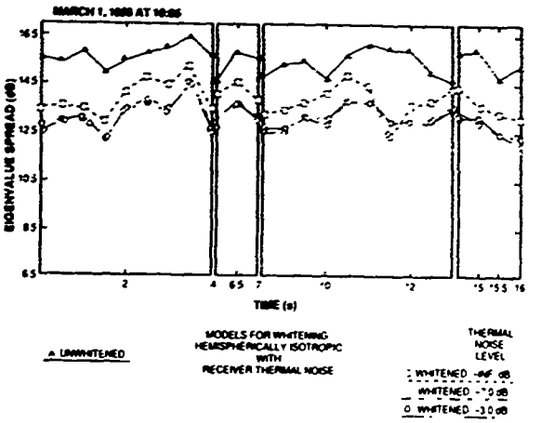


Figure 5: Eigenvalue spreads of unwhitened and whitened sampled covariance matrices constructed from 100 looks collected at 10.648 MHz using a 13 element array. Each model consists of both hemispherically isotropic noise and thermal noise.

spreads of the whitened matrices and the eigenvalue spread approximation of 7.1 dB for 15 antennas. Figure 4(b) shows differences of 6 to 10.5 dB between the eigenvalue spreads of the whitened matrices and the eigenvalue spread approximation of 6.56 dB (13 antennas). Thus, the results of the hemispherically isotropic noise field and thermal receiver noise model are only mildly encouraging.

Figure 5 compares models to 10 MHz data. Often, the whitening with the models brings the eigenvalue spread down by 2-3 dB.

Figure 6 shows sequences of eigenvalue spreads for sampled covariance matrices whitened with hemispherically isotropic noise field models with several different thermal noise levels. The best results are for the model without thermal noise included. This should not necessarily be interpreted as the channel at 16.1 MHz being strictly environmental noise limited. Instead, this result may be related to the low values that the eigenvalue spreads of the modeled covariance matrices have at 16.1 MHz. That is, the eigenvalue spread of the modeled covariance matrix without any thermal noise included is about 3.9 dB. The modeled covariance matrices with a non-zero thermal noise level included have even smaller eigenvalue spreads. For thermal noise levels 7.0 dB and 3.0 dB below the environmental noise, the eigenvalue spreads of the modeled covariance matrices are 3.18 dB and 2.49 dB, respectively. The amount of the decreases in eigenvalue spread shown in Figure 6 for the various thermal noise models is limited by the eigenvalue spread of the modeled covariance matrices.

MLM SPECTRUM

Figure 7(a) shows the MLM spectrum contours for a 100-look sampled covariance matrix taken from a 10.648 MHz data record. The difference between the maximum and minimum contour levels is only 5.57 dB; i.e. the noise power, when averaged over 0.13 seconds, does not differ by more than 5.57 dB over all azimuth-elevation coordinates. The regions from 80° to 160° and near 20° at elevations up to 40° are regions with the higher power levels.

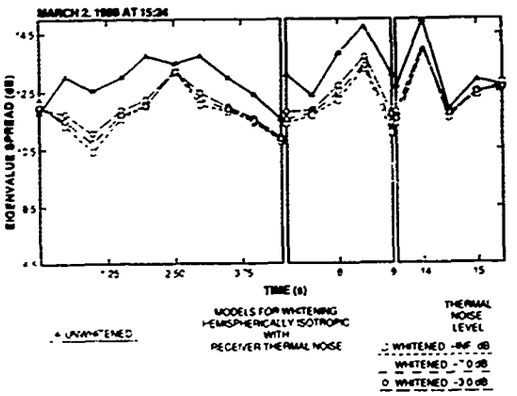


Figure 6: Eigenvalue spreads of unwhitened and whitened sampled covariance matrices constructed from 100 looks collected at 16.1 MHz using a 13 element array. Each model consists of both hemispherically isotropic noise and thermal noise.

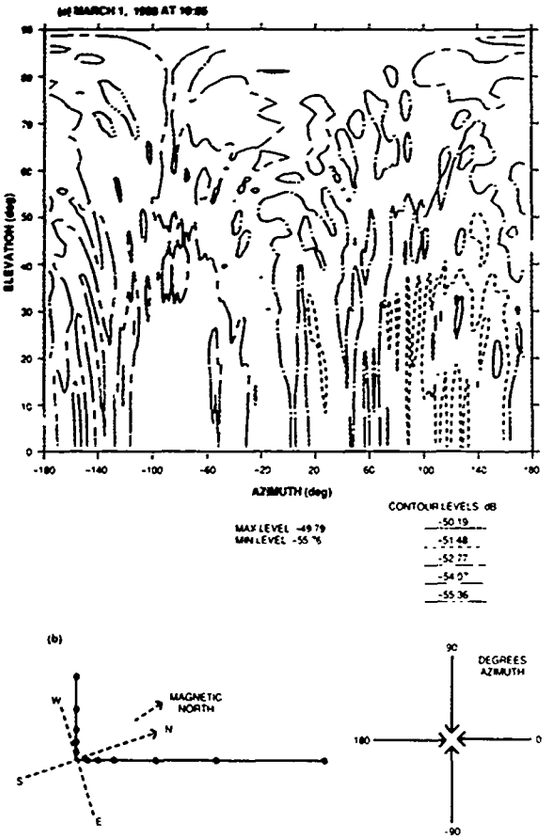


Figure 7: (a) MLM spatial spectrum contours of a sampled covariance matrix constructed from 100 looks starting at 7 seconds in the data record. The data record was collected at 10.648 MHz using a 13 element array. (b) Orientation of the array.

The minimum and maximum spectral contour levels do depend on the number of looks used to construct **R**. In general, the difference in the levels decreases as the number of looks increases—the discrete power sources are averaged over a longer time interval as the number of looks increases. The longer time averaging tends to decorrelate the sources at the array.

Many other MLM spectral contour plots were examined for data at 10.648 MHz as well as for data at 8.063 and 16.1 MHz. All of the contour plots have regions of higher power at nearly the same azimuth coordinates as the plot in Figure 7(a). By looking at MLM spectral contour plots for symmetrical noise field models, like the hemispherically isotropic model, it is verified that the higher power regions are not a figment of the array geometry. The higher power regions are attributed to noise sources. Thus, it is interesting to see what physical directions the higher contour levels of the MLM spectrum indicate. Figure 7(b) is a picture showing the orientation of the array in true and magnetic coordinates and the relationship of these coordinates to array azimuth coordinates. The region between 80° and 160° azimuth in array coordinates is in line with true west and southwest directions—these coordinates point down the east coast, including New York City, Baltimore, and Washington D.C. The region between 10 and 30° azimuth is approximately true north of the array. The city of Quebec, Canada, on the Saint Lawrence River is in this direction, and much further north is the Hudson Strait. It is possible that the higher noise power in these directions is due to ship navigation and communication systems; these systems are known to use several bands below 30 MHz [3]. Cities close to the array without a line-of-sight path, like Boston and Worcester, do not have mode support at the array; the signal paths would have to be almost vertical to arrive at the array.

It is clear that the environmental noise field is not strictly hemispherically isotropic, although the difference between maximum and minimum contour levels is only about 5 dB for 100 looks. Knowing that the actual noise field is not hemispherically isotropic and knowing what directions exhibit higher power than others should lead to better models. Asymmetric elevation spread noise models were tried that break the hemisphere up into four segments with different power levels. MLM spectra were used to determine the boundaries of the segments. Thermal noise was included in the models. Unfortunately, none of the asymmetric models tried did a better job of whitening the data than the hemispherically isotropic noise models.

CONCLUSIONS

Results have been presented on the characterization of the HF noise environment. It is shown that the noise environment is sufficiently stationary for noise models to be useful. Of the noise models compared, the model consisting of both the hemispherically isotropic noise field and thermal noise is the best fit to the HF data. The MLM spectra indicate that the actual noise field is not hemispherically isotropic. Thus, models were tried that have, in addition to the sky noise and thermal noise components, components that account for the asymmetry in the noise field due to long-term discrete HF transmissions that are present in certain directions. However, these models did not work better than the hemispherically isotropic noise models.

All of the data records examined in this report were taken on winter days. The noise environment on winter days in the continental United States may be the easiest to characterize. Due to sunspot activity at its maximum in winter, there is more available bandwidth to be shared by the emitter population; the emitter population in a narrow frequency band will be sparser in winter than in summer. Also, the emitter population in the continental United States is not as large as that in Europe. The characteristics of the HF channel in the continental US on a summer evening may better emulate the European HF channel [8]. Thus, it is of interest to study the HF noise environment using summer evening data.

ACKNOWLEDGEMENT

This work was performed under Air Force Contract No. F19628-90-C-0002 and is sponsored by the Department of Defense.

References

[1] J. E. Hudson, *Adaptive Array Principles*, pp. 55–58. New York: Peter Peregrinus, Ltd. and the IEEE, 1981.

[2] E. N. Skomal, *Man-Made Radio Noise*. New York: Van Nostrand Reinhold Co., 1978.

[3] E. N. Skomal and A. A. Smith, Jr., *Measuring the Radio Frequency Environment*. New York: Van Nostrand Reinhold Co., 1985.

[4] B. F. Cron and C. H. Sherman, "Spatial-correlation functions for various noise models," *J. Acoust. Soc. Amer.*, vol. 34, pp. 1732–1736, Nov. 1962.

[5] G. Strang, *Linear Algebra and Its Applications*, 3rd edition. San Diego: Harcourt Brace Jovanovich, 1988.

[6] A. Edelman, "Eigenvalue and condition numbers of random matrices," *SIAM J. Matrix Analysis and Applications*, vol. 9, pp. 543–560, Oct. 1988.

[7] R. A. Horn and C. R. Johnson, *Matrix Analysis*. New York: Cambridge University Press, 1985.

[8] B. D. Perry and R. Rifkni, "Interference and wideband communications," *Proc. Ionospheric Effects Symposium: The Effect of the Ionosphere on Communication, Navigation, and Surveillance Systems*, pp. 45–54, May 1987.

AD-P006 284

PROBABILITY DISTRIBUTION OF RECEIVED INTERFERENCE LEVELS IN THE HF BAND

Howard E. Nichols
Dennis J. Gooding
Engineering Research Associates
275 Wyman Street
Waltham, MA 02154

ABSTRACT

HF interference typically spans a wide dynamic range and often has a high density of detectable interferers. In the design and evaluation of communications systems it is often useful to be able to characterize the HF environment by means of an analytic model. Based upon a series of recent measurements of the HF environment, we have concluded that a lognormal distribution is a reasonable model of the observed HF interference. These data were collected using a wideband (2 MHz), high dynamic range spectrum analysis system, and consisted of short-term time averages of the magnitudes of FFT bin outputs spaced at 610 Hz. Spectral plots and associated probability distributions are presented spanning a frequency range from 1.5 to 19 MHz for various averaging parameters. In almost all cases, the distributions of the bin averages were found to be well modeled by the lognormal distribution for levels well above the background noise. In some cases, the fit was excellent over a 50 dB range of levels.

INTRODUCTION

While it is well known that received interference levels in the HF band typically span a wide dynamic range and the density of detectable interferers often is very high in active portions of the band, there seems to be relatively little quantitative data on the probability distribution of short-term average interference power over wide bandwidths. Furthermore, some of the reported results appear on the surface, at least, to be contradictory in that they seem to show different models for the interference cumulative distribution function (CDF).[1,2] Gibson and Arnett found that their data generally fit a lognormal distribution quite well, whereas Perry and Abraham, using somewhat different test conditions, obtained results that were linear functions of the form $p(x) = Ax^B$ over a wide range of interference power level x.

We have recently performed independent measurements that tend to support a lognormal model of interference CDF. Measurements of the CDFs over a 2 MHz bandwidth made in both summer and winter of 1989 near Boston, Massachusetts generally showed a good fit to lognormal distributions with standard deviation parameters ranging from about 19 to 23 dB. Furthermore, our results indicate that the CDF for a given 2 MHz band is almost unaffected by the averaging time of individual frequency bins and that they are stationary over periods of at least several minutes.

91-09677

The experimental equipment and the test approach used are described in the section below. In the next section, we present the results of the experimental measurements and in fourth section we discuss a possible rationale for the observed lognormal distribution of interferers. The conclusions of the research are presented in the final section.

EXPERIMENTAL SETUP AND APPROACH

Figure 1 shows a simplified block diagram of the experimental equipment. A 20-foot vertical whip antenna feeds a wideband HF receiver whose output is a band slightly wider than 2.0 MHz centered at 1.25 MHz. The result is digitized at a 5 MHz rate using a 14-bit A/D converter, blocked into 50% overlapping 8192-point blocks, windowed, and processed in real time using an 8192-point FFT processor. The magnitudes of the resulting frequency bin outputs are calculated and, optionally, are averaged over multiple FFTs to produce a single sample (revisit) of the spectrum. In the measurements presented here, only the central 2 MHz portion of the band was retained to assure that the receiver gain was flat over the analysis band. The resulting spectra are saved to a computer disk file for later non-realtime processing.

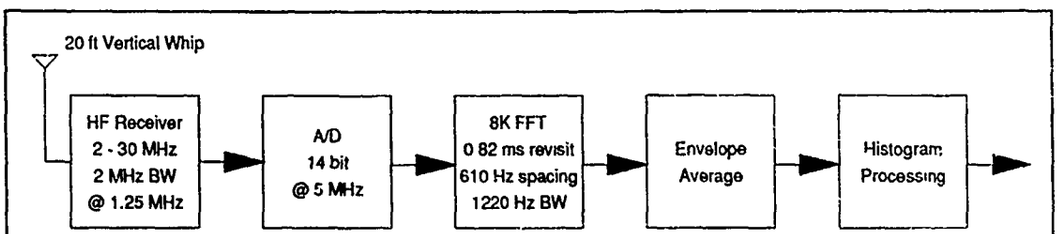


Figure 1. Block Diagram of Experiment Setup

The system is designed to provide an instantaneous dynamic range of about 120 dB at the output of the FFT. Before each collection, the input attenuator is adjusted manually to assure that overload does not occur in the receiver or the A/D converter and that the total output noise is dominated by input noise.

Because of the overlapped operation, a new FFT is produced every 0.82 ms. The resulting frequency bin spacing is about 610 Hz, but the effective bin width is about twice this because of the window, which is a Nuttal window with -92 dB worst-case sidelobes. The non-realtime processing consists of converting the averaged amplitude values into decibel values and forming histograms of all or downsampled portions of the data.

Data were collected in April and December of 1989. The earlier April measurements consisted of one second averages of magnitude spectra made at center frequencies 6, 8, 10, 12, 14, 15, 16, 18, and 22 MHz. The subsequent December collections were more comprehensive in that they used longer runs and used a variety of spectral averaging times. The December collections included experiments with a total observation interval spanning 7.1 minutes, measurements made at 5, 7, 9, 11, 13, 15, 17 and 19 MHz and envelope averaging times of 209 ms (256 FFTs). The December collections also included a separate series of collections made at 15 MHz using envelope averages over 1, 16, and 256 successive FFTs. The observation interval for these experiments also spanned 7.1 minutes with the total of the three experiments conducted within a 30 minute period. The data presented herein is primarily selected from the December experiments, although the conclusions are based upon both sets of data.

EXPERIMENTAL RESULTS

Figures 2 and 3 show typical examples of the CDFs obtained at 9 and 11 MHz plotted on a normal probability scale, and Figures 4 and 5 show snapshots of the corresponding interference spectra. In each of Figures 2 and 3, six CDFs formed from the same 7.1 minute collection are overlaid

to show the effect of changing the number of data points included in the CDF and to demonstrate stationarity of the results over that period of time. The raw data in these cases consisted of 2048 averaged spectra (revisits) each consisting of a spectral magnitude averaged over 256 FFTs (209 ms).

In one case, the CDF was formed from the first revisit alone (i.e., the first 209 ms of the 7.1 minute interval). The other five CDFs were obtained by downsampling the 2048 revisits to save 64, 128, 256, 512 or 1024 revisits (uniformly distributed in time over the 7.1 minute interval) and using the pooled data to construct the CDF. The CDFs based on 64 or more revisits are virtually indistinguishable, and the CDF based on a single revisit appears identical to within the random fluctuation expected from the small data set used in the histogram. This indicates stationarity of the CDF over periods of time of at least 7.1 minutes and indicates that 64 revisits are enough data to obtain stable results.

Before considering other experiment variations, we will consider the characteristics of the sets of CDF plots in Figures 2 and 3. These characteristics are similar among the other variations and are central to our conclusion that a lognormal model is applicable to HF interference. Observe that each CDF can be regarded as consisting of two distinct parts with a small transition region joining them. The upper part is the upper tail of the interference distribution, while the lower part is the lower tail of the background noise (received plus receiver generated). The composite distribution is well modelled in both cases by straight lines. To illustrate this, we have included a regression fit to the interference region of a typical CDF in each of the plots in Figures 2 and 3. Since the CDF plots use a normal probability scale, and the spectra data are all in logarithm format (i.e., in dB), each linear region can be regarded as approximately Gaussian. In the section following the description of the remaining experiment variations, we discuss rationale for the observed linearity for each region independently.

Figure 6 shows CDFs of data taken at 13 MHz that further demonstrate the stationarity of the CDF with time. Here, each of five CDFs was calculated using a different 128 revisit subset of the 2048 revisits collected in a 7.1 minute period. As before, each revisit consists of a 209 ms time average of the FFT bin output envelopes, with each 128 revisit subset spanning about 26.7 s. These subsets are uniformly spaced in time over the 7.1 minute total observation interval. Note that the CDFs are nearly identical except at the extreme upper tails of the distribution.

Another interesting result observed is that the CDF of the interference does not seem to depend on the envelope averaging time, at least for short averaging times. Figure 7 shows the results of measurements made at 15 MHz using three different values of envelope averaging time: no averaging (0.8 ms sample), 16 FFTs (13 ms), and 256 FFTs (209 ms). Each CDF was based on 128 revisits (of 0.8 ms, 13 ms, or 209 ms, as appropriate) equally spaced over a 7.1 minute total observation interval, and the three measurements were taken one after the other within a total period of about 30 minutes. The results are nearly identical in the interference region of the CDF, and the differences in the noise are as expected, since the distribution of an unaveraged magnitude spectrum will be Rayleigh (assuming white Gaussian input noise) whereas the averaged magnitude spectrum tends to Gaussian with increasing number of spectral samples averaged.

Figure 8 summarizes the results of measurements made over eight 2 MHz frequency bands on the evening of 28 December 1989. Each curve is the result of 128 revisits taken over a 7.1 minute period with 256 magnitudes averaged per revisit. The entire set of data was taken within a 90 minute period. Note that in almost all cases the interference portion of the CDF can be well approximated by a straight line except near the upper tail of the distribution. The noise portions of all of the CDFs also can be approximated by a straight line, indicating that the noise in each band was essentially white. No significance should be attached to the horizontal displacement of the curves since this was the result of operator adjustment of the input attenuator to optimize the dynamic range of the system at each frequency.

Figure 9 shows best and worst straight line fits to the data of Figure 8. These are the results of linear regression over data ranges chosen somewhat subjectively to emphasize the fit to a lognormal model. The best case (9 MHz) gave an excellent fit over a range of almost 60 dB. Even the worst case (19 MHz) gave an acceptable fit over a 25 dB range.

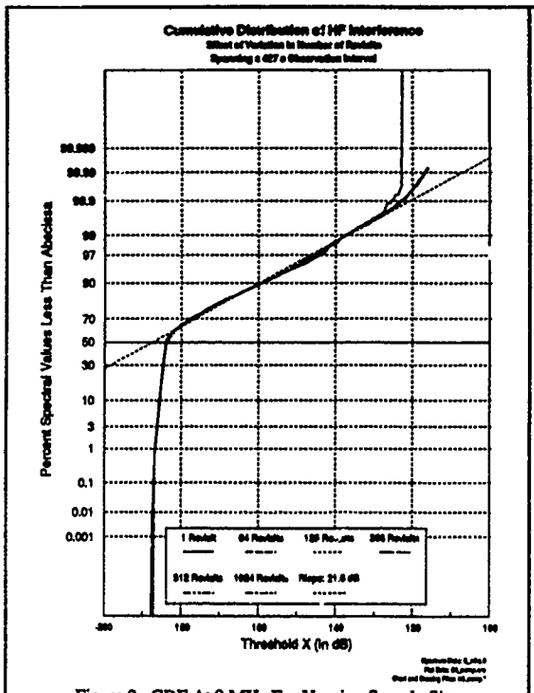


Figure 2. CDF At 9 MHz For Varying Sample Size (7.1 m Observation Interval)

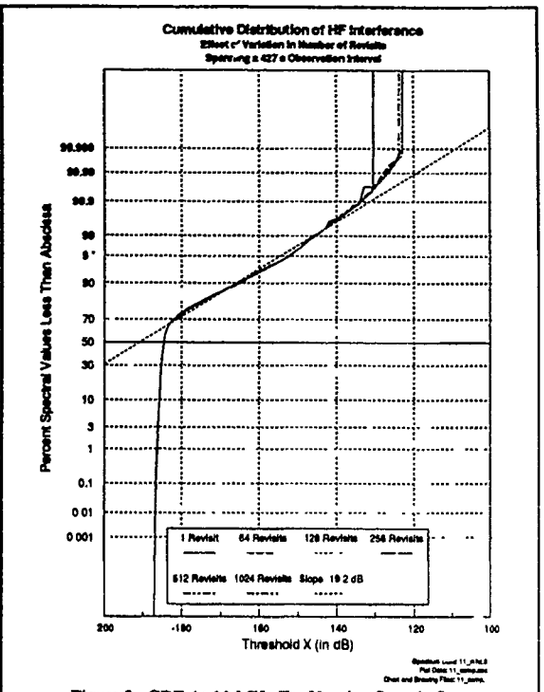


Figure 3. CDF At 11 MHz For Varying Sample Size (7.1 m Observation Interval)

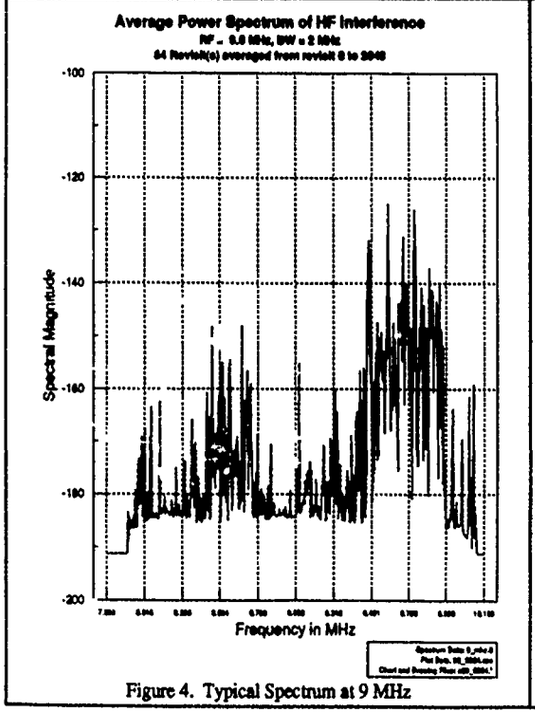


Figure 4. Typical Spectrum at 9 MHz

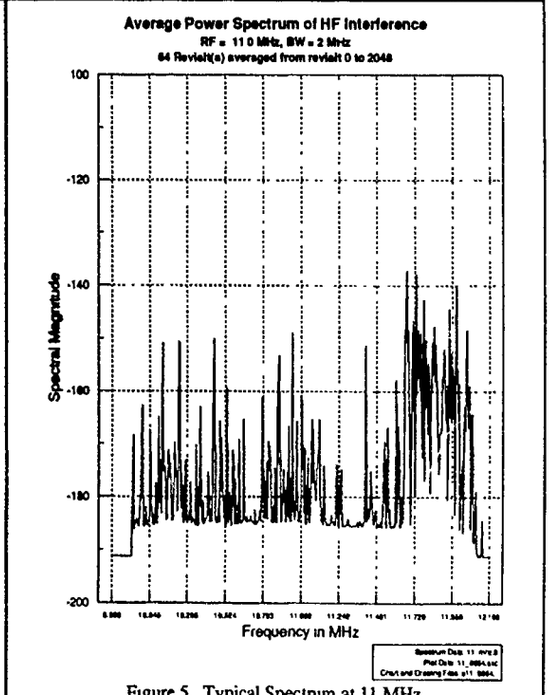


Figure 5. Typical Spectrum at 11 MHz

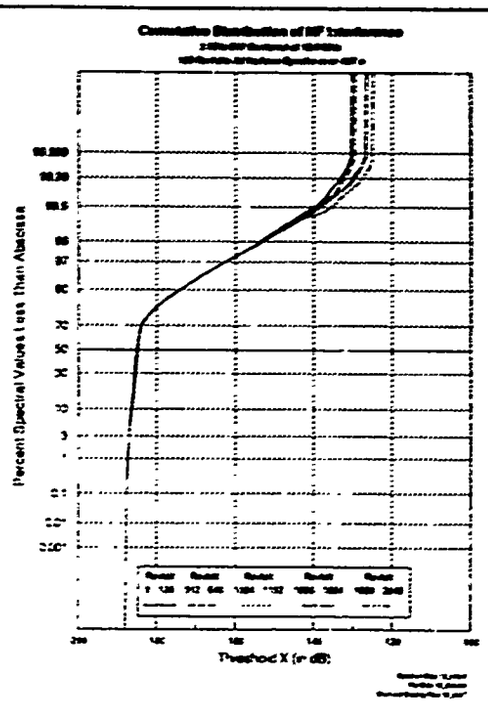


Figure 6. CDF at 13 MHz for Various 26.7 s Epochs Selected From 7.1 m

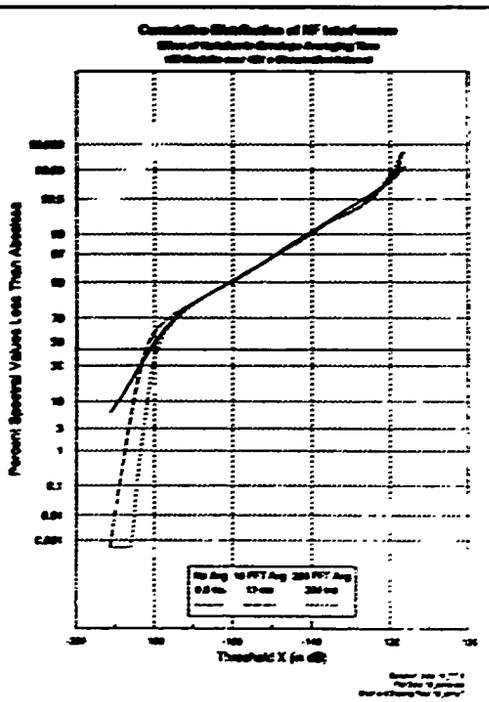


Figure 7. CDF at 15 MHz for Various Averaging Times Uniformly Distributed Over 7.1 m

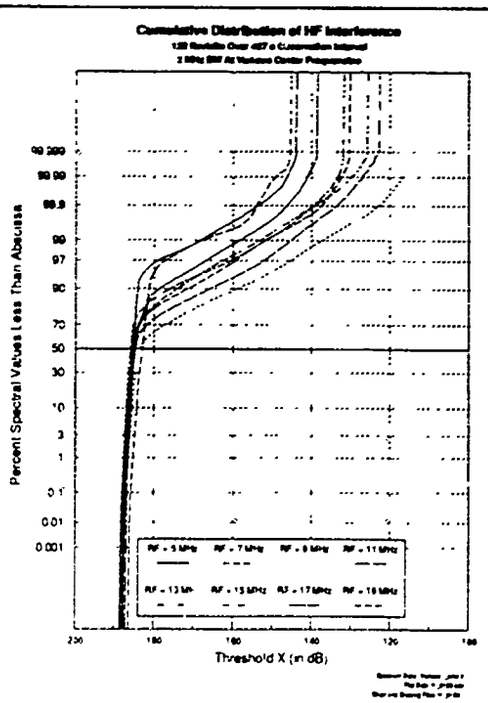


Figure 8. CDF For Various Center Frequencies

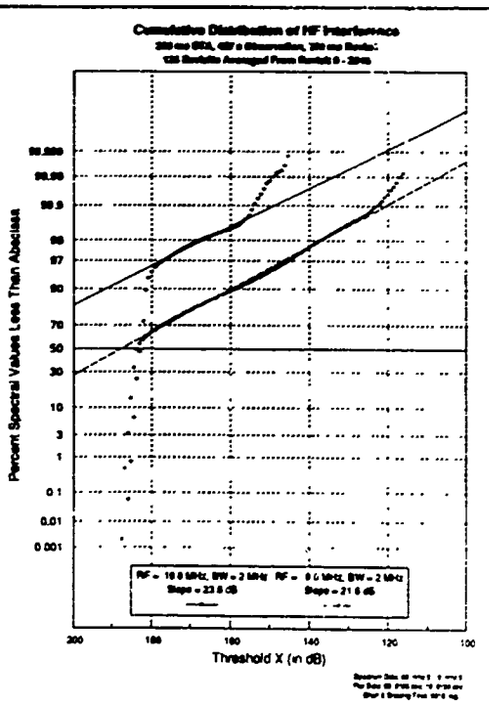


Figure 9. Regression Fit for CDF at 9 and 19 MHz

The slopes of the straight line fits are inversely proportional to the standard deviation S in decibels of the approximating Gaussian distribution. These standard deviations were found to be well clustered in both the April and the December data. The April data yielded an average value of 19.7 dB and a standard deviation of 2.4 dB for S. The December data yielded 20.1 dB and 1.8 dB for the mean and standard deviation of S. We have also estimated the slope of CDF data presented by Gibson and Arnett from the plots in Reference 1 to be approximately 20 dB. These results suggest that a Gaussian distribution of signal level in decibels with a 20 dB standard deviation would constitute a reasonable empirical model for the distribution of HF interference.

To complete the model, we need to specify the vertical axis intercept or other equivalent parameter of the straight-line fit. One convenient parameter is the probability value at which the extrapolated noise distribution intersects the straight-line approximation to the interference distribution. This point represents the fraction of all frequency bins that are dominated by noise rather than interference, and is an intuitively meaningful parameter. This parameter varied over a wide range in our experiments, as might be expected, since both the received signal level and the received noise vary significantly with frequency and time of day and year. Also the noise level is strongly dependent on the local noise environment, which in our case is a suburban industrial area.

ORIGIN OF THE LOGNORMAL DISTRIBUTION OF INTERFERENCE LEVELS

The observed fit of the interference CDFs to a lognormal distribution is too consistent over time and frequency to be chance and clearly must be the result of some robust mechanism.

Consider first the background noise region. In this case, the linearity of the log spectral data is indicative of a lognormal process; however, since the slope is very steep, it is also indicative of a distribution wherein the standard deviation is small relative to the mean. It is our supposition that the background noise spectrum is approximately Gaussian. To the extent that the noise in the 2 MHz band is white and suitably well behaved, the frequency bins of the average of a large number of magnitude spectra (i.e., 256 spectra) will be approximately jointly Gaussianly distributed by virtue of the central limit theorem. Furthermore, since the normalized standard deviation of the averaged bins will be very small, the distribution in dB also will be approximately lognormal and will plot as a straight line on a normal probability scale.

To show this, we consider a Gaussian variate x assumed to be normalized to unit mean. We write it as

$$x = 1 + u$$

where u is a zero-mean Gaussian random variable with standard deviation very much smaller than unity. The logarithm of x can be expressed in a Taylor's series as

$$\ln(x) = \ln(1+u) = u - \frac{u^2}{2} + \ldots$$

Since u is almost always much smaller than unity, all but the leading term of the series can be ignored, leading to the approximation $\ln(x) = u$. Therefore, since u is Gaussian, $\ln(x)$ will be approximately Gaussian.

Consider now the CDF region attributed to the interference levels. An obvious candidate here is the central limit theorem applied to various multiplicative random effects (additive in dBs) that determine the received interference level. There are more of these than might be supposed on first consideration and include:

- Rated transmitter output power
- Loading factor - fraction of transmitter power capability actually used, as determined by type of modulation and modulation source

- Transmit antenna gain in direction of receiver
- Average propagation loss from transmitter to receiver - product of inverse-square law loss and reflection losses at the ionosphere and ground
- Instantaneous fading gain for the path
- Receiver antenna gain in direction of the transmitter
- Fraction of interference power intercepted by a given frequency bin - depends on bin width and modulation type

While these are not all completely independent of each other, clearly there are several essentially independent multiplicative variables involved, each of which has a range of variation of several dB. Thus it would not be surprising that the convolution of their individual distributions (in units of dBs) would tend to a Gaussian distribution.

As a test, we postulated distribution functions for each of the above quantities expressed in decibels, and numerically convolved them to obtain a composite distribution for the received interference. The postulated distributions, described via the parameters in Table 1, are not offered as necessarily realistic, and in some cases are greatly oversimplified. However, they will serve to illustrate the principle. Figure 10 shows the resulting distribution plotted on a normal probability scale. It is seen to be reasonably linear over 120 dB of range. The standard deviation is about 15 dB, somewhat smaller than the experimental data, which suggests that some of the postulated distributions of Table 1 may have been too narrow. Note that this convolved distribution does not include a component representative of the broadband background noise so that a second steeply sloped region is not present.

Table 1
Postulated Component Distributions Contributing to Received Interference Distribution

Rated transmitter output peak envelope power: Six discrete values with following probabilities:			
Power	Probability	Power	Probability
10 W	0.05	1 kW	0.3
100 W	0.1	5 kW	0.25
500 W	0.2	10 kW	0.1

Loading factor: Four discrete values with following probabilities:		
Load Factor	Probability	Signal Type
1 0	0.3	FSK
0.5	0.5	CW Morse, AM Voice
0.1	0.1	VFT
0.05	0.1	SSB Voice

Transmit antenna gain: Cosine directivity pattern (short horizontal dipole)
Average propagation loss: Two factors assumed. One was inverse-square loss corresponding to transmitters uniformly distributed over earth surface. The other was ionospheric and ground reflection loss modelled as uniformly distributed from 0 to 6 dB.
Instantaneous fading gain: Rayleigh envelope distribution
Receiver antenna gain: Omnidirectional
Fraction of interference power in a single bin: Uniform distribution in dB ranging from 0 to 10 dB.

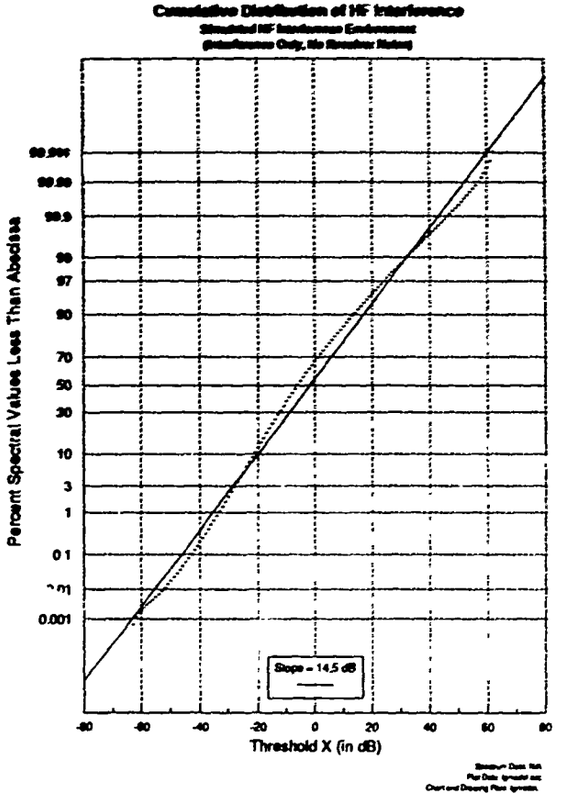


Figure 10. CDF of Simulated HF Model

CONCLUSIONS

Based upon both experimental data collected in the metropolitan Boston area and a theoretical distribution obtained by convolving the distribution of some hypothetical emitter characteristics, we have concluded that the typical HF interference levels can be well modeled as arising from a lognormal random process. Furthermore we have observed that the model is surprisingly robust with respect to several observation parameters including the envelope averaging time used to estimate the instantaneous interference level, the time epoch of the observation (at least over the space of several minutes), and the frequency of the particular 2 MHz wide observation spectrum. Our proposed model is consistent with the observations and conclusions by Gibson and Arnett drawn from HF spectrum occupancy measurements in Southern England and reported in [1]. This suggests that the model is not somehow tied into the HF environment conditions specific to the New England area of the U.S., and may therefore be globally applicable. An advantage of using the proposed lognormal model is that the combination of the background noise and interference characteristics can be specified with three parameters: the mean and variance of the lognormal process representing the interference and the mean power level of a broadband process representing the background and receiver noise.

1. A. J. Gibson and L. Arnett, "New HF Spectrum Occupancy Measurements in Southern England", Fourth International Conference on HF Radio Systems and Tecnniques, 1988.

2. B. D. Perry and L. G. Abraham, "A Wideband HF Interference and Noise Model Based on Measured Data", Fourth International Conference on HF Radio Systems and Techniques, 1988.

91 ... 50

91-09718

AD-P006 285

91-09717

HF CHANNEL OCCUPANCY AND SUBBAND CONGESTION: THE OTHER-USER INTERFERENCE PROBLEM

Roy H. Stehle
Information and Telecommunications Sciences Center
SRI International
333 Ravenswood Avenue
Menlo Park, CA 94025

George H. Hagn
Information and Telecommunications Sciences Center
SRI International
1611 N. Kent Street
Arlington, VA 22009

ABSTRACT

Information on other-user interference is important in predicting the performance of high-frequency (HF) systems. For many systems, especially frequency-hopping systems, other-user interference frequently is more important than man-made noise from incidental radiators (e.g., power lines) or atmospheric noise from lightning. Occupancy of individual channels and congestion of allocated HF subbands (e.g., fixed and mobile, amateur) vary with type of service, frequency, time of day, season, angle of arrival, type of receiving antenna, bandwidth, threshold, geographic location, and sunspot number. This paper summarizes channel occupancy and subband congestion observations made in the continental United States (CONUS) and Europe in 1987 with SRI International's Mobile Spectrum Monitoring Unit (MSMU). The data on subband congestion are compared with the empirical model developed by researchers at the University of Manchester Institute for Science and Technology (UMIST), Manchester, United Kingdom (UK). An approach for extending the UMIST model for congestion at a point to a model for congestion within a region is described.

INTRODUCTION

This paper describes high-frequency (HF) occupancy and subband congestion data taken by SRI International (SRI) and the University of Manchester Institute of Science and Technology (UMIST). It also describes how well a mathematical model of subband congestion developed at UMIST fits the measured data and proposes an extension from the UMIST model for a single location to a model for a region.

SRI MSMU AND RELATED DATA

A Mobile Spectrum Monitoring Unit (MSMU) constructed by SRI International was used at four locations in the continental United States (CONUS) and at two locations in the Federal Republic of Germany (FRG) in 1987 to measure HF signal levels, channel occupancy, and allocated subband congestion. The MSMU and the signal level and occupancy results for the shortwave broadcast bands have been described by Hagn et al. (1). The key features of the MSMU are the antenna, the bandwidth, the sample dwell time, and the channel revisit rate. The MSMU features an unmatched 16-ft vertical monopole mounted on the roof of the measurement van. The groundwave antenna factor for this

91

antenna was measured, and the results are given in Figure 1. The groundwave antenna factor data in Figure 1 (solid line) differ from the data in Figure 4 of Reference 1 (shown as a heavy broken line for comparison), with the data given here representing a refined calibration.

The MSMU made most of its measurements with a 3-kHz-bandwidth IF bandpass filter with excellent selectivity characteristics; Figure 2 is a plot of the passband characteristic of this filter under the dynamic conditions of fast scanning. The MSMU also had IF bandpass filters with bandwidths of 10, 20, and 60 kHz, which were used to obtain data on the bandwidth sensitivity of the mean signal level. The MSMU receiver is a triple-conversion, superheterodyne design utilizing a frequency synthesizer for channel tuning with random frequency selection capability. The log-compressed signal of the final IF section was averaged and digitally sampled. The MSMU receivers typically exhibited a sensitivity of -20 dBuV (-137 dBm) for a 3-dB signal-plus-noise-to-noise ratio when using the 3-kHz bandwidth. The 1-dB compression point was typically 86 dB above the sensitivity level. The two-signal, third-order intermodulation (IM) intercept was at least +102 dBuV (-5 dBm).

The system used a dwell time of 2 ms for each data sample, and each of the 9,333 channels in the band 2-30 MHz was revisited 150 times per hour. The scans were organized into subbands according to the international frequency tables for Europe (ITU Region 1) and North America (ITU Region 2) for data taken in the FRG and CONUS, respectively. Figure 3 shows the difference in subband mean signal level for samples taken with the 3- and 60-kHz-bandwidth filters; better matching to the ITU allocations was possible with the narrower bandwidth filter. Since the signals in the broadcast service subbands were typically 20 dB stronger in Europe than in CONUS during the day, and 5 dB stronger at night, a 20-dB attenuator was added to the system for the European measurements to mitigate intermodulation problems. The attenuation required was determined through field measurements.

UMIST SYSTEM AND RELATED DATA

The primary occupancy data for the HF band published to date have been those of Professor Geoffrey Gott et al. of the University of Manchester Institute for Science and Technology (UMIST) (2-6). Both the Gott and SRI systems have excellent filters to help avoid measuring signals in adjacent channels. However, the SRI system samples much more rapidly; that is, the SRI system has a 2-ms dwell time (vs. 1 s for Gott's system) and uses a 3-kHz bandwidth (vs. 1 kHz for Gott's). As a result, Gott's system revisits the same 1-kHz channel during the same hour (± 1 h) once every 3 to 5 days, depending on the sampling plan, for a 1-s duration; the SRI system revisits the same 3-kHz channel 150 times in a 5-minute period each hour. As a result, the SRI data are useful for channel and subband occupancy, and the Gott data are useful only for subband occupancy, which Gott has called congestion. Channel occupancy can be defined as the percentage of samples in an individual channel that exceed a given threshold (7). Gott has defined congestion as the percentage of subband channels with an occupancy exceeding a given threshold. SRI's congestion data represent the percentage of individual channel samples, accumulated over a subband, that exceed a given threshold.

The UMIST measurements have been made twice a year since 1982, around the winter and summer solstices at a site near Pershore, UK. During this time, the Belgian sunspot number ranged from 2 to 130. A 1-m active vertical monopole with good intermodulation characteristics was used. A groundwave field strength of 2 μ/m produced a 50-Ω receiver reading of 1 μV (-107 dBm). Additional data have been presented recently by Gibson et al. of the Rutherford Appleton Laboratory (RAL) (8, 9). A comparison of characteristics of the SRI, UMIST, and RAL systems is given in Table 1 of Reference 1. In 1989, UMIST activated a new site for occupancy data acquisition at Baldock, UK. The new site is remotely operated from Manchester, where the data processing and analysis facility is located. In 1990, a Swedish system, patterned on the UK system, began operation at Linköping.

UMIST SUBBAND CONGESTION MODEL

Laycock et al. (10) have used Gott's data to develop a mathematical model for HF subband congestion, Q, that is a function of subband (type of service and frequency), time of day, season, sunspot number, and threshold for the bandwidth of interest. For each set of measurements since January 1982, Laycock et al. divided the spectrum into 95 frequency allocations; congestion results were obtained for each allocation for five threshold levels. Hence, the total number of recorded values of congestion from 1982 up to the summer of 1986 equals 9500 (corresponding to about 3 million channel observations), at a measurement bandwidth of 1 kHz. They sought to represent the experimental data by a mathematical model in which values of congestion, Q, are given by

$$Q = f(x1, x2....) \quad , \tag{1}$$

where x1, x2..... represent parameters on which occupancy may be expected to depend.

Consider the following choice of a linear function of the variables,

$$Q = \sum_i A_i X_i \quad . \tag{2}$$

where X_i represents functions of the parameters and A_i represents coefficients. In the model, the measured values of congestion are proportions, and consequently their values must lie in the range 0 to 1. Since a significant fraction of the measured values lie on, or close to, one or another of these boundary values, there is a risk that a linear model, as defined above, may give estimated values of congestion that lie outside this range.

Laycock et al. used a logit transformation to overcome this risk, where

$$\text{logit}(Q) = \ln\left\{\frac{Q}{1-Q}\right\} - \sum_i A_i X_i = y \quad . \tag{3}$$

$$\text{Hence,} \quad Q = \frac{e^y}{1+e^y} \quad \text{where } 0<Q<1 \text{ and } -\infty < y < \infty \quad . \tag{4}$$

Thus, the linear function of the variables has been included as the index of an exponent, and

$$y = A_k + B * \text{threshold (dBm)} + (C_0 + C_1 f_k + C_2 f_k^2) * \text{sunspot number} \quad , \tag{5}$$

where A_k has 95 values, corresponding to the 95 frequency allocations from 1.606 to 30.000 MHz. B is a single coefficient to be multiplied by the threshold; C_0, C_1, C_2 are coefficients in a quadratic expression for frequency, to be multiplied by the sunspot number; and f_k is the arithmetic mean center frequency of the k-th allocation. The magnitude of adjustment to the subband congestion, based on sunspot number, is shown in Figure 4. Laycock et al. (10) provide a coefficient set for a summer day model (local noon, ± 1 h) for 95 allocations. The accuracy of the model is limited to the range -107 dBm to -77 dBm at the output of the active antenna. For the corresponding 1900 measured values of congestion, 56% are given by the model to an accuracy of 0.01 (1%), 91% are given to an accuracy of 0.05, and 98% are given to an accuracy of 0.1.

COMPARISON OF UMIST MODEL AND DATA WITH SRI DATA

Baumann (11) has suggested that the SRI and UMIST data be compared on the basis of mean measured power (in dBW) by converting SRI's data to Gott's calibration using SRI's groundwave antenna factor as given in Reference 1. Assuming that the directivity patterns of the two vertical

monopoles are similar, the conversion applied is as follows:

$$\text{UMIST (dBm)} = \text{SRI (dB}\mu\text{V)} - 107 + \text{BWF} - \text{AF}_u + \text{AF}_s \quad (6)$$

where: BWF = bandwidth factor (-4.77 dB for 3 kHz)
AF_u = UMIST groundwave antenna factor (6 dB)
AF_s = SRI groundwave antenna factor (from Figure 1) .

The UMIST model for a summer day given in Reference 10 (the only model available to SRI) was compared with the SRI MSMU data taken at Geinsheim, FRG (near Mannheim). A more appropriate comparison would have been a winter day, since the SRI data were taken in November 1987. The upper portions of Figures 5 through 8 show the predicted and measured congestions for each of the subbands; the lower portion of the figure shows the numerical difference between the predicted and measured values. Moderately good agreement can be seen, with the difference between the model predictions and the measured data increasing as the signal strength threshold is lowered. Also, as the signal level decreases, the effects of SRI's antenna factor and the system noise floor cause large variations below approximately 5 MHz. As expected, the fit is not as good as with the data used to develop the model.

The model was also compared with data measured by Gott (13) at Pershore, UK, in November 1987. Here one would expect the seasonal match to be better, but the geographical differences may make the fit worse. Figures 9 through 12 present the data in a fashion similar to the previous comparison. The figures show that significantly stronger signals were observed at Pershore, UK than at Geinsheim, FRG. Particularly noticeable is the difference in the broadcast subbands. Some of the large differences between the two sets of data may be attributed to uncertainties in antenna patterns and bandwidth and to video averaging differences, but the biggest difference is probably attributable to geographic siting. More significant, however, is the fact that the measured November data are significantly larger than those predicted for a summer day.

The probability of geographic effects attributing to the large differences in received signal level is based on the fact that Gott (12) and Laycock et al. (10) have reported good agreement with the congestion data acquired at Pershore and Baldock, UK, which are located approximately 140 km apart. Approximately 770 km distant, at an east-south-easterly bearing, are the sites of Seckenheim and Geinsheim, FRG; approximately 40 km separates these two sites. Figures 13 and 14 document the diurnal variation of the mean signal level measured in the fixed service subband from 9.040 to 9.500 MHz. Except for an occurrence of a strong emitter at Geinsheim, between the hours of 0300 and 0700 GMT, the curves exhibit great similarity.

Similar measurement data for four sites in the continental United States are presented in Figures 15 through 18. Generally, the mean signal levels are 10 dB lower than those observed in Europe, which is consistent with a much heavier usage of HF transmissions in Europe than in the United States. The general character of the US curves deviates markedly from those for Europe, assumably because of the diurnal propagation effects from the sources of the emitters and the differing usage of the spectrum in the vicinity.

Further comparative data may be found in Figures 19 through 36, which show the UMIST summertime day predicted congestion, the UMIST measured congestion in Pershore, UK, and the SRI measured congestion in Geinsheim, FRG. All data are for the 1-h timeframe around local noon. Cases of both good and poor agreement are observable. For this example set of measured congestion data, it appears that the broadcast and fixed services show the most agreement between measurements and predictions, while the aeromobile and maritime services exhibit much more variability. The scheduled nature and high occupancy of the broadcast service and the non-mobile characteristic of the fixed service may account for the better agreement in these subbands.

CONCLUSIONS

Knowledge of channel occupancy and subband congestion is important to HF system developers and to spectrum usage planners. Other–user interference is becoming a significant factor in determining the signal levels encountered in the HF spectrum. We believe that it would be useful to develop regional models of HF subband congestion as an aid in this planning process.

The UMIST summertime, local noon model has been demonstrated to give good agreement to the geographic region and temporal range applicable to the available data; it does not exhibit such agreement when applied to data acquired in in another season. The SRI congestion measurements just as often agree with the UMIST summertime model predictions as they do with the UMIST measurements taken in the same winter month. The large geographic distance between these sites is assumed to be the determining factor for this disagreement, although some differences might be attributable to the difference in the measurement equipment. A distance correlation test is needed to determine the maximum permissible spatial separation between monitoring sites. The comparisons between SRI and UMIST data suggest that this distance exceeds 140 km. Ionospheric propagation studies suggest that the minimum spacing, for correlation, might be 500 km or less. The effects of the daylight terminator may imply different values in an east–west direction from those in a north–south direction.

A regional network of monitoring stations should be established to collect the large volume of data that is required to develop a regional congestion model. Stations in the UK and in Sweden could be augmented with a site in the northern FRG to provide a preliminary continuous–monitoring network. The distances involved may be too great for spatial correlations, and additional sites should be considered.

Ideally, the data from each station should be taken with identical equipment and procedures, to remove equipment dependencies from the analysis. A low–cost equipment suite would allow more sites to become operational and contributory to the research. For specialized research, the higher sampling rate of the existing SRI MSMUs would be applicable to studying the diurnal aspects of channel occupancy and subband congestion. The two MSMUs, because of their mobility, could be used to derive spatial correlation data by synchronized monitoring of individual channels, since one MSMU would be moved to locations at increasing distances from the other (fixed) MSMU.

The UMIST model's algorithms appear to be a reasonable approach to developing a regional model with coefficients for each subband, type of service, sunspot number, season, and time of day. These coefficients could be interpolated to estimate a subband congestion value at an arbitrarily chosen regional location, time of day, and season.

ACKNOWLEDGMENTS

These data were collected under Contract DAAB07–86–D–A035. The work was performed for the Program Manager Regency Net, Col. James Fields, Mr. Ronald Durkel, and Mr. G. Allen Hart. The US Army CECOM technical monitor was Mr. Paul Major. The views expressed in this paper reflect those of the authors, who are solely responsible for any errors of commission or omission.

The authors wish to acknowledge the efforts of Mr. Larry O. Harnish, Mr. Richard D. Sarran, and Mr. Rick A. Page of SRI, who participated in the operation of the MSMU at the field sites during these measurements. We also appreciate the assistance of Professor Geoffrey Gott of UMIST, who provided us with his November 1987 data, enabling us to make a more complete comparison of measured congestion.

REFERENCES

1. G.H. Hagn, R.H. Stehle, and L.O Harnish, Shortwave Broadcasting Band Spectrum Occupancy and Signal Levels in the Continental United States and Western Europe, IEEE Trans. Broadcasting, Vol. 34, No. 2, pp. 115–125, June 1988.

2. S. Dutta and G.F. Gott, Correlation of HF Interference Spectra with Range, Proc. IEE, Part F, Vol. 128, No. 4, 1981.

3. G.H. Gott, N.F. Wong, and S. Dutta, Occupancy Measurements Across the Entire HF Spectrum, Proceedings of AGARD Conference on Propagation Aspects of Frequency Sharing, Interference, and System Diversity, Paris, 1982.

4. G.F Gott, S. Dutta, and P. Doany, Analysis of HF Interference with Application to Digital Communications, Proc. IEE, Part F, Vol. 130, No. 5, 1983.

5. N.F. Wong, G.F. Gott, and L.W. Barclay, Spectral Occupancy and Frequency Planning, Proc. IEE, Part F, Vol. 132, No. 7, 1985.

6. G.F. Gott and M.J.D. Staniforth, Characteristics of Interfering Signals in Aeronautical HF Voice Channels, Proc. IEE, Vol. 125, No. 11, pp. 1208–1212, 1978.

7. A.D. Spaulding and G.H. Hagn, On the Definition and Estimation of Spectrum Occupancy, IEEE Trans. EMC, Vol. EMC–19, No. 3, pp. 269–280, August 1977.

8. A.J. Gibson, P.A. Bradley, and J.C. Schlobohm, HF Spectrum Occupancy Measurements in Southern England, Proc. 3rd Int. Conf. on HF Communication Systems and Techniques, IEE Conf. Pub. 245, pp. 71–75, 1985.

9. A.J. Gibson and L. Arnett, New HF Spectrum Occupancy Measurements in Southern England, Proc. 4th Int. Conf. on HF Radio Systems and Techniques, IEE Conf. on HF Radio Systems and Techniques, IEE Conf. Pub. 284, Institution of Electrical Engineers, London, pp. 159–164, April 1988.

10. P.J. Laycock, M. Morrell, G.F. Gott, and A.R. Ray, A Model for HF Spectral Occupancy, Proc. 4th Int. Conf. on HF Radio Systems and Techniques, IEE Conf. Pub. 284, Institution of Elecltrical Engineers, London, pp. 165–171, April 1988.

11. E. Baumann, private communication, March 1989.

12. G. Gott, private communication, 27 April 1990.

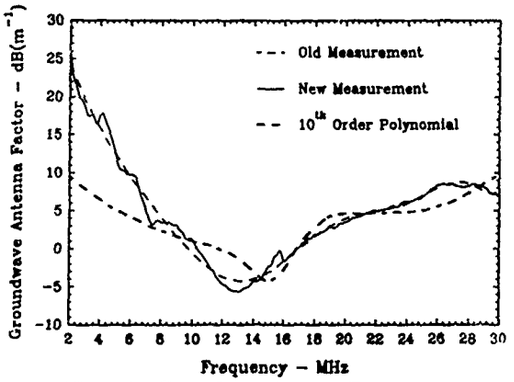


Figure 1. Groundwave Antenna Factor for SRI 15-foot Whip

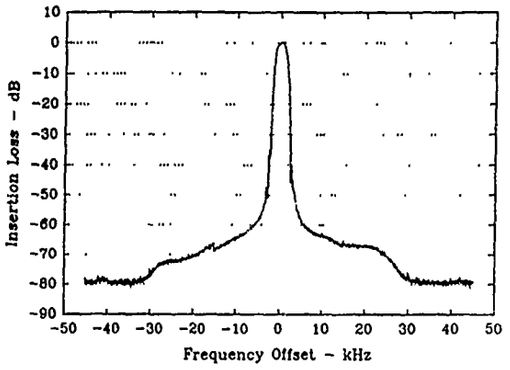


Figure 2. 3-kHz-Bandwidth IF Filter Frequency Response

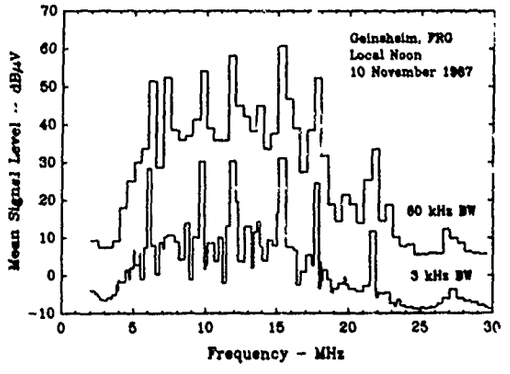


Figure 3. Subband Mean Signal Level as a Function of Bandwidth

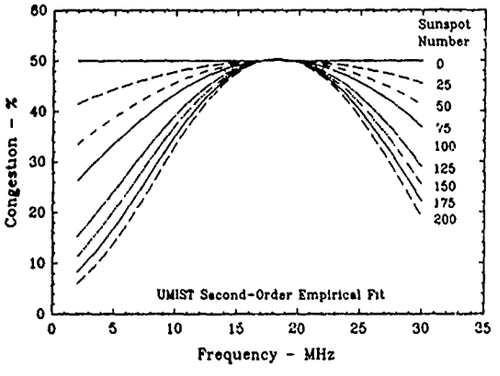


Figure 4. Influence of Sunspot Number on Predicted Congestion

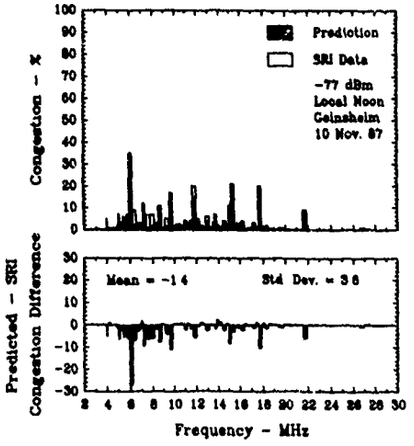


Figure 5. UMIST Congestion Predictions vs. SRI Data -77 dBm Threshold, Local Noon, Geinsheim, FRG

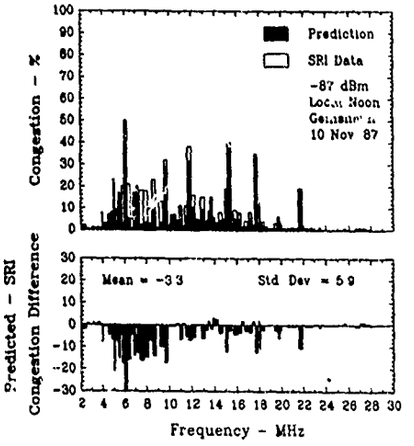


Figure 6. UMIST Congestion Predictions vs. SRI Data -87 dBm Threshold, Local Noon, Geinsheim, FRG

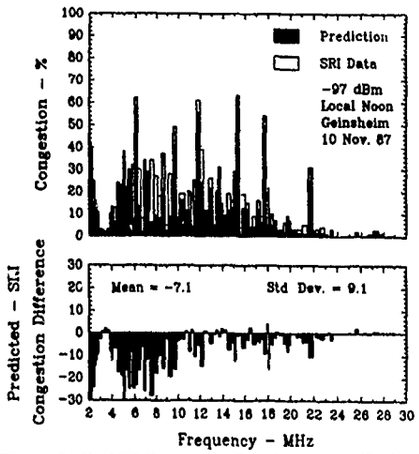


Figure 7. UMIST Congestion Predictions vs. SRI Data -97 dBm Threshold, Local Noon, Geinsheim, FRG

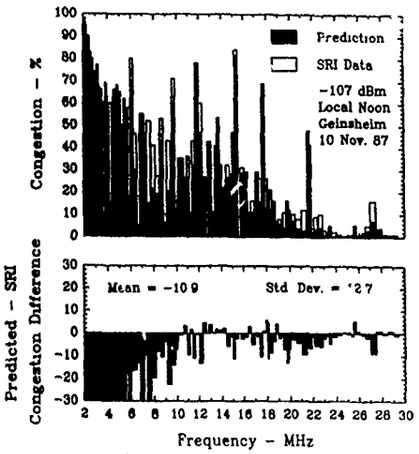


Figure 8. UMIST Congestion Predictions vs. SRI Data -107 dBm Threshold, Local Noon, Geinsheim, FRG

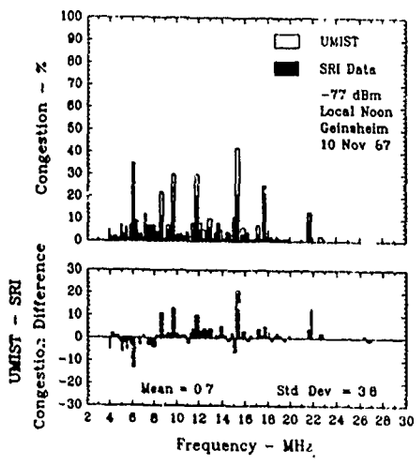


Figure 9. UMIST and SRI Congestion Data -77 dBm Threshold, Local Noon, Geinsheim, FRG

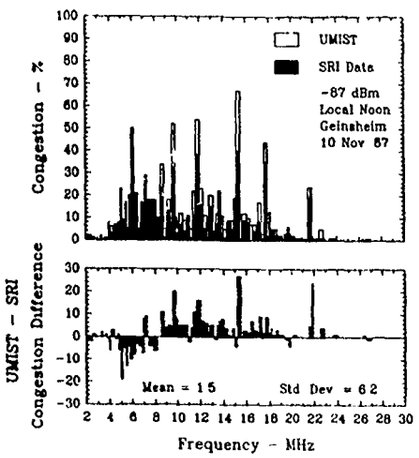


Figure 10. UMIST and SRI Congestion Data -87 dBm Threshold, Local Noon, Geinsheim, FRG

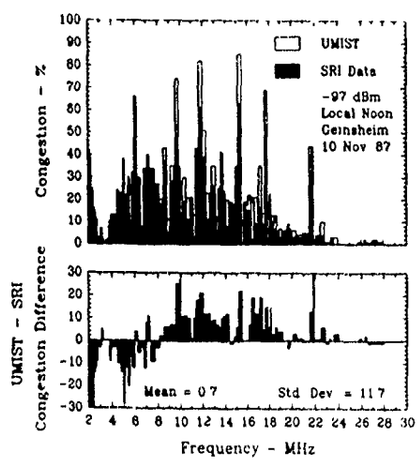


Figure 11. UMIST and SRI Congestion Data -97 dBm Threshold, Local Noon, Geinsheim, FRG

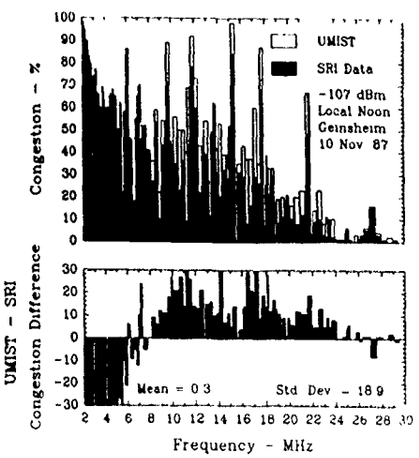


Figure 12. UMIST and SRI Congestion Data -107 dBm Threshold, Local Noon, Geinsheim, FRG

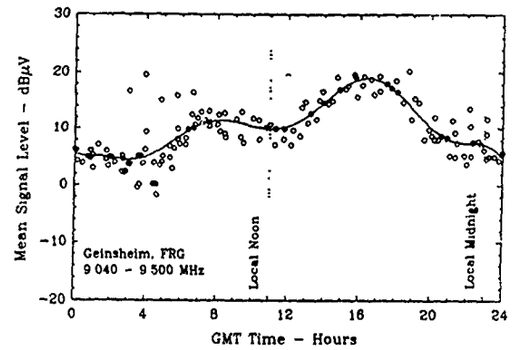


Figure 13. Mean Signal Distribution at Geinsheim, FRG Fixed Service Subband (9.040 to 9.500 MHz)

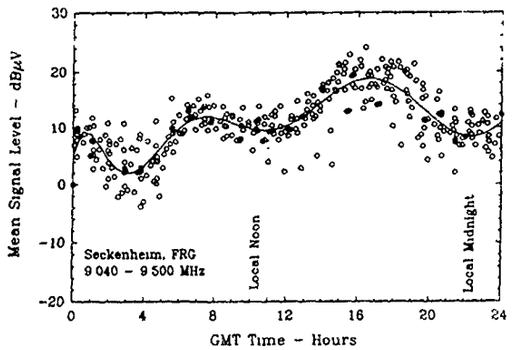


Figure 14. Mean Signal Distribution at Seckenheim FRG Fixed Service Subband (9.040 to 9.500 MHz)

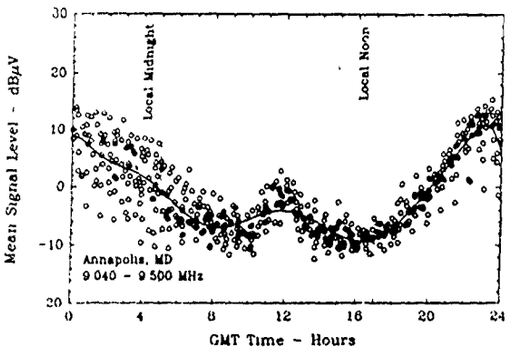


Figure 15. Mean Signal Distribution at Annapolis MD Fixed Service Subband (9.040 to 9.500 MHz)

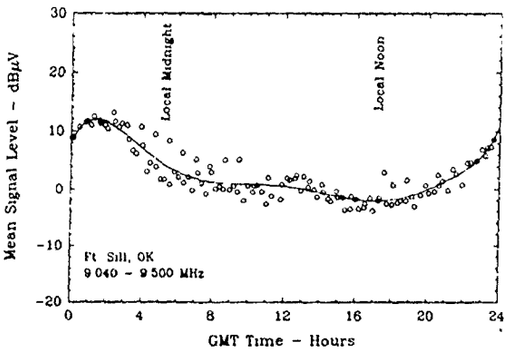


Figure 16. Mean Signal Distribution at Ft. Sill, OK Fixed Service Subband (9.040 to 9.500 MHz)

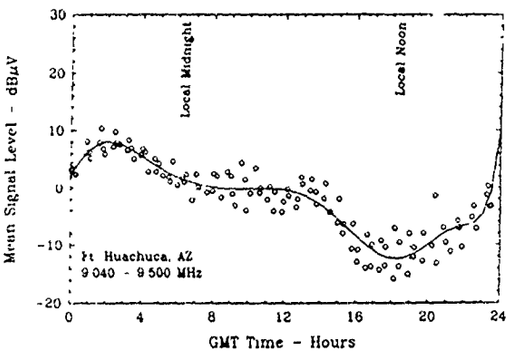


Figure 17. Mean Signal Distribution at Ft. Huachuca, AZ Fixed Service Subband (9.040 to 9.500 MHz)

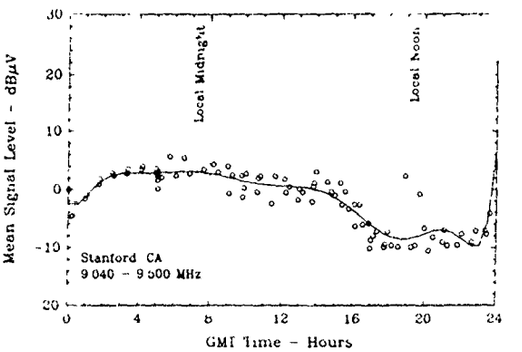


Figure 18. Mean Signal Distribution at Stanford, CA Fixed Service Subband (9.040 to 9.500 MHz)

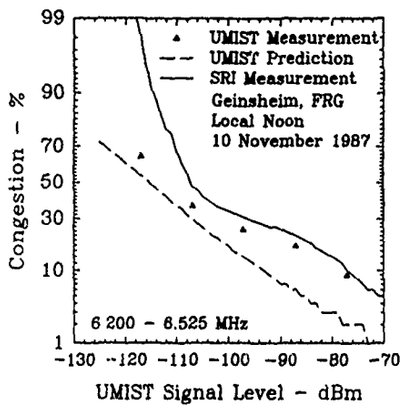


Figure 19. Measured and Predicted Congestion at Maritime Mobile Service Subband (6.200 to 6.525 MHz)

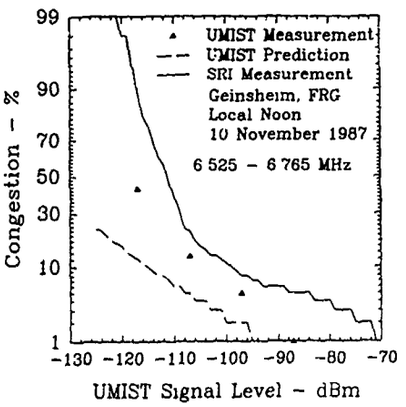


Figure 20. Measured and Predicted Congestion at Aeromobile Service Subband (6.525 to 6.765 MHz)

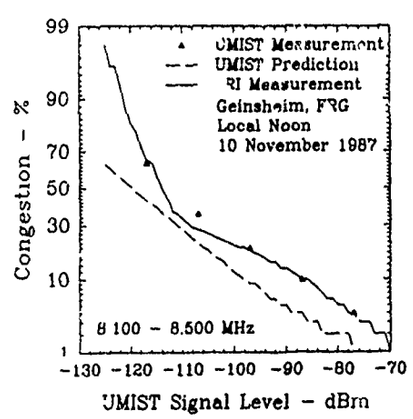


Figure 21. Measured and Predicted Congestion at Fixed and Maritime Mobile Service Subband (8.100 to 8.500 MHz)

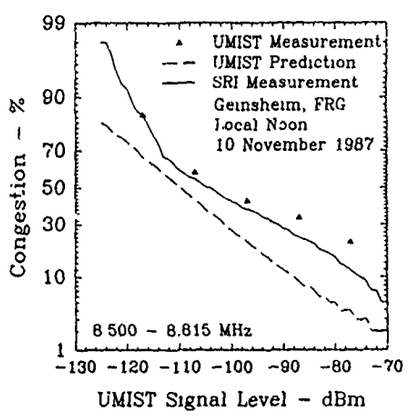


Figure 22. Measured and Predicted Congestion Maritime Mobile Service Subband (8.500 to 8.815 MHz,

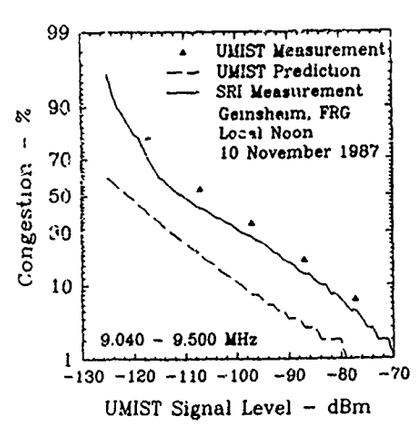


Figure 23. Measured and Predicted Congestion Fixed Service Subband (9.040 to 9.500 MHz)

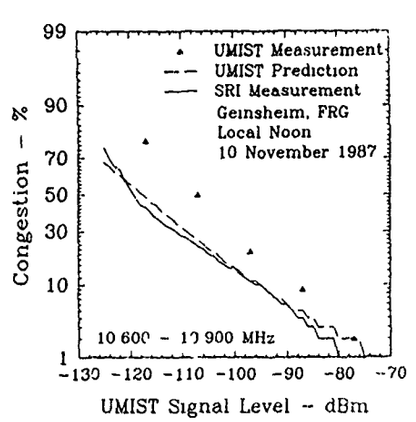


Figure 24. Measured and Predicted Congestion Fixed Service Subband (10.600 to 10.900 MHz)

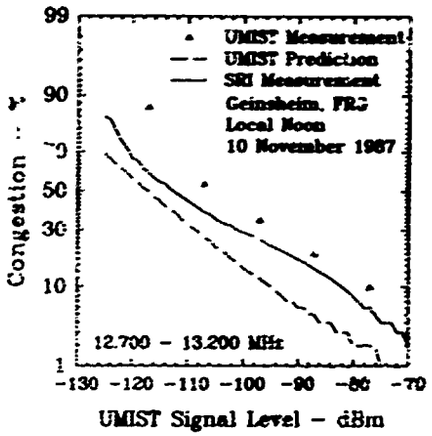


Figure 25. Measured and Predicted Congestion Maritime Mobile Service Subband (12.700 to 13.200 MHz)

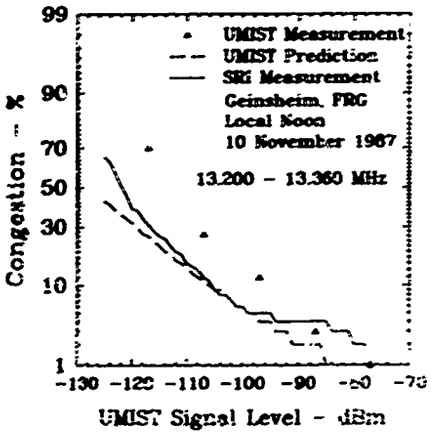


Figure 26. Measured and Predicted Congestion Aeronautic Service Subband (13.200 to 13.360 MHz)

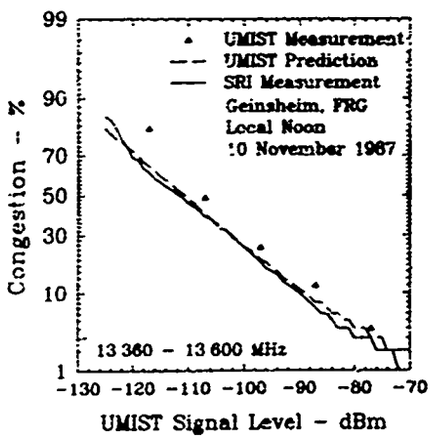


Figure 27. Measured and Predicted Congestion Fixed and Mobile Service Subband (13.360 to 13.600 MHz)

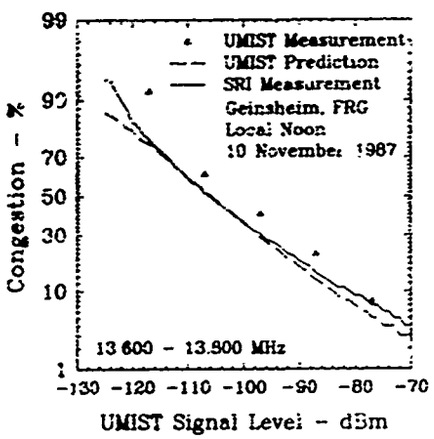


Figure 28. Measured and Predicted Congestion Broadcast Service Subband (13.600 to 13.800 MHz)

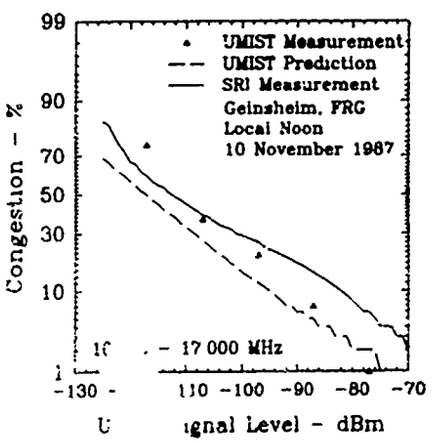


Figure 29. Measured and Predicted Congestion Maritime Mobile Service Subband (16.600 to 17.000 MHz)

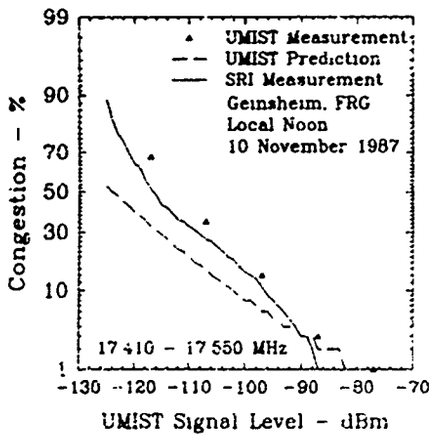


Figure 30. Measured and Predicted Congestion Fixed Service Subband (17.410 to 17.550 MHz)

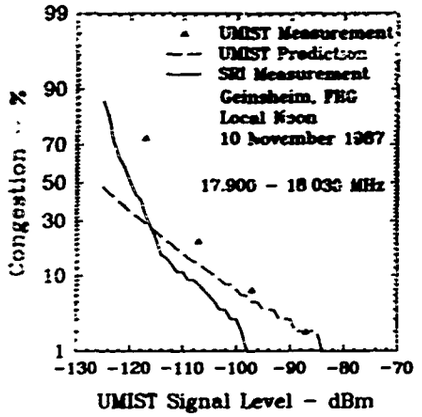


Figure 31. Measured and Predicted Congestion Aeromobile Service Subband (17.900 to 18.030 MHz)

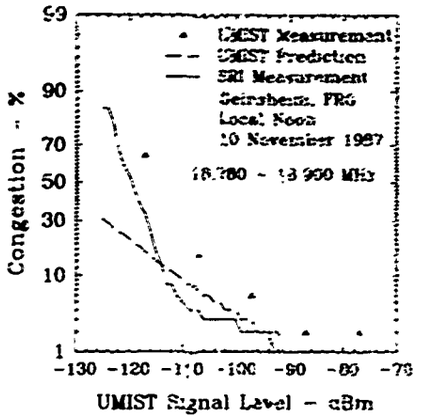


Figure 32. Measured and Predicted Congestion Maritime Mobile Service Subband (18.780 to 18.900 MHz)

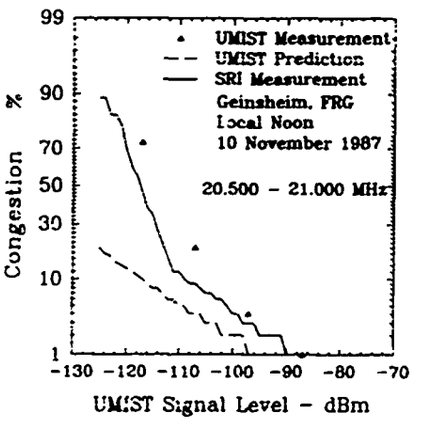


Figure 33. Measured and Predicted Congestion Fixed and Mobile Service Subband (20.500 to 21.000 MHz)

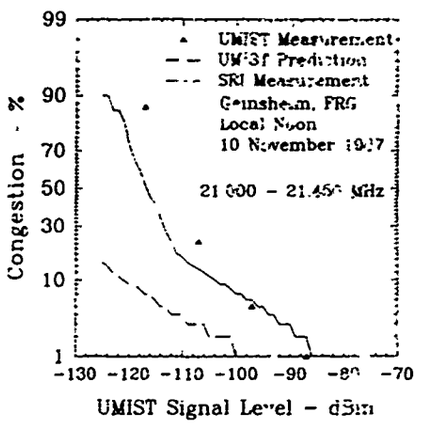


Figure 34. Measured and Predicted Congestion Amateur Service Subband (21.000 to 21.450 MHz)

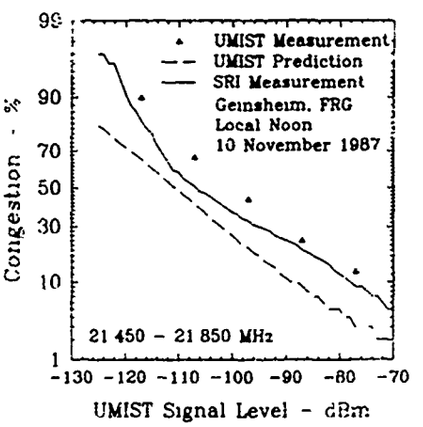


Figure 35. Measured and Predicted Congestion Broadcast Service Subband (21.450 to 21.850 MHz)

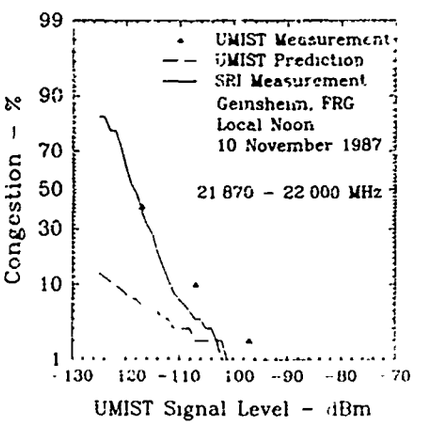


Figure 36. Measured and Predicted Congestion Aeromobile Service Subband (21.870 to 22.000 MHz)

AD-P006 286

91-09716

TRANSFER FUNCTIONS AND PULSE DISTORTION FOR AN IONOSPHERIC REFLECTION CHANNEL WITH EMBEDDED RANDOM IRREGULARITIES

Timothy W. Rand and K. C. Yeh

Department of Electrical and Computer Engineering, University of Illinois at Urbana-Champaign

ABSTRACT

The ionospheric reflection channel is known to be dispersive, time-varying, and random. The dispersive nature of the ionosphere and the propagation geometry make the transfer function frequency selective, i.e. frequency dependent; while the ionospheric motions cause the communication channel to be time dependent or fading. Frequently, the background ionosphere is permeated by random irregularities. A radio wave propagating through such an irregular ionosphere will have its phase randomized, followed by phase mixing through diffraction. As a result, the transfer function also acquires a random character. The transfer function for a random, dispersive, and time-varying ionospheric reflection channel can be simulated numerically using the phase-screen-diffraction-layer method. The detailed behavior of the calculated transfer function is examined in several different ways. First the spatial behavior of the transfer function, both magnitude and phase, is depicted in three-dimensional plots. This spatial behavior can also be interpreted as a temporal behavior for a moving ionosphere. A curious phase jump of 180° accompanied by a deep fade at some frequency within the pass band of the signal is found to occur. To facilitate the examination of this phase jump, locus plots are made. Using real and imaginary parts of the transfer function as coordinates, the transfer function traces out a locus with increasing frequency on this complex plane. In the presence of severe selective fading, the locus is shown to pass through the origin occasionally. When this happens, the magnitude of the transfer function dips down to zero with an accompanied phase jump of 180°. Such an event is caused by destructive interference among multipath components. Secondly, the channel behavior is examined by taking temporal moments of the received pulse. This is done for several sample pulses, each up to the moments of the fourth order, giving the arrival time, pulsewidth, skewness, and kurtosis. Each of these four quantities is then treated as a time series in the computation of the mean, standard deviation, histogram, power spectrum, and autocorrelation function. The stochastic behavior of these four quantities is thus thoroughly studied

INTRODUCTION

Ionospheric radio propagation has been used as a means of long-distance communication since the 1930s. Although the quiet background ionosphere can be represented by a deterministic model, random ionospheric turbulence and irregularities cause deviations from the background ionization levels which can severely affect radio waves propagating in the ionosphere.

The effect on waves after propagating through an assembly of random irregularities was initially treated using a diffraction theory. The importance of mixing an angular spectrum of waves through the diffraction process was first pointed out by several authors [Booker et al., 1950; Fejer, 1953]. An early review on the subject can be found in Ratcliffe [1956]. In these works, the irregular ionosphere is modeled as a phase-changing screen. A method of treating large phase changes known as a deep phase screen has also been developed [Mercier, 1962] and has been applied to the case when the irregularities have a power-law spectrum [Rino, 1979]. The relation of a simple phase screen model to the more exact treatment under the forward scatter approximation has been made more transparent by Dashen [1979] who expresses the solution in terms of Feynman's path integrals. The theory demonstrates the importance of diffraction in addition to phase perturbations within the irregular layer especially when the layer is thick, which is frequently the case at the equatorial ionosphere. Use of differential operators has been successful in mathematically justifying a procedure known as the split-step algorithm [DiNapoli and Deavenport, 1979]. In their application to underwater acoustics, they show that the diffraction operator and the phase perturbation operator can be split and applied sequentially if the whole irregular layer is divided into steps. In the meantime, the continued interest in ionospheric scintillation problems helps develop an implicit, numerically-stable method to solve the parabolic equation either for the moments up to the fourth order [Liu et al., 1974] or for the random field itself for each realization [Wernik et al., 1980]. The phenomenon of ionospheric scintillation, its historical development and the mathematical methods in treating the problem can be found in the review of Yeh and Liu [1982]. The scintillation effect on temporal signals can be found in Knepp [1983]. In all of above works, the problem treated pertains to the line-of-sight geometry, such as transionospheric propagation. For high frequencies, a radio signal can also be reflected. In this case, the radio wave experiences the reflection process in addition to diffraction and phase perturbation. The first attack on this problem was made by Kiang and Liu [1985a and b], which was later refined and extended to the oblique propagation [Wagen and Yeh, 1986]. In this paper the transfer functions computed by Wagen and Yeh [1989a, 1989b] are used to study the stochastic behavior of pulse propagation.

This paper is divided into three sections. The Ionospheric Modeling section describes the model ionospheres and ionospheric transfer functions used in this study. The section on Pulse Propagation and Temporal Moments describes the pulses used in the simulation, defines temporal moments, and includes figures showing pulse distortion. Observations and conclusions are presented in the Discussion section.

91

IONOSPHERIC MODELING

The ionospheric reflection channel is modeled as a horizontally-stratified medium so that the electron density is a function of altitude. Horizontal variations in the background ionosphere, the magneto-ionic effect, and the curvature of the earth are not considered. The model of the ionosphere includes a background electron density profile that is perturbed by ionospheric irregularities. Wagen [1988] produced transfer functions for several ionospheric models. Three sets of Wagen's transfer functions are investigated in detail in Rand and Yeh [1989]. The first case is the vertical-incidence case with a carrier frequency of 5 MHz. This case uses a linear ionospheric profile. The second case is an oblique-incidence case where the incidence angle is $\theta_0 = 60°$ and the carrier frequency is 6 MHz. The ionospheric profile is a piecewise-linear approximation of the International Reference Ionosphere—IRI 79 [Lincoln and Conkright, 1981]. This reference ionosphere is for Boston, Massachusetts, in March at midnight with a Zurich sunspot number of R = 100. For the propagation conditions given, the ground distance is 998 km. The third case is the same as the second, except the Zurich sunspot number for the third case is R = 10, the carrier frequency is 3.5 MHz, and the ground distance is 981 km. All presentations in this paper correspond to case 2. Readers interested in other cases or more complete presentations should consult the report [Rand and Yeh, 1989].

A quantity of interest in the ionospheric model [Wagen, 1988] is the fractional fluctuation in the electron density, σ_{Nf}. The fractional fluctuation in the electron density is the standard deviation of $\Delta N/N_0$, where ΔN is the fluctuation in the electron density and N_0 is the electron density in the background ionosphere. For the first case, the fractional fluctuation in the electron density is one tenth of one percent. For the second and third cases, σ_{Nf} is one half of one percent. In this report, the second case is chosen as a representative example.

Turbulence and plasma instabilities in the ionosphere cause irregular deviations from the background ion density. A wave incident on the ionosphere from free space at an incidence angle of θ_0 is described by the following wave function.

$$E(\vec{r}) = U(\vec{r})\, E_0(z)\, e^{-j\, k_0 \times \sin\theta_0} \tag{1}$$

In (1), $U(\vec{r})$ is the complex amplitude, $E_0(z)$ is the electric field due to the background ionosphere, and the exponential factor is the description of the horizontal phase shift according to Snell's law. The complex amplitude is the stochastic part of the expression. It models the irregularities that occur in the ionosphere. The field at height z, $E_0(z)$, satisfies the wave equation and is completely determined by the background ionospheric profile and the frequency.

The phase-screen-diffraction-layer method is used to compute the complex amplitude, $U(\vec{r})$, in order to model the effects of the turbulent ionosphere on high frequency waves [Wagen and Yeh, 1989a, 1989b]. The ionospheric irregularities cause phase fluctuations. These phase fluctuations and diffraction effects are computed sequentially along the ray path in the ionospheric reflection channel. In the region of the ray path near the turning point, the classical theory given by WKB solutions is not valid. A more careful analysis using the Airy functions is needed. The details of the model for the turning-point region appear in Wagen and Yeh [1989a].

For a particular point on the ground, the value of $U(\vec{r})$ is computed at the carrier frequency for incidence angle $\theta_0 = 60°$. Then values of $U(\vec{r})$ are found at the same point on the ground for ten other frequencies. Because the different frequencies are reflected at different heights, homing is necessary to find the correct launch angle so that the signal may be received at the same ground location. The set of eleven values of $U(\vec{r})$, one for each of eleven frequencies, makes up the transfer function for the ionospheric reflection channel at a particular location on the ground. Wagen calculates transfer functions for each of 8192 distance locations for each of the cases considered. The spatial information may be interpreted as temporal behavior for a moving ionosphere.

The complex amplitudes computed by Wagen [1988] are used to produce the received pulses presented in the next section. The complex amplitudes represent the effect of the ionospheric irregularities on the received pulses, but do not include the effects due to the background ionosphere.

Individually-realized transfer functions and their mean have been previously published [Wagen and Yeh, 1989b]. In computing the transfer function, the diffracted wave for a plane wave incident on the ionosphere is calculated on the ground, which is expressed as a function of distance. An ensemble of such transfer functions on the ground varying in distance from 0 through 3750 m over the frequency band from 5.90 MHz through 6.10 MHz is plotted in Fig. 1 under ionospheric and propagation conditions described as case 2 earlier. The slow change in both the magnitude and the phase of the transfer function with distance is very clear. If it is assumed that the ionosphere is undergoing a frozen flow in the horizontal direction of velocity v_0, the distance scale in Fig. 1 can be translated into a time scale by dividing it by v_0. In so doing, the transfer function can be visualized as its realization at successive discrete times. In examining Fig. 1, it is interesting to note the abrupt phase changes, such as those occurring at distances of about 0, 900 and 2900 m. Occurrence of such phase jumps is found to be accompanied by the vanishing of the transfer function. To investigate this phenomenon further, the transfer function is plotted in a complex plane as a locus plot shown in Fig. 2. As the frequency is varied across the band from 5.90 MHz to 6.10 MHz, the point corresponding to the transfer function traces out a locus on the complex plane. The plot of Fig. 2a is made for a distance of 750 m. As the distance (or the time after being translated for a frozen flow as explained above) is varied, the locus plot will change slowly. Occasionally. the locus can pass through the origin as shown in Fig. 2b for a distance of 14,475 m. In this example, owing to multipath, severe selective fading occurs at 5.99 MHz at which the amplitude fades to zero and near which the phase shows a jump of π rad. Since the frequency of 5.99 MHz is near the center of the signal bandwidth, severe fading at this frequency is expected to have a pronounced effect on distorting the signal. This will be presented in the next section.

The sharp phase discontinuity in the transfer function is different from the smooth phase changes that occur when the background ionosphere is modeled, for example, as a quadratic phase channel [Romanova and Ryzhkina, 1976]. The phase for a quadratic phase channel is expressed as

$$\phi(\omega) = \phi(\omega_c) + \phi'(\omega_c)\,(\omega - \omega_c) + \frac{1}{2}\,\phi''(\omega_c)(\omega - \omega_c)^2 \tag{2}$$

where the first three terms of the Taylor series expansion represent phase shift, time delay, and dispersive effects, respectively, and the higher-order terms are ignored. The transfer function of a quadratic phase channel, when plotted as a locus plot in the same manner as that in which Fig. 2 is plotted, is a circle since the transfer function maintains a constant magnitude. The total effect can then be treated as the product of this circle with the locus such as shown in Fig. 2.

PULSE PROPAGATION AND TEMPORAL MOMENTS

For each of the three cases for which Wagen developed transfer functions, two pulse shapes were investigated: Gaussian and triangular. The two pulses experienced similar distortion when passed through the reflection channel. The triangular pulse and the distortion it experienced are discussed in this section.

The idealized triangular pulse of half width 25 μs to the null is transformed to the frequency domain using Fast Fourier Transform. The resulting amplitude spectrum is then discretely sampled and band limited to a bandwidth of ±0.1 MHz about the carrier of 6 MHz. When this truncated discrete spectrum is transformed back to time domain, the triangular envelope remains reasonably intact with slight rounding of the triangular tip and the appearance of minor lobes about 40 dB or more below the main lobe. The half width of the first null is still about 25 μs and the pulsewidth defined by using the temporal moment (see later in this section) is 7.91 μs. This non-idealized triangular but bandwidth-limited pulse is then passed through the reflection channel represented by the typical transfer functions of Fig. 1. The resulting sample output pulse envelopes from the channel are shown in Fig. 3. As can be seen in Fig. 3a, there is considerable distortion: the triangular peak being higher than unity, the shift of the peak in time from the origin, the broadening of the pulse, and the rise and smearing of the minor lobes. All these behaviors can be quantified as done later by using temporal moments. One interesting phenomenon alluded to earlier is concerned with phase jumps accompanied by deep fading. This occurs at a distance of 14,475 m, for which the locus plot is depicted in Fig. 2b. Corresponding to this distance, the output pulse envelope is shown in Fig. 3b. Severe distortion and broadening are clearly manifested as expected.

Fig. 4 shows the envelopes of the received pulses versus time for successive distances. By viewing this figure, one can visualize how the received pulses change with distance. If the ionosphere is undergoing a frozen flow, one can alternately view the figure as successive pulses at discrete times.

To quantify the pulse distortion by a few numbers, the temporal moments are found to be useful. The n^{th} temporal moment is defined as

$$\text{(Temporal Moment)} \qquad M_n = \int I(t)\, t^n\, dt \tag{3}$$

where I(t) is the intensity of the pulse. The intensity is the square of the pulse envelope. The zeroth moment is computed first and I(t) is normalized so that

$$M_0 = \int I(t)\, dt = 1 \tag{4}$$

Once the intensity is normalized as in (4), the first moment is calculated. The first moment is the mean time of arrival, t_a.

$$\text{(Mean Time of Arrival)} \qquad M_1 = \int I(t)\, t\, dt = t_a \tag{5}$$

With the arrival time calculated by (5), the n^{th} central temporal moments are defined as

$$\text{(Central Temporal Moment)} \qquad \overline{M}_n = \int I(t)\, (t - t_a)^n\, dt \tag{6}$$

The zeroth central moment, $\overline{M}_0$, is the same as M_0 and is normalized to unity. The first central moment, $\overline{M}_1$, is equal to zero. Higher-order central moments yield pulsewidth, τ; skewness, s; and kurtosis, K. These quantities are defined as

$$\text{(Pulsewidth)} \qquad \tau = \sqrt{\overline{M}_2} \tag{7}$$

$$\text{(Skewness)} \qquad s = \frac{\overline{M}_3}{\tau^3} \tag{8}$$

$$\text{(Kurtosis)} \qquad K = \frac{\overline{M}_4}{\tau^4} - 3 \tag{9}$$

The computation of each of the quantities t_a, τ, s or K is weighted by pulse intensity which is a stochastic quantity. Accordingly, each of t_a, τ, s or K can be treated as a time series (assuming appropriate translation to time behavior for a frozen flow). These are presented in Fig. 5 through Fig. 8. Note that these plots are actually made up of many tiny dots. Each dot pertains to the quantity at a particular point. The two neighboring dots are separated by 15 m, the resolution of the computation. There are 8192 dots for each plot. The time of arrival as shown in Fig. 5 ranges from –14 to +12 μs. The mean is 1.24 μs and the standard deviation is 4.37 μs. These values are shown in Table I.

If a set of data is approximately Gaussian and if the mean and standard deviation are known, then a confidence interval within which the true mean probably lies can be calculated. The Rule of Three Sigma [Shreider, 1966] is given by

$$\text{(Rule of Three Sigma)} \qquad \delta \le \frac{3\sigma}{\sqrt{\eta}} \tag{10}$$

where δ is the one-sided confidence width, σ is the standard deviation of the data, and η is the number of independent sample points in the data set. If the Rule of Three Sigma is applied, the 99.7-percent confidence interval of the mean is given by $\mu \pm \delta$, where μ is the mean. Statistically, there is a 99.7 percent certainty that the mean lies within δ of the mean calculated for a particular data set. Note that η is the number of independent sample points, a number different from N, the total number of points. The value of η can be estimated by dividing N by the number of sample points in the correlation distance. Table I shows the 99.7-percent confidence intervals for the mean values calculated for each moment.

Table I. Mean, μ, and standard deviation, σ, for each moment.

Moment	Mean, μ	Standard Deviation, σ
Mean Time of Arrival (μs)	1.24 ± 1.19	4.37
Pulsewidth (μs)	13.5 ± 1.18	3.72
Skewness	-0.143 ± 0.113	0.447
Kurtosis	3.96 ± 0.612	2.49

As shown in Table I, the effect of irregularities is to retard the pulse arrival time by 1.24 μs. This value is small when compared with arrival time for the background ionosphere (a few ms for the case under study). Furthermore, a frequency histogram can be constructed and fitted by a Gaussian curve with this mean and standard deviation values. The fit is quite good, indicating that the fluctuation in arrival time is Gaussian-like.

The square root of the second central moment, according to (7), is equal to the pulsewidth. The pulsewidth versus distance appears in Fig. 6. The pulsewidth of the transmitted pulse is 7.91 μs. This creates a lower bound for the received pulsewidths because a pulse propagating through an ionospheric medium is not expected to become narrower than the original pulse. The range of pulsewidths varying from about 8 to 27 μs can be divided into bins to produce the histogram. Such a histogram is found to be nonsymmetrical about a mean of 13.5 μs. This indicates that on the average, the received pulses are stretched by more than 70% above the transmitted pulsewidth due to the stochastic effect alone.

Fig. 7 shows the skewness versus distance. The skewness represents the asymmetry of the pulses. Negative skewness means that energy is shifted to the leading (left) edge of the pulse. Positive skewness indicates an energy shift toward the trailing (right) edge. The average skewness (see Table I) is found to be slightly negative, but its value is comparable to the confidence interval so that no significance can be attached to this mean value. A plot of the histogram would show that the fluctuations of skewness are Gaussian-like.

The kurtosis reflects the concentration of energy near the mean time of arrival. Fig. 8 shows that the received pulses are mostly leptokurtic. That is, they have positive kurtoses, indicating that the received pulses are more peaked than Gaussian distributions. Since kurtosis is such a high-order moment, numerical errors and integration width can become very important. Therefore, the kurtosis should be interpreted conservatively.

The power spectrum for each moment is produced by multiplying the data by a Hanning window [Austen, 1983] defined as

$$W_{\text{Hanning}}(n) = \frac{1}{2}\left[1 - \cos\left(\frac{2\pi n}{N-1}\right)\right] \quad \text{for } 0 \le n \le N\text{-}1 \tag{11}$$

The length of the data array is increased by padding zeros at the beginning and end. Tapering the data with a window and padding zeros are important when computing spectral estimates and autocorrelation functions [Schwartz and Shaw, 1975]. The power spectrum is computed by taking the Fast Fourier Transform of the windowed and padded data and squaring the magnitude of the result. The power spectra for the times of arrival, pulsewidths, skewnesses, and kurtoses can all be computed similarly, but only the first two will be shown here in Figs. 9 and 10.

The power spectrum plots look like they are derived from data that is approximately Gaussian. On Fig. 9, the power spectrum of the time of arrival, the power drops approximately 10 dB from 10^{-4} to 10^{-3} rad/m. In other words, most of the power is concentrated at wavenumbers below 10^{-3} rad/m. Let l_{ta} be the scale distance of the mean time of arrival on the ground. Then

$$\begin{matrix}\text{(Scale Distance}\\ \text{of the}\\ \text{Time of Arrival)}\end{matrix} \qquad l_{ta} \approx \frac{2\pi \text{ rad}}{10^{-3} \text{ rad/m}} \approx 6 \text{ km} \tag{12}$$

The scale distance of about 6 km seems reasonable when looking at Fig. 5, the time of arrival data. A similar analysis can be applied to the higher moments. The 10 dB point on the pulsewidth power spectrum plot, Fig. 10, is also at about 10^{-3} rad/m. This gives $l_\tau \approx 6$ km. For skewness and kurtosis, the 10 dB point is at about 2×10^{-3} rad/m for both, giving $l_S \approx 3$ km for the skewness and $l_K \approx 3$ km for the kurtosis.

Assume that the ionosphere with imbedded irregularities is moving at a velocity of 50 m/s so that v_0, the velocity on the ground, is 100 m/s. The wavenumber, k_0, and the angular frequency, ω, can be related through v_0 by

$$k_0 = \frac{\omega}{v_0} \tag{13}$$

Thus, for $v_0 = 100$ m/s, the horizontal scales on Fig. 9 through 10 can be replaced by angular frequency scales. The scales of 10^{-4}, 10^{-3}, 10^{-2}, 10^{-1}, and 10^{0} rad/m can be replaced with 10^{-2}, 10^{-1}, 10^{0}, 10^{1}, and 10^{2} rad/s. From Fig. 9, it can be seen that the fluctuation power in the time of arrival at 1 rad/s is about 40 dB weaker than at 0.01 rad/s. Most of the power is concentrated at low frequencies. The fluctuation in the time of arrival has a temporal scale of approximately 2 minutes for a drift of diffraction pattern on the ground of 100 m/s. With angular frequency as the horizontal scale, the power spectrum plots of the pulsewidth, skewness, and kurtosis can be interpreted analogously.

The autocorrelation function of the data can be computed by taking the inverse Fast Fourier Transform of the power spectrum. The autocorrelation function is usually normalized so that its value is unity at the first index point, corresponding to zero shift. They are depicted in the original report [Rand and Yeh, 1989] and are not reproduced in this paper. They all show a correlation distance (defined as a distance at which the normalized correlation function is 0.5) of 1 to 1.5 km.

DISCUSSION

The propagation of high frequency radio waves through an ionospheric reflection channel has been studied. It has been shown with locus plots and plots of the received pulses that the pulses propagating through the channel may experience severe fading due to interference from a continuum of multi-path components.

Three-dimensional plots of the transfer functions and the received pulses give a qualitative view of the distortions that occur. For more quantitative analysis, temporal moments of the received pulses have been calculated. For an ionosphere with a constant drift velocity, the spatial variation can be equivalently treated as a time series. Histograms produced from the moment data show distributions that may be compared to Gaussian distributions.

Power spectra and autocorrelation functions have been computed from the moment data. The power spectra plots have been used to estimate 10 dB scale distances which are of the order of 6 km. When the flow velocity is known, the spatial scales can be translated to temporal scales. For a drift velocity of 100 m/s, the temporal scales approximate to a value of 2 minutes.

The work presented may be extended in the future by investigating techniques to produce compensation filters. These compensation filters could be used to minimize the distortion that the pulse experiences as it propagates through the ionospheric reflection channel.

Acknowledgments. This material is based upon work supported under a National Science Foundation Graduate Fellowship. Further support came from the U.S. Army Center for Command, Control, and Communication Systems under contract DAAB-07-87K-A006. The help given by J.-F. Wagen is gratefully acknowledged.

REFERENCES

Austen, J. R., Analysis of scintillation data from a laser propagation experiment, M.S. thesis, University of Illinois at Urbana-Champaign, 1983.

Booker, H. G., J. A. Ratcliffe, and D. H. Shinn, Diffraction from an irregular screen with applications to ionospheric problems, *Philos. Trans. Roy. Soc. London, Ser. A, 242,* 579-607, 1950.

Dashen, R., Path integrals for waves in random media, *J. Math. Phys., 20(5),* 894-920, 1979.

DiNapoli, F. R. and R. L. Deavenport, Numerical models of underwater acoustic propagation. In Topics in Current Physics, Ocean Acoustics, edited by J. A. DeSanto, 79-185, Springer-Verlag, Berlin, 1979

Fejer, J. A., The diffraction of waves in passing through an irregular refractive medium, *Proc. R. Soc. London, Ser. A, 220,* 455-471, 1953.

Kiang, Y. -W. and C. H. Liu, Wave scattering by an irregularity slab embedded in a stratified medium: Applications to ionospheric propagation, *Radio Sci., 20,* 80-94, 1985a.

Kiang, Y. -W. and C. H. Liu, Multiple phase-screen simulation of HW wave propagation on the turbulent stratified ionosphere, *Radio Sci., 20,* 652-668, 1985b.

Knepp, D. L., Multiple phase-screen calculation of the temporal behavior of stochastic waves, *Proc. IEEE, 71* (6), 722-737, 1983.

Kreyszig, E., Advanced Engineering Mathematics, Fifth Edition, John Wiley & Sons, New York, 1983.

Lincoln, J. V., and R. O. Conkright, International Reference Ionosphere—IRI 79 *Report UAG-82,* NOAA, Boulder, November 1981.

Liu, C. H., A. W. Wernik, K. C. Yeh and M. Y. Youakim, Effects of multiple scattering on scintillation of transionospheric radio signals, *Radio Sci., 9,* 599-607, 1974.

Mercier, R. P., Diffraction by a screen causing large random phase fluctuations, *Proc. Camb. Phil. Soc., 58,* 382-400, 1962.
Rand, T. W., and K. C. Yeh, Transfer functions and pulse distortion for an ionospheric reflection channel with embedded random irregularities, Technical Report, U.S. Army Communications-Electronics Command, Center for Command, Control, and Communications Systems, Fort Monmouth, New Jersey, 1989.
Ratcliffe, J. A., Some aspects of diffraction theory and their application to ionosphere, *Rep. Prog. Phys., 19,* 188-267, 1956.
Rino, C. L., A power law phase screen model for ionospheric scintillation 2. Strong scatter, *Radio Sci., 14,* 1147-1155, 1979.
Romanova, T. S., and T. Y. Ryzhkina, Calculation of the dispersion coefficients of pulsed signals for the oblique propagation of radio waves in the ionosphere, *Radio Eng. and Electron. Phys.* (English Translation), *21,* 1-7, 1976.
Schwartz, M., and L. Shaw, Signal Processing: Discrete Spectral Analysis, Detection, and Estimation, McGraw-Hill Book Company, New York, 1975.
Shreider, Yu. A., Editor, The Monte Carlo Method, Pergamon Press, Oxford, 1966.
Wagen, J.-F., Simulation of HF waves reflected from a turbulent ionosphere for spread spectrum systems, Ph.D. dissertation, University of Illinois at Urbana-Champaign, 1988.
Wagen, J.-F., and K. C. Yeh, Simulation of HF propagation and angle-of-arrival in a turbulent ionosphere, *Radio Sci., 24,* 196-208, 1989a.
Wagen, J.-F., and K. C. Yeh, A numerical study of waves reflected from a turbulent ionosphere, *Radio Sci., 21,* 583-604, 1986.
Wagen, J.-F., and K. C. Yeh, Numerical investigations of two-frequency mutual coherence functions of an ionospheric reflection channel, *Radio Sci. 24,* 209-223, 1989b.
Wernik, A. W., C. H. Liu, and K. C. Yeh, Model computations of radio wave scintillation caused by equatorial bubbles, *Radio Sci., 15,* 559-572, 1980.
Yeh, K. C., and C. H. Liu, Radio wave scintillation in the ionosphere, *Proc. IEEE, 70* (4), 324-360, 1982.

T. W. Rand and K. C. Yeh, Department of Electrical and Computer Engineering, University of Illinois, 1406 West Green Street, Urbana, IL 61801.

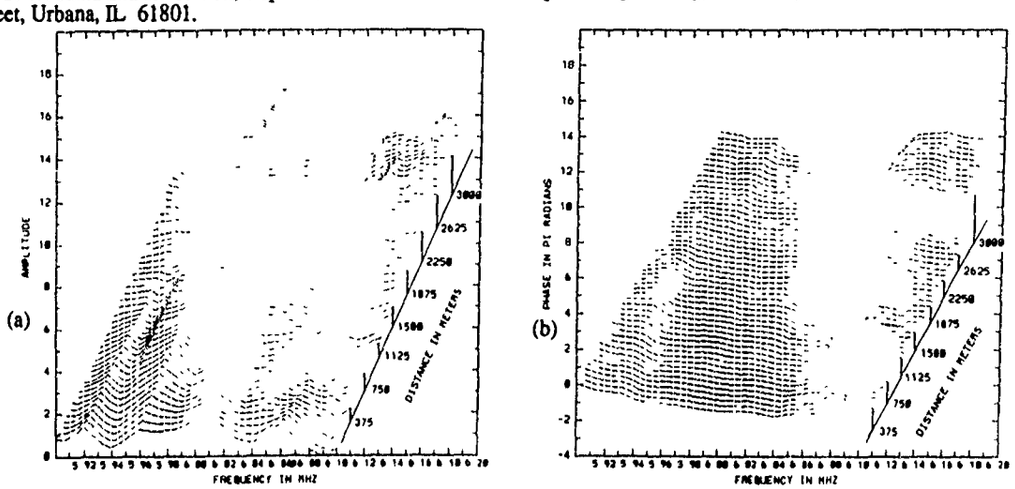


Fig. 1. The complex transfer functions ranging in distance from 0 through 3750 m. The ionospheric and propagation conditions are referred to as case 2 in the text. (a) Magnitude of the transfer function. (b) Phase of the transfer function.

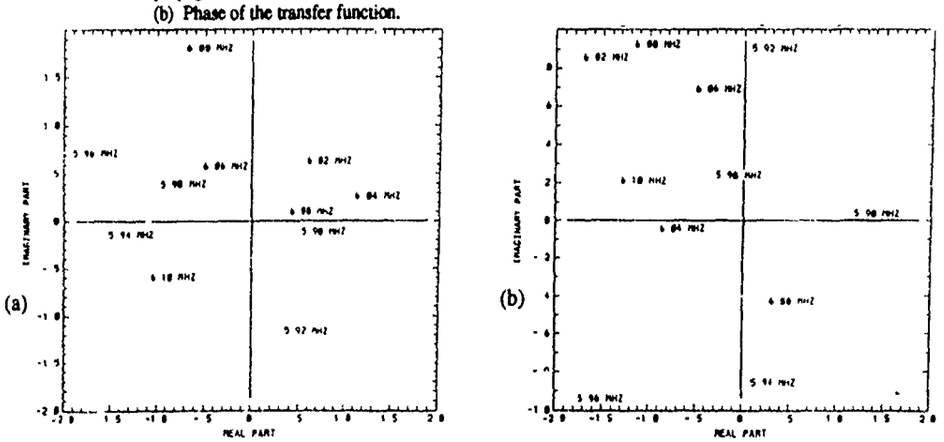


Fig. 2. Locus plot showing the real and imaginary parts of the transfer function as the frequency varies from 5.90 to 6.10 MHz for the case considered in Fig. 1. (a) At 750 m. (b) At 14,475 m.

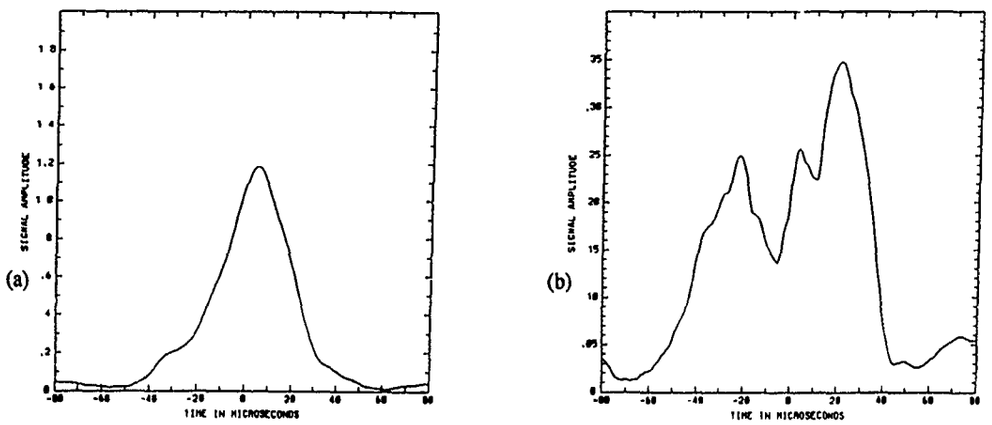


Fig. 3. Output envelope from the ionospheric channel considered in Fig. 1 when the input envelope is triangular. (a) At 750 m. (b) At 14,475 m.

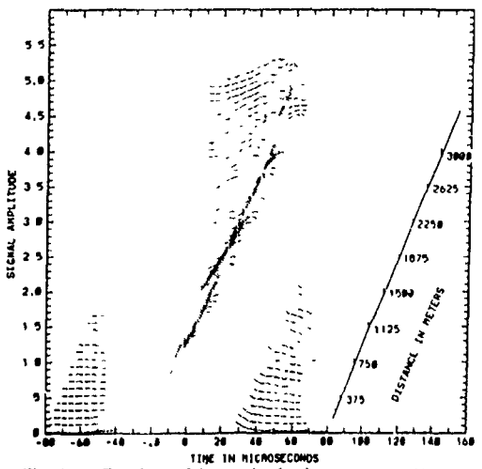


Fig. 4. Envelope of the received pulse in response to a transmitted triangular pulse ranging in distance from 0 through 3,750 m for the case considered in Fig. 1.

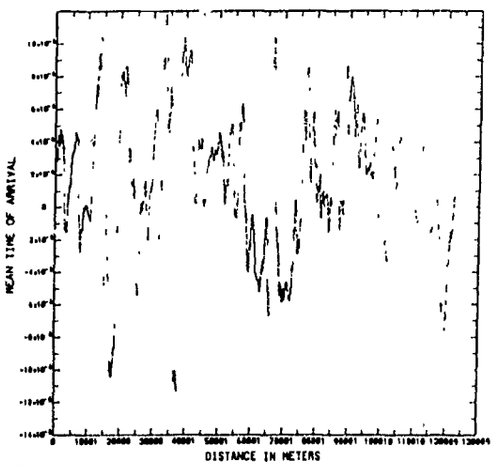


Fig. 5. The time of arrival (in seconds) of pulses received on the ground when the transmitted pulse is triangluar. The ionospheric conditions are identical to those considered in Fig. 1.

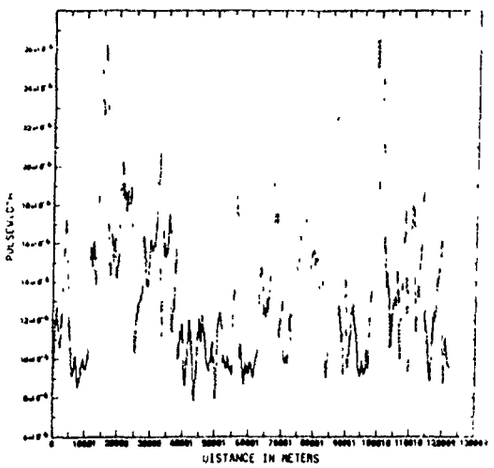


Fig. 6. The pulsewidth (in seconds) of pulses received on the ground when the transmitted pulse is triangluar. The ionospheric conditions are identical to those considered in Fig. 1.

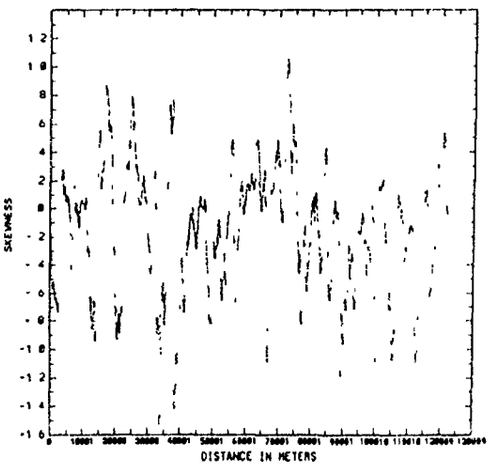


Fig. 7. The skewness of pulses received on the ground when the transmitted pulse is triangluar. The ionospheric conditions are identical to those considered in Fig. 1.

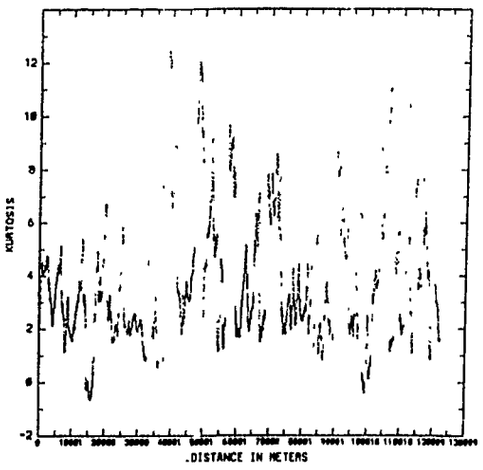


Fig. 8. The kurtosis of pulses received on the ground when the transmitted pulse is triangluar. The ionospheric conditions are identical to those considered in Fig. 1.

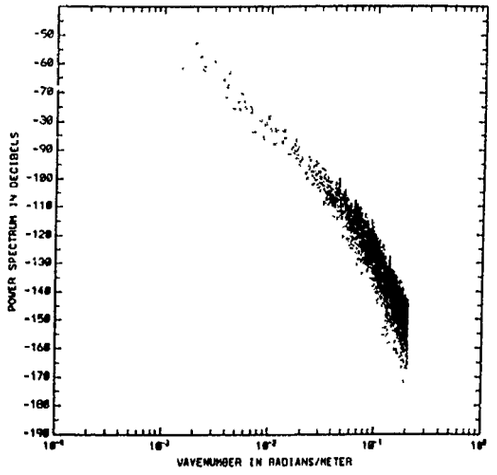


Fig. 9. Spatial power spectrum of the arrival time for triangular pulses.

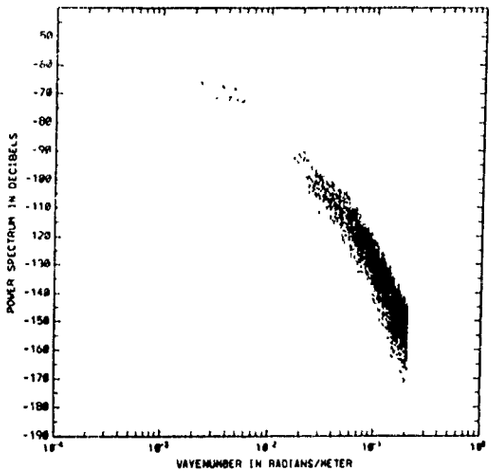


Fig. 10. Spatial power spectrum of the pulsewidth for triangular pulses.

AD-P006 287

KNOWLEDGE-BASED FADING CHANNEL ANALYZER FOR ADAPTIVE RADIO LINKS

Jeffrey S. Yalowitz, Frank G. Baird, and James M. Morrison
SRS Technologies
990 Explorer Blvd. NW
Huntsville, AL 35806

ABSTRACT

A new analysis technique proposed for fading radio channels combines object-oriented, knowledge-based processing algorithms with conventional quantitative methods to provide on-line estimates of channel dynamics and permit real-time formulation of adaptive link control strategies. Temporal features of a digitized signal envelope power record from the receiver are classified by statistical properties and threshold crossing behavior into elemental descriptive objects (e.g., increasing signal, decreasing signal), which have both symbolic and parametric attributes. An expert-system inference engine operates on the elemental objects to construct more complex descriptive objects that characterize complete signal fade waveforms and fading sequences. Techniques are being developed to appress additional rules and statistical system identification procedures to generate multipath propagation models which are consistent with the signal record and other available information. The models can be used for derivation of adaptive equalization and link control directives. The proposed technique integrates use of automated heuristic procedures and adaptive signal processing to permit effective interpretation of channel conditions in cases where time and information are limited. Applications of knowledge-based ionospheric propagation analysis include adaptive HF communications links, ionospheric monitoring, and over-the-horizon radar.

INTRODUCTION

The work reported in this paper has the objective of developing automated methods to predict channel fading events in real time, for intelligent selection of adaptive control actions by communications link terminals. The record of radio frequency (RF) signal power received over a fading channel contains, in implicit form, significant amounts of information about the state of the channel. Additional relevant information can be obtained a priori from published seasonal/ diurnal/geographic statistics and propagation forecasts. Recognition of features in the received signals, such as fade entry and recovery times, periodicity of fading sequences embedded in generally random signal variations, and correlation between signal features, can become the means to extract information important to prediction of the channel dynamics. Such pattern recognition procedures tend to depend on heuristic rules and qualitative assessment of the data, rather than on formal physical and statistical models.

91-09715

Ionospheric fading processes, as observed by the receiver, possess complex statistical properties. Certain deterministic characteristics of multipath propagation are also apparent. The statistical properties are nonstationary, but they often can be approximated over useful time intervals by stationary process models.[1,2] Prediction of fading channel behavior details a few seconds into the future may be sufficient to support critical real-time adaptive control functions in HF radio links. The conditions that cause individual fading events are not directly measurable, so any predictions about them must be based on empirical or intuitive knowledge. However, stored models of the channel can be tested against measured channel performance to estimate the validity of the models. If enough measured data is available, it may be practical to estimate the parameters of a hypothetical model accurately enough to support fade predictions using that model.

Automatic gain control (AGC) circuits incorporated in virtually all HF receivers attempt to adjust the receiver RF and IF amplifier gains to maintain the output signal level within a narrow range, despite variations of RF input signal strength. AGC responses typically follow preprogrammed gain-versus-signal-power relationships, adapting to channel states of the recent past, not to predictions of future channel states. The AGC dynamic response characteristics in most receivers represent compromises in efforts to simultaneously counter fading, interference, and differences in signal levels received successively from different locations, all without adversely affecting reception of low-frequency modulation components in the signals.

Adaptive equalization methods construct filters having impulse response related to the inverse of channel impulse response estimates. The dynamics of HF ionospheric propagation limit capability of most estimation techniques which adjust the equalizer coefficients by minimizing mean square errors under assumptions of stationary statistics.[3] Decision feedback equalizers which use recursive least squares algorithms have been proposed for HF channels, but such algorithms must exhibit rapid convergence and be computationally efficient to track a rapidly time-varying channel.[4,5] Nor do such algorithms provide information about the channel in forms directly usable for control of communications link functions external to the equalizers. One might also expect severe fading to stress the dynamic range of the signal conditioning and digitization circuits which drive equalizer algorithms.

The knowledge-based fading channel analyzer as defined for this project uses pattern recognition principles and symbolic processing techniques to identify fading channel characteristics that are of value in prediction of channel behavior. One of the results we desire is the expected variation in received signal power with time, which can be used to schedule future receiver gain settings and link error control parameters. A more ambitious desired result is a sequence of estimates of future channel impulse response, which can be used to enhance operation of adaptive equalizers and diversity transmission schemes.

APPROACH

The proposed fading channel analysis method takes advantage of object-oriented knowledge representation principles and heuristic reasoning techniques to develop temporal records of received signal power into successively higher levels of abstraction to characterize the process which produces fading phenomena contained in the signal power records. The stages of development are: (1) identification and classification of signal power variations into symbolic features or elemental objects, (2) synthesis of complex object representations of signal fade events from the elemental objects, (3) generation of multipath propagation models consistent with the

representations, and (4) prediction of future channel behavior. Figure 1 is a flow diagram of the analysis process.

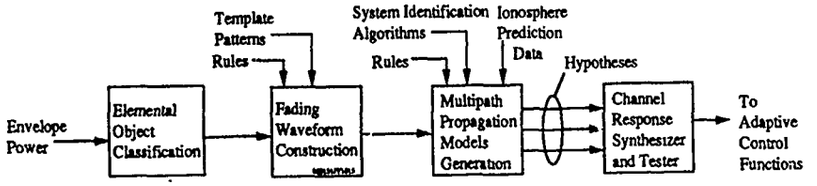


Figure 1. Knowledge-Based Fading Channel Analyzer

Our intent in formulation of the analysis approach is to adopt, where practical, techniques that are efficient in use of computation time and input data. Real time implementation of the fading channel analyzer requires that results be produced in at most a few seconds after signal measurements are taken. Simple arithmetic operations and table lookups are preferable to matrix inversions and iterative optimization algorithms when computation time is limited. The knowledge-based processing methodology, upon which the present work is based, provides structures and mechanisms compatible with these time-efficient techniques. This methodology organizes relevant information that is available *a priori* into sets of production rules and template patterns which can be accessed quickly and which permit significant results to be obained through straightforward logic and arithmetic comparison operations. Further, knowledge base constructs can be made to compensate for conditions of sparse, missing, or conflicting input data.

Combinations of numerical algorithms and expert system inferences are used at each stage of analysis. Template patterns for construction of objects to represent waveforms, fading sequences, and propagation models reside in the knowledge base, as do rules for using them to operate on the data. Patterns in objects stored in the knowledge base may be compared to corresponding patterns in observed events as steps in identification and classification. New objects may be created by transferring parameters of observed signal features to template objects from the knowledge base. Templates and rules are invoked by algorithm calls or by earlier calculations, primarily in response to occurrences of particular types of events, such as threshold crossings, local extrema in waveforms, or recognition of symbolic patterns. Heuristic relationships incorporated in the knowledge base permit construction of objects and generation of hypotheses in situations where algorithmic solutions are not practical.

ANALYSIS TECHNIQUES

Object-Oriented Signal Representation. An object is defined as a package of information and descriptions of its manipulation. In the present application, an object consists of a data structure that represents an entity. We permit both symbolic and quantitative information elements to be incorporated in objects related to the fading channel analyzer. Some of the elements may describe attributes of the physical or conceptual entity represented by the object. Other elements may identify its relationship with things in its environment. Still other elements may be rules or procedures for performing operations on the object. For example, we define a "fade" as an object that consists of an interval of "decreasing signal" followed by an interval of "increasing signal." Hierarchical arrangements are used to classify objects according to concepts such as commonality of attributes or inheritance of properties. Objects can thus be grouped into classes and subclasses, with subclasses inheriting all

characteristics of their superclasses.[6]

Figure 2 illustrates a hierarchy of objects associated with fading signals encountered in HF communications links.[7] In this example taxonomy, the classification is performed on the basis of physical phenomena which cause the observed fading effects. Significant descriptive attributes of the objects are noted. Subordinate objects possess attributes of their parent objects (e.g., skywave fading tends to have V-shaped signal power vs. time and also tends to have pseudoperiodic occurrence intervals). Knowledge-based fading channel analysis makes use of the hierarchical object structure to select hypotheses and analysis algorithms upon observation of distinctive patterns in the received signal-power.[8]

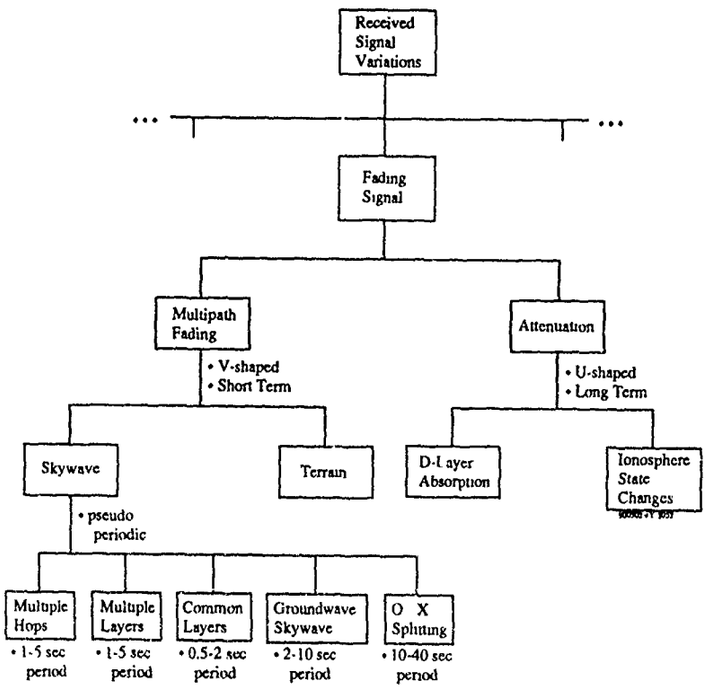


Figure 2. Hierarchy of Objects Associated with HF Fading

Fade Recognition Techniques. The received signal power waveform is sampled at a frequency much higher than the fading frequencies to be analyzed. We assign symbolic values such as "decreasing signal, increasing signal," and "level signal" to appropriate segments of the waveform. Additionally, we identify as quantitative values the magnitudes of the changes and the times of transition between segments. Figure 3 is a hypothetical fading signal record. The reader will visualize the fade within this record intuitively and have little difficulty identifying its component segments. The automated techniques proposed here identify features in the record through comparisons between signal samples. The authors have investigated classification techniques based on detection of threshold crossings and techniques based on detection of local maxima and minima to find fade components. Threshold crossing methods require maintenance of moving averages by which to set the

threshold levels. They work well in analysis of relatively stable channels, such as microwave line-of-sight links.[9]

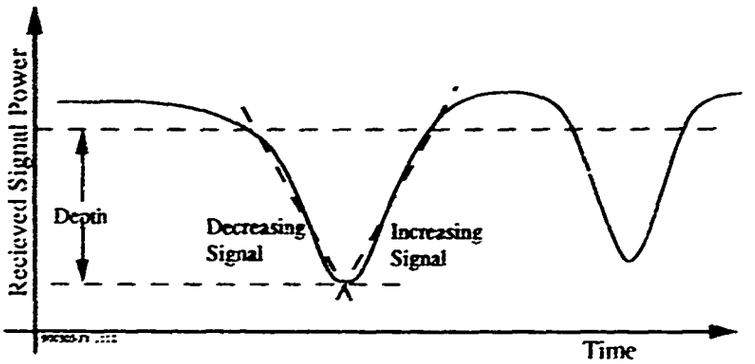


Figure 3. Fading Signal Envelope Power Record

Techniques based on extrema detection appear to be better suited to the volatile HF channel conditions, where large variations predominate the received signal record and prevent meaningful threshold signal levels from being assigned. The technique used identifies components of fade waveforms by searching for sign reversals in the first difference of the sampled signal power record, as follows: Samples $s(t_i)$ are taken at intervals $\Delta t = t_i - t_{i-1}$. The first difference of $s(t)$ at t_i is computed by

$$\Delta s(t_i) = s(t_i) - s(t_{i-1}).$$

If $\Delta s(t_{i-1}) < 0$ and $\Delta s(t_i) > 0$, and if $\Delta s(t)$ has no other sign changes within $t_i \pm m\Delta t$ (where m is a design constant), then t_i is classified as the time of a "signal null" event. If $\Delta s(t_{i-1}) > 0$ and $\Delta s(t_i) < 0$, and if $\Delta s(t)$ has no other sign changes within $t_i \pm m\Delta t$, then t_i is classified as the time of a "signal peak" event.

Given a margin M, a "decreasing signal" condition is declared to exist for a time segment $\{t_a, t_b\}$ which begins at a "signal peak" and ends at a "signal null" if $s(t_a) - s(t_b) > M$. In like manner, $s(t_c) - s(t_b) > M$ defines an "increasing signal" condition for the time segment $\{t_b, t_c\}$. A fade is declared upon recognition of "decreasing signal" followed by "increasing signal." The time of occurrence of the fade is considered to be the time of the included "signal null" event, t_b in this case. We have, somewhat arbitrarily, defined the depth of the fade to be the decrease in signal from peak to null. We have defined the fade duration as the time from peak to succeeding peak. If $|\Delta s(t)| < L$ over a time segment, where L is a relatively small constant, then the signal is considered to be level during that time segment.

Fade Prediction Techniques. Individual fades can be designated as objects using the methodology described above. Also, key descriptive attributes, such as time of occurrence, depth, and duration, can be estimated. The next step in analysis of multipath fading is to determine the presence and repetition rate of fade sequences which exhibit periodicity. Two or more such sequences may exist simultaneously. Deinterleaving capability is necessary to permit recognition of each of the embedded fade sequences. To accomplish the deinterleaving, the presence of discrete fade repetition rates is determined, then tracking filters are established to capture fades that occur at the detected repetition rates. Tracked fade sequences can be treated as objects for knowledge-based processing operations. Our goals are to establish tracks of fade sequences within a small number of fade periods and to maintain the tracks

for prediction of future fades in the sequence.

Techniques investigated for detecting fade repetition rates include (1) Fast Fourier Transforms (FFTs) of the raw sampled signal power record and (2) histograms of the time intervals between pairs of previously recognized individual fade objects. In real-time analyzer operation, we expect the histogram technique to be computationally more efficient, since it requires less complex arithmetic operations and uses fewer data inputs. In either case, the time basis for identifying repetition rates should be of the order of a few periods of the minimum rate of interest. A time basis which is too short may not reduce measurement or fading process jitter sufficiently to provide reliable initialization of tracking filters. A time basis which is too long may attempt to use obsolete data in cases where ionospheric conditions are changing rapidly.

A limited-memory least squares estimator/predictor algorithm [10] was selected to implement the tracking filter. The algorithm fits a quadratic polynomial to the occurrence times of the newest members of a series of fade objects, as follows: Let a series of J fade occurrence times be expressed as a vector $\underline{t}$, such that

$$\underline{t} = [t_1\ t_2\ \ldots\ t_J]^T,$$

where $t_j = \theta_1 + \theta_2\xi_j + \theta_3\xi_j^2,\ 1 \le j \le J$

Assume that $\underline{t}$ is of the form

$$\underline{t} = H\underline{\theta} + \underline{n},$$

where

$$H = \begin{bmatrix} 1 & 1 & 1 \\ 1 & \xi_1 & \xi_1^2 \\ & \ldots & \\ 1 & \xi_J & \xi_J^2 \end{bmatrix}$$

$$\xi_j = j,\quad 1 \le j \le J,$$
$$\underline{\theta} = [\theta_1\ \ \theta_2\ \ \theta_3],$$

and $\underline{n}$ is a noise vector. The polynomial coefficients are estimated by

$$\hat{\underline{\theta}} = [\hat{\theta}_1\ \hat{\theta}_2\ \hat{\theta}_3] = (H^TH)^{-1}H^T\underline{t}.$$

The matrix $(H^TH)^{-1}H^T$ is a constant matrix that can be computed when H is selected. This tracking filter gives equal weight to the J fade occurrences used, and zero weight to any previous samples. The use of fading memory and growing memory filters was considered, but statistical properties were not found that would justify the additional complexity associated with their implementation.

Fade sequences with constant periods and those with periods increasing or decreasing with time can be tracked by the limited-memory least squares filter. Expected fade occurrences that are missing from the sequence can be bridged by letting the filter coast through the gaps by using values of $\hat{\underline{\theta}}$ computed with previous occurrence times, or substituting estimated times for the missing samples.

Once $\hat{\theta}$ is computed for fades k through k + J of the sequence, the occurrence time of the (k + J + i)-th fade can be predicted by substituting $\xi = J + i$ into the quadratic polynomial to obtain

$$t_{J+i} = \hat{\theta}_1 + \hat{\theta}_2 (J + i) + \hat{\theta}_3 (J + i)^2$$

Limited confidence can be placed on predictions for large values of p when dealing with rapidly changing ionospheric fading conditions.

Propagation Model Estimates. Object-oriented waveform models of observed fade sequences are produced by the techniques described above. These models appear to provide for short term prediction occurrence times of future fading events with moderate confidence level. However, problems remain in making practical use of the techniques. For example, predictions of the depth of future fades are not accomplished when multiple fade sequences exist in a signal record. Two interleaved fade sequences may represent interaction of three signal components at different doppler shifts. Deepest fades can be expected to occur when the phasor sum of the three components is minimum. The phase relationships between signal component pairs are different in this case than in fades which occur when the phasor sum of the components pair is minimum.

Thus an incentive exists to fit phasor models of multipath signal components to the fade sequence models for assistance in predicting fade depths. We may wish to abstract further by fitting mathematical models of tapped delay lines with phase and amplitude modulation of the taps, as suggested by Maslin [7], to the fade sequences. Investigation of knowledge-based processing techniques for production of phasor and delay line propagation models is planned in continuation of the current work. The approach formulated is as follows: Multiple-ray template models will be designed in phasor or tapped delay line form. A rule set will be defined to select candidate templates consistent with ionospheric propagation forecasts and fade sequence on-line repetition rates computed from signal power measurements. Parameters will be estimated for the selected templates from the fade sequence repetition rates. The resulting propagation models are maintained in the knowledge-base and tested against new data as it arrives.

Considerable interest has arisen recently in the behavior of wideband HF channels (bandwidth of 1 MHz or greater). Hausman, et.al. reported on a technique to infer the impulse response of a wideband HF skywave channel from a set of narrowband measurements.[11] Wagner, et.al. reported on use of "scattering functions" (density of signal power on the delay-doppler channel plane) to describe HF channel characteristics and discussed skywave propagation modes that produce distinctive patterns in scattering function plots.[12] Adaptations of these techniques have potential applicability to knowledge-based fading channel, particularly in real-time heuristic mapping of fading channel models to ionospheric effects.

EXPERIMENTS

The knowledge-based fading channel analysis techniques were tested on HF signals received at the SRS Technologies laboratory in Huntsville, Alabama. The setup described in Figure 4 was used to conduct the tests. A short vertical monopole antenna (~3m length) was connected to the HF communications receiver An amplitude detector with a bandwidth of approximately 1 Hz was used to sample the receiver audio output. The resulting signal envelope waveform was digitized at a rate

of 5 samples/second and analyzed using the Macintosh II computer, which has National Instruments A/D and DMA controller boards installed. A NEXPERT *OBJECT* expert system shell installed in the computer was used to house the analysis rules and object structures. Individual fades were recognized in real-time during the tests. Fade sequence computations were performed off line because histogram and tracking filter algorithms had not yet been integrated into the real-time software.

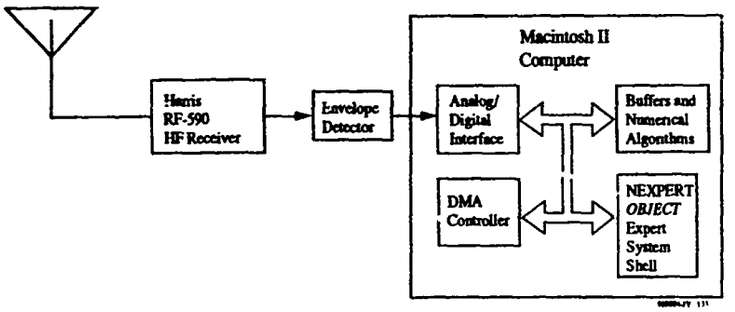


Figure 4. Laboratory Test Setup

Figure 5 is a plot of a record of signal envelope power taken from the RF carrier of CHU at 14.670 MHz. (Receiver AGC was disabled during the test.) Table 1 lists times of occurrence of individual fades in that signal power record as identified by the extreme detection technique. Figure 6 is a histogram of intervals between fade pairs in the first several seconds of the record. A period of 11 seconds (0.91 Hz) is identified as a fade repetition rate from the histogram. A least squares tracking filter initialized for that repetition rate identified the fades listed in Table 2 by occurrence time as members of a fade sequence.

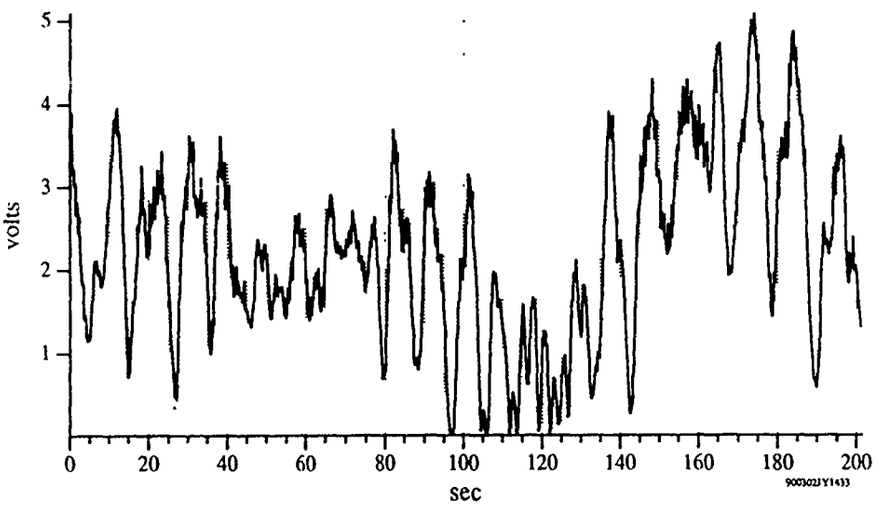


Figure 5. Record of Signal Envelope Power (RF Carrier of CHU, 14.670 MHz, Received at Huntsville, Alabama, 2 March 1990, 1433 CST)

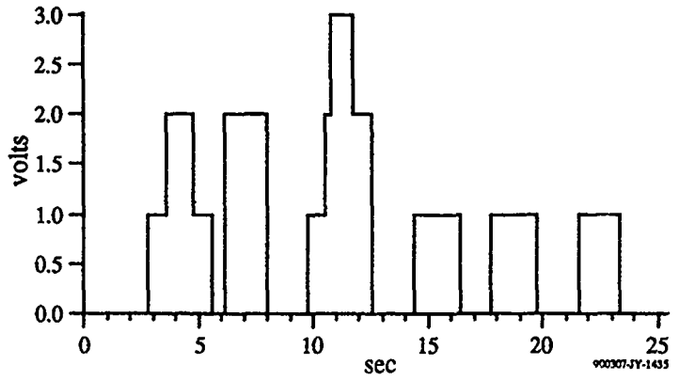


Figure 6. Histogram of Time Intervals Between Pairs of Fades

Time(sec)
4.4
8.2
15.2
19.8
27.0

Table 1. Times of Occurrence of Individual Fades Used to Initialize Tracking Filter (First 30 Seconds of Received Signal)

Actual Occurrence Time(sec)	Predicted Occurrence Time(sec)	Prediction Error(sec)
4.4	-	-
15.2	-	-
27.0	-	-
36.2	-	-
45.8	45.5	0.3
54.8	53.6	1.2
64.2	64.0	0.2
74.8	73.3	1.5
-	86.0	-
97.6	98.0	0.4
-	109.8	-
122.4	122.4	0
133.0	135.6	2.6
143.0	143.5	0.5
152.4	151.3	1.1

Table 2. Sequence of Fade Occurrence Times - Comparison of Actual Times with Predictions

CONCLUSIONS

Knowledge-based processing techniques were used successfully fo. real time identification of fade occurrences in received HF skywave signals and for short term predictions of fade occurrences subsequent to those identified. An approach has been formulated for construction of multiple-ray channel models from these results. The work performed to date is in the nature of a feasibility study. Future tasks include integration of fade recognition algorithms into an automated real-time software process, testing over a range of fading conditions, and development of heuristic mapping functions to derive propagation models from the present fade occurrence models.

REFERENCES

1. Kenneth Davies, *Ionospheric Radio Propagation*, National Bureau of Standards monograph 80, U.S. Government Printing Office, Washington, 1965.

2. Seymour Stein, "Fading Channel Issues in System Engineering", Vol. SAC-5, No.2, *IEEE Journal on Selected Areas in Communications*, February 1987.

3. A.P. Clark and S. Hariharan, "Adaptive Channel Estimator for an HF Radio Link", *IEEE Transactions on Communications*, Vol. 37, No. 9, September 1989.

4. Evangels Elftherion and David D. Falconer, "Adaptive Equalization Techniques for HF Channels", *IEEE Journal on Selectea Areas in Communications*, Vol. SAC-5, No. 2, February 1987.

5. Shalid U.H. Qureshi, "Adaptive Equalization", *Proceedings of the IEEE*, Vol. 73, No. 9, September 1985.

6. David Robson, "Object-Oriented Software Systems", *Byte*, Vol. 6, No. 8, August 1981.

7. Nichols Maslin, *HF Communications: A Systems Approach*, Plenum Press, New York, 1987.

8. Jeffrey S. Yalowitz, Frank G. Baird, James Morrison, and Rex D. Geveden, "Integrated Expert System for Sentient Radio Reoives", Rome Air Development Center Report No. RADC-TR-88-115, May 1988.

9. Jeffrey S Yalowitz, Frank G. Baird, and James M. Morrison, "Intelligent Receiver Technology Development Testbed", in *Proceedings of the 1990 Tactical Communications Conference*, in Fort Wayne, In., April 24-26, 1990.

10. James L. Melsa and David L. Cohn, *Decision and Estimation Theroy*, McGraw-Hill, New York, 1978.

11. Cynthia L. Housman, Donald R. Uffelman, and Thomas L Walton, "Wideband High Frequency (HF) Skywave Channel Parameters from Narrowband Measurements" in *Conference Record of MICOM '87*, Vol. 3, Washington, DC, October 19-22, 1987.

12. L.S. Wagner, J.A. Goldstein, and W.D. Meyers, "The HF Skywave Channel: Measured Scattering Functions for Medlatitude and Aurral Channels and Estimates for Short-Term Wideband HF Rake Modern Performance", in *Conference Record of MILCOM '89*, Vol. 3, Boston, MA, October 15-18, 1989.

AD-P006 288

PROPAGATION EFFECTS IN EXTENDED RANDOM MEDIA

L. J. Nickisch
Mission Research Corporation
Carmel, California

INTRODUCTION

In this paper the effects of scatter on fields propagated through an ionosphere with random fluctuations in electron density are examined. In particular we consider those effects associated with the finite extent of the region of irregularities (e.g. multiple scatter effects) and effects due to the bulk motion of the medium (generally nonuniform over the propagation path). The problem of electromagnetic wave propagation in structured ionization has been the topic of numerous studies. A particularly fruitful approach for the forward scatter problem involves the use of the parabolic wave equation in the Markov approximation. Weak and strong scattering solutions of the second moment equation (the equation for the mutual coherence function) have been obtained in the Rytov [*Tatarskii*, 1961, 1971], phase screen [*Knepp*, 1983a; *Rino*, 1979a, 1979b, 1981, 1982], and temporal moments [*Yeh and Liu*, 1977] theories. In none of these, however, has both a variable medium density and variable medium velocity over the propagation path been considered, and rarely has the mutual coherence function been expressed simultaneously for spaced frequency, position, and time. We will present a formulation incorporating these generalities which results in an analytical expression for the two frequency - two position - two time mutual coherence function.

The method we present is essentially a multiple phase screen (MPS) approach in that a phase-screen/diffraction (PDM) procedure is employed. It differs from the usual MPS procedure, however, in that rather than solving the parabolic wave equation for the wave field in a random realization and obtaining the mutual coherence function by ensemble averaging, the parabolic equation for the second moment is solved directly and analytically using a quadratic structure function approximation. (The exact structure function can also be employed using this phase-screen/diffraction method at the inconvenience of numerical solution rather than analytic.)

We first motivate the need for such a formulation by examining scattering functions measured by the DNA HF Channel Probe in the polar region. The scattering function is the two dimensional Fourier transform of the two frequency - two time mutual coherence function, and it describes the spreading of signals by the channel in delay and doppler frequency. These measured scattering functions display a wide variety of shapes in delay and doppler which do not conform to the simple single thin phase screen - uniform motion result. Next the phase-screen/diffraction solution is presented. Finally, three DNA HF Channel Probe examples are analyzed using PDM.

MEASURED SCATTERING FUNCTIONS

The high latitude ionosphere is characterized by structured ionization typically aligned along the earth's magnetic field. Structure scale sizes parallel to the field lines may be more than ten times as large as the perpendicular scale sizes. It is the structure perpendicular to the field which dominates in producing scintillations of radio signals. These amplitude and phase fluctuations are the result of distortions imposed on the wave as it propagates through the irregularities.

A commonly used formulation for the calculation of stochastic effects on electromagnetic waves due to structured ionization has been developed under the sponsorship of the Defense Nuclear Agency (DNA) [*Wittwer*, 1979, 1982] for transionospheric links. In this regime the frequency of the wave is always substantially above the plasma frequency of the ionosphere so that approximations which make use of the large carrier to plasma frequency ratio are valid. While this is not generally true for HF propagation, it remains true in moderately to highly oblique propagation, and the transionospheric propagation theory remains valid.

91-09714

Measurements of signal strength as a function of frequency and delay were made under the sponsorship of DNA using the HF Channel Probe. The measurements we will consider were made over a one hop path in the polar region. The link, located in Greenland, consisted of a transmitter at Narssarssuaq and two receivers at Thule, 1913 km to the north. Details on the experimental setup and measurements are contained in Basler et al., [1985, 1987, 1988]. The measurements yield information on the degree of spreading in delay and doppler frequency suffered by signals at HF due to the naturally occurring ionization structure of the polar region.

We have previously applied the transionospheric scintillation model of Wittwer to the polar channel probe data [*Nickisch*, 1986; *Nickisch and McCartor*, 1987]. In these studies only the overall delay and doppler spreads were considered (i.e., no consideration was given to the delay-doppler coupling or scattering function shape). Wittwer's formulation provides expressions for these quantities for extended media. The doppler spread (or decorrelation time) result is based on a solution of the parabolic equation for the two position (and hence two time via Taylor's hypothesis) mutual coherence function. The delay spread (or frequency coherence bandwidth) solution uses the temporal moments theory. We found that the model was capable of reproducing the measured spreads adequately for reasonable values of the physical parameters of the model. An alternate, though substantially similar model [*Basler et al.*, 1985; *Price*, 1985] produced similar results.

Basler et al., however, eventually considered the full scattering function shape. Their model [*Basler et al.*, 1987, 1988] utilized a single thin phase screen orthogonal to the ray at midpath. They found that the single thin screen formulation was incapable of describing the variety of delay-doppler couplings evident in the measured data. They resorted to using an *ad hoc* spread function to smear the scattering functions in doppler. With this generality they were able to reproduce many of the varied scattering function shapes. The results were unsatisfying, however, in that no physical basis for the required doppler spread function could be found. We will show that the varied scattering function shapes can be explained in a completely physical way by considering extended medium effects (multiple scatter) and nonuniform plasma velocity profiles over the signal path.

Several examples of the DNA HF Channel Probe measured scattering functions are shown in Figure 1. Two and sometimes three modes are apparent in the cases shown. The latter two modes correspond to the high and low ray propagation paths for one hop F-layer reflection. The occasionally appearing early delay mode is a two hop E-layer return. In single phase screen theory, a simple parabolic delay-doppler coupling is predicted under the usual Taylor's hypothesis (uniformly drifting frozen plasma structure). This arises because all ionization structure is collapsed to a single screen at some point along the signal trajectory. All angular scatter occurs at that point, providing a direct geometrical relation (of parabolic form) between angle of signal arrival at the receiver and signal delay. The assumption of uniform plasma drift then provides a direct (delta function) relation between angle of arrival and doppler shift. In only one case in Figure 1 is the parabolic structure associated with uniform plasma motion realized. Most of the cases instead show a quite different delay-doppler relationship. Little or no wing-like structure appears, and instead the scattering functions range from teardrop shapes to cases exhibiting no delay-doppler correlation.

The fact that the channel probe measured scattering functions do not generally conform to the single screen result is not surprising. Typical HF ray paths spend a large portion of their trajectories in significant ionization. Thus multiple scatter effects are likely to be important and the single thin screen approximation is suspect. Also, because of the plasma convective flow pattern of the polar-auroral ionosphere and the long propagation path to the north, it is very likely that the wave will encounter changing plasma velocity structure over the path. Multiple scatter will reduce the correlation of angle of arrival with delay. Furthermore, if the plasma velocity is nonuniform over the ray trajectory, the delta correlation between ray angle and doppler shift is destroyed. In the extreme, the plasma motion could be completely turbulent. In this case no correlation between ray angle and doppler exists, and the scattering function will be the uncorrelated product of a delay profile function with a doppler profile function. Examples of this nature are apparent in Figure 1. More generally one would expect to see scattering functions representing intermediate states between these extremes. We will show in the next section that these expectations are born out in the multiple screen formulation.

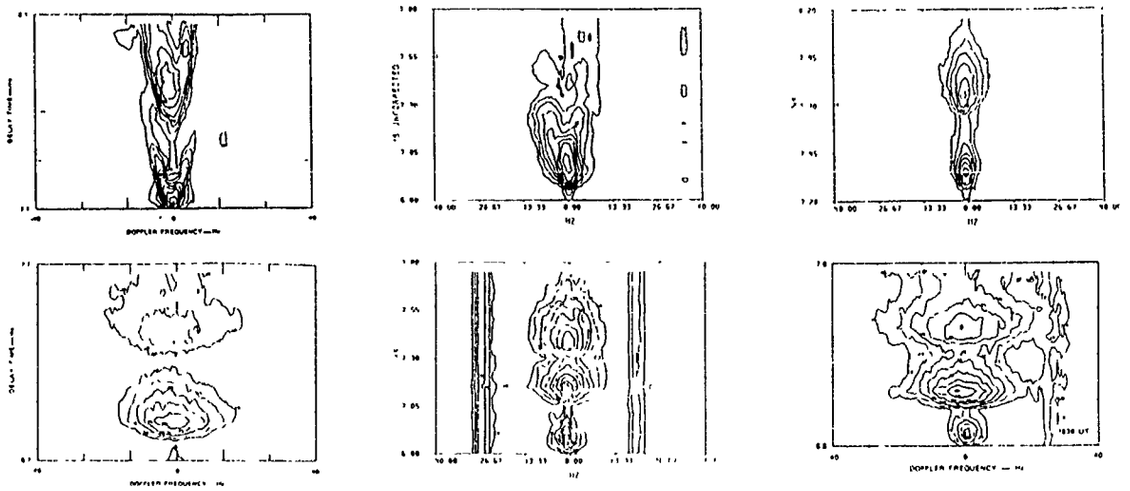


Figure 1. Examples of DNA HF Channel Probe measured scattering functions showing departure from the single thin screen uniform plasma motion form. In general two modes corresponding to low and high ray F-layer returns are evident. The occasional third early delay mode is the two hop E-layer return.

THE PHASE-SCREEN/DIFFRACTION METHOD

The scattering function (delay-doppler power spectrum) is the two-dimensional Fourier transform of the two frequency - two time mutual coherence function. In order to include the effects of multiple scatter and varying plasma density and velocity, the parabolic equation for this mutual coherence function must be solved without resorting to the single thin screen approximation.

The phase-screen/diffraction method (PDM) is one approach which, as we will show, allows this generality to be incorporated. In this method, a solution for the two frequency - two position - two time mutual coherence is obtained for an extended medium approximated by an arbitrary number of thin screens. The solution converges to the continuum result rather rapidly with screen number. An analytical solution for the two frequency - two position mutual coherence function for the special case of a uniform slab of electron density was derived by Knepp [*Knepp*, 1983b]. This solution serves as a check on the convergence properties of the phase-screen/diffraction solution.

In the Markov approximation, the point source parabolic equation for the mutual coherence function (actually the ratio of the full mutual coherence function to the MCF in the absence of irregularities) is [*Knepp*, 1983b]

$$\frac{\partial \Gamma}{\partial z} - \frac{i}{2}\frac{k_d}{k^2}\nabla_d^2\Gamma + \frac{1}{z}\left[x\frac{\partial}{\partial x} + y\frac{\partial}{\partial y}\right]\Gamma + A(\omega_d, x, y, t; z)\Gamma = 0 \tag{1}$$

In this equation, z is the coordinate in the propagation direction, x and y are separation variables transverse to the propagation direction, t is the temporal separation, k_d is the wavenumber associated with the angular frequency separation ω_d, and k is the wavenumber of the average frequency. It has been assumed that the wave carrier frequency is everywhere large compared to the plasma frequency, that $k_d \ll k$, and that the source term A depends on transverse spatial and temporal separation only. Certain near zone terms have been neglected as well. The source term can be written as

$$A = \frac{k_p^4}{8k^2}d_\xi(x,y) + \frac{k_d^2 k_p^4}{8k^4}A_\xi(0,0) \tag{2}$$

where

$$d_\xi(x,y) = 2\left[A_\xi(0,0) - A_\xi(x,y)\right] \tag{3}$$

is the differential structure function of the relative electron density fluctuations $\xi = \Delta N_e/\langle N_e\rangle$, and $A_\xi(x,y)$ is the transverse autocorrelation function of the irregularities. Here k_p is the wavenumber of the local plasma frequency, generally z dependent.

A change of variables simplifies the spherical wave contribution to Equation 1. Let $\theta_x = x/z$ and $\theta_y = y/z$. The parabolic equation becomes

$$\frac{\partial \Gamma}{\partial z} - \frac{i}{2}\frac{k_d}{k^2}\frac{1}{z^2}\left[\frac{\partial^2}{\partial \theta_x^2} + \frac{\partial^2}{\partial \theta_y^2}\right]\Gamma + A(\omega_d, \theta_x z, \theta_y z, t; z)\Gamma = 0 \tag{4}$$

In Equation 4, the transverse derivative terms account for diffraction effects (angular scatter).

The phase-screen/diffraction method replaces the continuous irregularity medium by a number n_s of thin screens. Diffraction is ignored at a screen, but is allowed to develop in the free space between. We write the multiple screen source term as

$$A(\omega_d, x, y, t; z) = \sum_{i=1}^{n_s} A_i(\omega_d, \theta_x, \theta_y, t; z)\delta(z - z_i) \tag{5}$$

where z_i is the location of screen i. Let $z_{i,-}$ ($z_{i,+}$) be the propagation coordinate just before (after) screen i. Ignoring the diffraction terms at a screen, we have

$$\Gamma(\omega_d, \theta_x, \theta_y, t; z_{i,+}) = \Gamma(\omega_d, \theta_x, \theta_y, t; z_{i,-}) \exp\left\{-A_i(\omega_d, \theta_x, \theta_y, t; z_i)\right\} \tag{6}$$

Between screens, the source term vanishes and the solution can be obtained by Fourier transforming,

$$\tilde{\Gamma}(\omega_d, K_{\theta_x}, K_{\theta_y}, t; z_{i,-}) = \tilde{\Gamma}(\omega_d, K_{\theta_x}, K_{\theta_y}, t; z_{i-1,+}) \exp\left\{-i\omega_d\left[K_{\theta_x}^2 + K_{\theta_y}^2\right]P_{i-1}\right\} \tag{7}$$

where the free space propagation is described by

$$P_i = \frac{1}{2c}\frac{1}{k^2}\left[\frac{1}{z_i} - \frac{1}{z_{i+1}}\right] \tag{8}$$

The boundary condition at the first screen is $\Gamma(\omega_d, \theta_x, \theta_y, t; z_{1,-}) = 1$ since the full MCF is equal to the irregularity free MCF before the first screen.

The above equations constitute the phase-screen/diffraction method for the calculation of the two frequency - two position - two time mutual coherence function. One simply Fourier transforms back and fourth between (6) and (7), propagating from screen to screen. This must in general be done numerically for general irregularity structure functions. In the quadratic structure function approximation, however, an analytic solution can be obtained for an arbitrary number of screens.

Assuming arbitrarily oriented cylindrical structure, the transverse autocorrelation function of the irregularities in the quadratic structure function approximation can be written as

$$A_\xi(x, y) = A_o + A_{2x}x^2 + A_{2y}y^2 \tag{9}$$

The temporal dependence of the medium is incorporated via Taylor's hypothesis, but each screen is allowed to take on an independent velocity $\vec{v}_i$. Defining $\vec{\nu}_i = \vec{v}_i/z_i$, the source term for each screen is then

$$A_i(\omega_d, \theta_x, \theta_y, t; z) = S_{xi}(\theta_x - \nu_{xi})^2 + S_{yi}(\theta_y - \nu_{yi})^2 + B_i\omega_d^2 \tag{10}$$

where for propagation through a medium of extent L we have

$$S_{xi} = -\frac{L}{n_s}\frac{k_{pi}^4}{4k^2}A_{2x}z_i^2, \quad S_{yi} = -\frac{L}{n_s}\frac{k_{pi}^4}{4k^2}A_{2y}z_i^2, \quad B_i = \frac{A_o}{c^2}\frac{L}{n_s}\frac{k_{pi}^4}{8k^4} \tag{11}$$

The factor L/n_s is the length of the medium assumed collapsed to the i-th screen, here taken to be the same for all screens. It can be allowed to vary if desired.

Note that the differential structure function is $d_\xi = -2(A_{2x}x^2 + A_{2y}y^2)$. The quadratic structure function approximation allows the required Fourier transforms for propagation through multiple screens to be performed analytically. The solution is expressed in terms of a number of recursive quantities. These are

$$D_{xi} = \left[\frac{1}{D'_{xi}} + 4i\omega_d P_i\right]^{-1}, \quad D'_{xi} = D_{x,i-1} + S_{xi}, \quad D_{yi} = \left[\frac{1}{D'_{yi}} + 4i\omega_d P_i\right]^{-1}, \quad D'_{yi} = D_{y,i-1} + S_{yi} \tag{12}$$

with $D_{x0} = D_{y0} = 0$, and

$$\nu'_{xi} = \frac{1}{D'_{xi}}\left[D_{x,i-1}\nu'_{x,i-1} + S_{xi}\nu_{xi}\right], \quad \nu'_{yi} = \frac{1}{D'_{yi}}\left[D_{y,i-1}\nu'_{y,i-1} + S_{yi}\nu_{yi}\right] \tag{13}$$

By beginning at the first screen and analytically applying the phase-screen/diffraction method through two or three screens, the pattern of the mutual coherence function solution becomes apparent, and it is possible to write down the general solution. For propagation to z_R, it is

$$\Gamma(\omega_d, x, y, t; z_R) = \left\{ \prod_{j=1}^{n_s} \sqrt{\frac{D_{xj} D_{yj}}{D'_{xj} D'_{yj}}} \right\} \exp\left[-\omega_d^2 \sum_{j=1}^{n_s} B_j \right] \exp\left[-D_{xn_s} \left(\frac{x}{z_R} - \nu'_{xn_s} t \right)^2 - D_{yn_s} \left(\frac{y}{z_R} - \nu'_{yn_s} t \right)^2 \right]$$

$$\exp\left[-t^2 \sum_{j=1}^{n_s-1} \left\{ D_{xj} \nu'^2_{xj} + S_{x,j+1} \nu^2_{x,j+1} - D'_{x,j+1} \nu'^2_{x,j+1} + D_{yj} \nu'^2_{yj} + S_{y,j+1} \nu^2_{y,j+1} - D'_{y,j+1} \nu'^2_{y,j+1} \right\} \right] \tag{14}$$

Propagation through a medium with varying electron density and plasma velocity over the propagation path can be approximated by this two frequency - two position - two time mutual coherence function. Aside from the assumptions inherent in the parabolic equation (including the neglect of refractive effects) and the local use of Taylor's hypothesis, the accuracy of the solution is limited only by the number of screens used to describe the medium.

The spatial wavenumber spectrum, or equivalently the angle of arrival spectrum, is obtained from Equation 14 by setting the temporal separation to zero and Fourier transforming over spatial separation and frequency difference. The intermediate result (before transforming over ω_d) is

$$\tilde{\Gamma}(\omega_d, K_x, K_y, t = 0; z_R) = \frac{\pi z_R^2}{\sqrt{D'_{xn_s} D'_{yn_s}}} \left\{ \prod_{j=1}^{n_s-1} \sqrt{\frac{D_{xj} D_{yj}}{D'_{xj} D'_{yj}}} \right\} \exp\left[-\omega_d^2 \sum_{j=1}^{n_s} B_j \right] \exp\left[-\frac{z_R^2}{4} \left\{ \frac{K_x^2}{D_{xn_s}} + \frac{K_y^2}{D_{yn_s}} \right\} \right] \tag{15}$$

The final transform must be done numerically in general,

$$S(\tau, K_x, K_y) = \frac{1}{2\pi} \int_{-\infty}^{\infty} \tilde{\Gamma}(\omega_d, K_x, K_y, t = 0; z_R) e^{-i\omega_d \tau} d\omega_d \tag{16}$$

The case of propagation orthogonal to irregularities highly elongated in the $\hat{y}$ direction (one dimensional structure) is obtained by setting both y and A_{2y} to zero (i.e. $D_{yj}/D'_{yj} = 1$) in Equation 14 and transforming over x and ω_d. Figure 2 displays the resulting angle of arrival spectrum in 5 dB contours for a 2000 km propagation path with 1000 km of centered uniform density intervening plasma, represented as a single screen, three screens, and twelve screens. The effect of the extended medium is apparent in the multiple screen results. At each spatial wavenumber the spectrum rises quickly and decays more slowly. The fact that this is true even at $K_x = 0$ implies that multiple scatter is important. Single scatter occurring at only one of the spatially separated screens (summed over all screens), would have an effect only at nonzero K_x.

Most of the extended medium effect is already manifested at two screens. Comparisons to Knepp's thick layer result [*Knepp*, 1983b] have shown that by six screens only small differences remain, and the twelve screen result is essentially indistinguishable from this exact solution of the parabolic equation. Although convergence is attained with a fairly small number of screens, larger numbers of screens allows more flexibility in representing rapidly varying electron density and plasma velocity profiles.

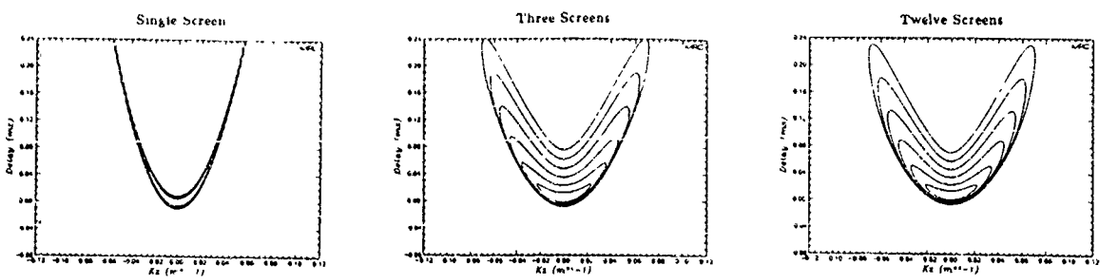


Figure 2. PDM angle of arrival spectra for a 10 MHz signal propagating over a 2000 km path with 1000 km of centered 4 MHz plasma represented by (a) one screen, (b) three screens, (c) twelve screens (5 dB contours). The irregularity structure parameters were chosen to be typical of the polar disturbed ionosphere. The twelve screen result is indistinguishable from the exact continuum solution for a uniform slab.

The doppler spectrum is obtained by setting the spatial separations in Equation 14 to zero and Fourier transforming over t and ω_d. Defining

$$\begin{aligned}\alpha &= D_{xn_s}\nu'^2_{xn_s} + \sum_{j=1}^{n_s-1}\left\{D_{xj}\nu'^2_{xj} + S_{x,j+1}\nu^2_{x,j+1} - D'_{x,j+1}\nu'^2_{x,j+1}\right\} \\ &+ D_{yn_s}\nu'^2_{yn_s} + \sum_{j=1}^{n_s-1}\left\{D_{yj}\nu'^2_{yj} + S_{y,j+1}\nu^2_{y,j+1} - D'_{y,j+1}\nu'^2_{y,j+1}\right\} \end{aligned} \tag{17}$$

we have

$$\Gamma_{dop}(\omega_d, x=0, y=0, \omega_{dop}; z_R) = \sqrt{\frac{\pi}{\alpha}}\left\{\prod_{j=1}^{n_s}\sqrt{\frac{D_{xj}D_{yj}}{D'_{xj}D'_{yj}}}\right\}\exp\left[-\omega_d^2\sum_{j=1}^{n_s}B_j\right]\exp\left[-\frac{\omega_{dop}^2}{4\alpha}\right] \tag{18}$$

Again the final transform must be performed numerically in general,

$$S(\tau, \omega_{dop}) = \frac{1}{2\pi}\int_{-\infty}^{\infty}\Gamma_{dop}(\omega_d, x=0, y=0, \omega_{dop}; z_R)e^{-i\omega_d\tau}d\omega_d \tag{19}$$

The doppler spectrum corresponding to the angle of arrival spectrum of Figure 2 for the case of a uniformly moving plasma is shown in Figure 3. It is similar in form to the angle of arrival spectrum, having a basic parabolic or wing-like structure for the delay-doppler coupling, but drawn out on the late delay side. As an approximation to a totally turbulent medium, Figure 4 shows the scattering function when the velocity of each successive screen is reversed in direction. There is essentially no delay-doppler coupling, and the scattering function can be viewed as the uncoupled product of a delay profile with a doppler profile. Figure 5 is an example of two stream plasma flow across the ray path. Again no wing-like structure exists (as could have been anticipated by symmetry arguments). In Figure 6, a sharp velocity shear is imbedded in an otherwise smoothly varying velocity profile. Some wing-like structure remains due to the dominating positive plasma flow, but the velocity shear causes a broad doppler spread at early delays.

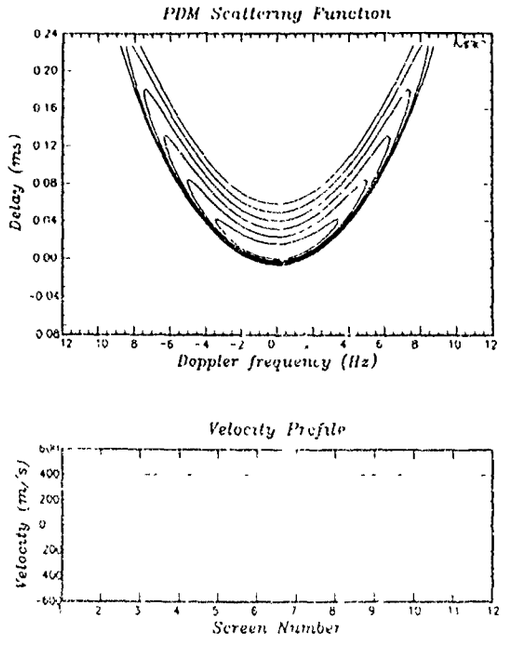


Figure 3. PDM scattering function corresponding to Figure 2 for a uniformly moving plasma represented by twelve comoving screens (5 dB contours).

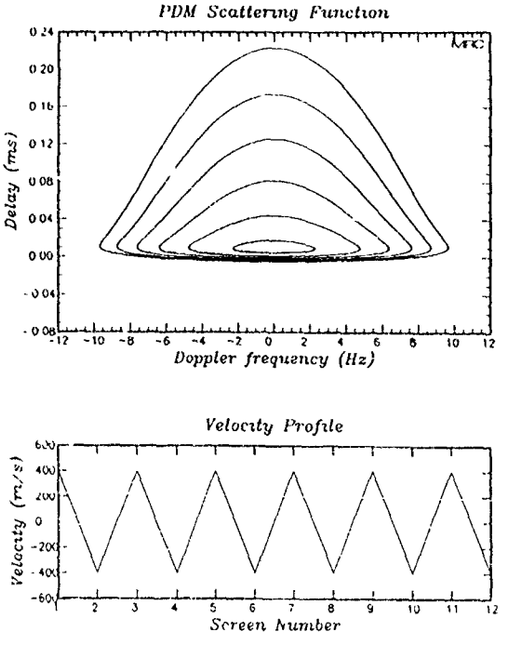


Figure 4. PDM scattering function for a turbulent plasma represented by twelve screens which successively reverse velocity.

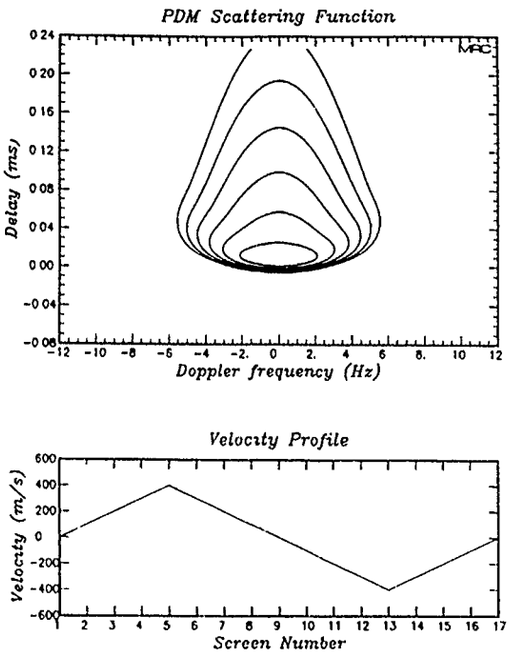


Figure 5. PDM scattering function for two stream flow.

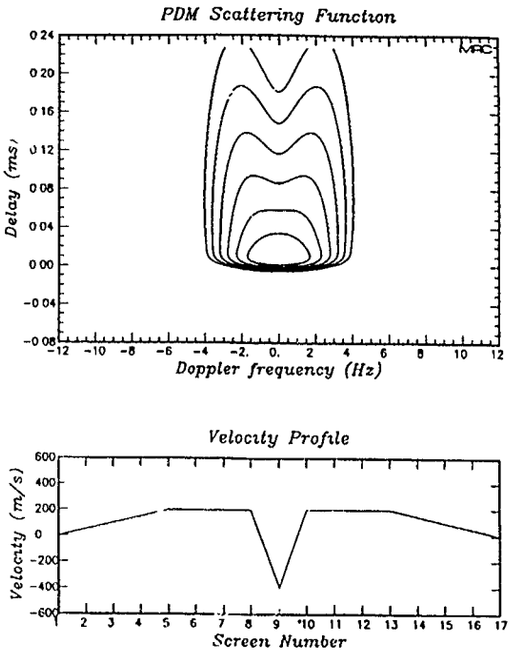


Figure 6. PDM scattering function showing the effect of velocity shear.

PDM ANALYSIS OF POLAR DATA

The phase-screen/diffraction method with nonuniform velocity profiles produces a variety of scattering function shapes. We now look at some specific examples of the measured polar scattering functions to see the extent to which they can be reproduced by PDM. Our approach is to utilize two-dimensional electron density profiles measured by the Sondre Stromfjord incoherent backscatter radar (nearly in the plane of the Channel Probe trajectories and nearly concurrent in time) to construct stratified model ionospheres. Once model ionospheres are constructed, numerical ray tracing is used to propagate signals over the DNA HF Channel Probe polar link via one hop F-layer low and high ray paths. The electron densities and required path geometry information output by the ray trace are then used in conjunction with an irregularity structure model (detailed in [*Nickisch*, 1990]) to construct a series of screens. The irregularity structure model parameters are adjusted until a reasonable representation of the measured delay profile is attained. Then various velocity profiles are applied until one is obtained which yields (more or less) the observed delay-doppler coupling.

Our irregularity structure model assumes that the electron density fluctuations scale directly with electron density, i.e. $\sigma_{N_e} = \xi\langle N_e \rangle$. The two component power law form used has an *in situ* spectral index of 1.8 at larger scales and breaks to a slope $n' = 2$ at the assumed freezing scale of $\ell_f = 1$ *km*. The structure is assumed elongated along the earth's magnetic field with an orthogonal outer scale L_o and parallel outer scale $L_\parallel$.

The first measured scattering function we examine is from March 20, 1985 (Day 79) of the HF Channel Probe polar spring campaign, shown in Figure 7 (5 dB contours). The carrier frequency is 10.57 MHz. Using a modeled ionosphere as described above, we found the high and low ray paths corresponding to the two modes of the scattering function and used the electron densities along these paths to construct twenty-five screens over each path, equidistant in ground range. We first fit the delay profile of this measurement obtained by integrating over doppler. The delay profile was well reproduced using PDM with $\xi = 0.22$, $L_o = 30$ *km*, $L_\parallel = 100$ *km*. We then tried various velocity profiles in an attempt to reproduce the measured delay-doppler coupling. Figure 8 is the PDM scattering function for one example velocity profile. The fit is not perfect, but reasonable, especially considering that we have assumed the same structure parameters and plasma velocities for both modes. We point out again that the true velocity profile was probably not this one. The PDM scattering function is fairly insensitive to the velocity of the screens near the endpoints since the ionization level is low there (the peak electron densities

occur over the central portion of the propagation path corresponding to the apex of the trajectory). The strict symmetry of the velocity profile about midpath is also unlikely. However the weakly parabolic form of the measured scattering function requires a dominating bulk plasma motion over the path as a whole, and the foreshortened wings require some shear in the region of strong ionization.

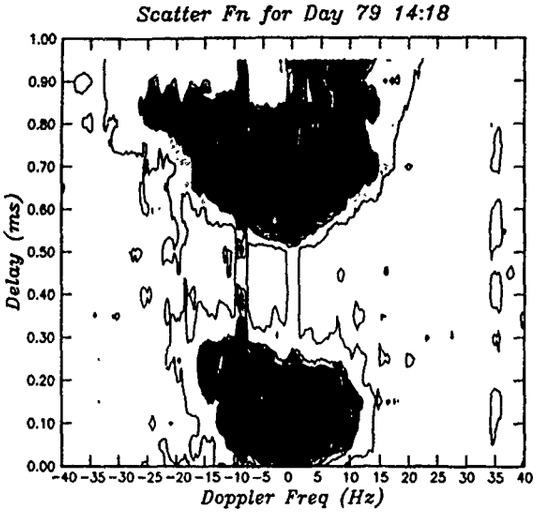


Figure 7. DNA HF Channel Probe measured scattering function for March 20, 1985, 14:18 UT (5 dB contours).

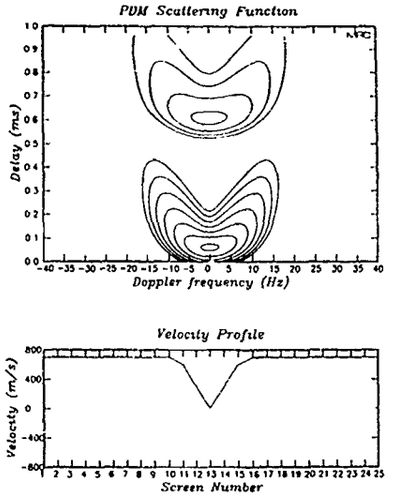


Figure 8. PDM scattering function fit of the measured scattering function in Figure 7.

The second case is from the same day (Day 79, 1985) as the first example, but several hours later and at a carrier frequency of 10.265 MHz. The measured scattering function, Figure 9, has less doppler spread than the first example, and shows no parabolic delay-doppler coupling whatsoever. Instead the modes are more or less teardrop shaped. A simple two stream flow does very well at reproducing the observed form of delay-doppler coupling, as shown in Figure 10. The structure parameters in this case are $\xi = 0.2$, $L_o = 30\ km$, $L_{\parallel} = .00\ km$.

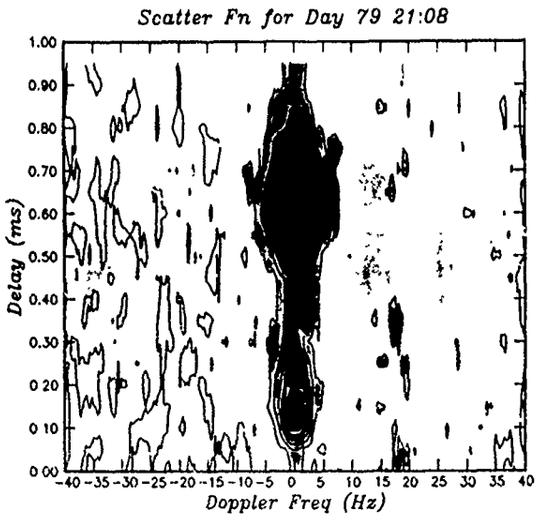


Figure 9. DNA HF Channel Probe measured scattering function for March 20, 1985, 21:08 UT (5 dB contours)

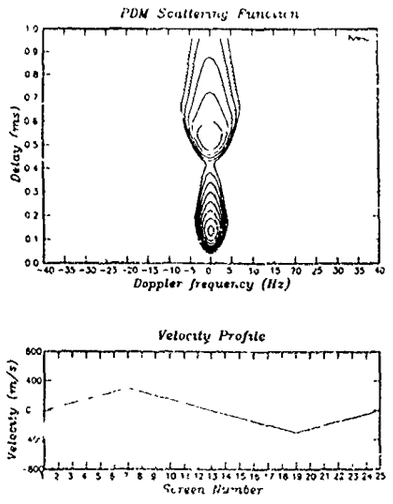


Figure 10. PDM scattering function fit of the measured scattering function in Figure 9.

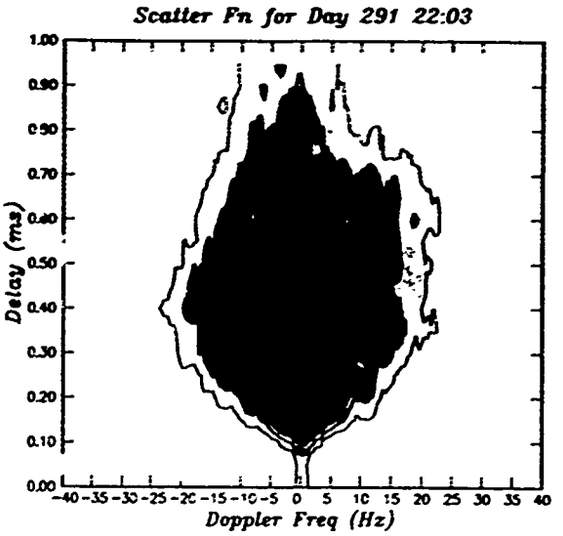


Figure 11. DNA HF Channel Probe measured scattering function for October 17, 1984, 22:03 UT (5 dB contours).

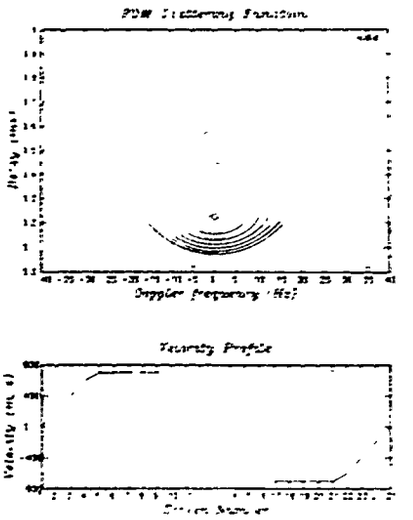


Figure 12. PDM scattering function fit of the measured scattering function in Figure 11.

The third scattering function, Figure 11, was measured during the 1984 fall campaign on October 17 (Day 291). In this case the two modes have almost coalesced into one, implying that the 11.26 MHz carrier frequency is very near the MUF of the link. Figure 12 shows that again two stream flow suffices to reproduce the observed delay-doppler coupling. The structure parameters used here are $\xi = 0.4$, $L_o = 100\ km$, $I_\cdot = 100\ km$. Note that the transverse outer scale is equal to the parallel outer scale in this fit. Elongated structure produced too rapid a decay in delay, and it was necessary to go all the way to isotropic structure in order to match the measured slow decay.

CONCLUSION

Using the phase-screen/diffraction method, an analytical expression for the two frequency - two position - two time mutual coherence function has been obtained. The solution places no restrictions on the number of screens used, and holds for differing electron densities and plasma velocities from screen to screen. Thus propagation through extended random media with nonuniform electron density and plasma velocity profiles can be considered. Spectra obtained as Fourier transforms of the mutual coherence function have been studied. The effect of the extended medium (particularly the effect of multiple scatter) was shown to be considerable in the delay - angle of arrival spectrum for HF propagation, and the PDM result rapidly converges with screen number to the continuum result. The PDM delay - doppler spectrum (scattering function) depends very sensitively on the plasma velocity profile over the signal path. Our comparisons to the measured scattering functions of the DNA HF Channel Probe show that the varied scattering function shapes observed can in large part be explained in terms of nonuniform velocity profiles. For example, broad doppler spreads at early delay are explained by turbulent plasma velocity structure and plasma shear. Parabolic delay-doppler coupling signifies dominant plasma motion in some direction, but more typically the measured delay-doppler contours exhibit no such wings and are better represented by transitions from positively to negatively moving plasma, whether turbulent or smooth (e.g.two stream flow).

The phase-screen/diffraction method provides a simple means of incorporating the generalities of extended media, nonuniform electron density, and nonuniform plasma velocity. Since the mutual coherence function is obtained directly rather than requiring ensemble averaging of random realizations, it is comparatively efficient numerically, requiring only the evaluation of certain recursive quantities. The quadratic structure function approximation (which provided the analytical expression for the mutual coherence function) need not be made, but at the expense of performing Fourier transforms at each screen and hence substantially reducing numerical efficiency. The quadratic approximation does, however, appear to be adequate for disturbed ambient HF propagation.

REFERENCES

Basler, R. P., P. B. Bentley, G. H. Price, C. L. Rino, and D. K. Rust, "HF Channel Probe," Technical Report DNA-TR-85-247, Contract No. DNA-001-83-C-0325, SRI Project 6185, SRI International, Menlo Park, CA, 31 May 1985.

Basler, R. P., G. H. Price, R. T. Tsunoda, and T. L. Wong, "Ionospheric distortion of HF signals," *Radio Science*, Vol. 23, No. 4, pp. 569-579, July-August 1988.

Basler, R. P., P. B. Bentley, G. H. Price, R. T. Tsunoda, and T. L. Wong, "Ionospheric Distortion of HF Signals," DNA-TR-87-246, SRI International, Menlo Park, CA, September 1987.

Knepp, D. L., "Multiple Phase-Screen Calculation of the Temporal Behavior of Stochastic Waves," *Proc. IEEE*, Vol. 71, No. 6, pp. 722-737, June 1983(a).

Knepp, D. L., "Analytic Solution for the Two-Frequency Mutual Coherence Function for Spherical Wave Propagation," *Radio Science*, Vol. 18, No. 4, pp. 535-549, July-August 1983(b).

Nickisch, L. J., "Stochastic Effects on Oblique HF Propagation," DNA-TR-86-74 (MRC-R-974), Mission Research Corporation, Santa Barbara, CA, 1986.

Nickisch, L. J. and G. D. McCartor, "Stochastic Effects on Oblique HF Propagation II," DNA-TR-87-160 (MRC-R-1074), Mission Research Corporation, Monterey, CA, 1987.

Nickisch, L. J., "Non-uniform motion and extended media effects on the mutual coherence function: An analytic solution for spaced frequency, position, and time," submitted to *Radio Science*, 1990.

Price, G. H., "Striation Effect on HF Signals," DNA-TR-85-250, Final Technical Report, Contract DNA-001-83-C-0413, SRI Project 6381, SRI International, Menlo Park, CA, June 1985.

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(a).

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(b).

Rino, C. L., V. H. Gonzalez, and A. R. Hessing, "Coherence bandwidth loss in transionospheric radio propagation," *Radio Science*, Vol. 16, No. 2, pp. 245-255, March-April 1981.

Rino, C. L., "On the application of phase screen models to the interpretation of ionospheric scintillation data," *Radio Science*, Vol. 17, No. 4, pp. 855-867, July-August 1982.

Tatarskii, V. I., Wave Propagation in a Turbulent Medium, Dover Publications, Inc., New York (1961).

Tatarskii, V. I., The Effects of the Turbulent Atmosphere on Wave Propagation, translated by the Israel Program for Scientific Translations, National Technical Information Service, U. S. Department of Commerce (1971).

Wittwer, L. A., "Radio Wave Propagation in Structured Ionization for Satellite Applications," DNA 5304D, Defense Nuclear Agency, Washington D.C., 1979.

Wittwer, L. A., "Radio Wave Propagation in Structured Ionization for Satellite Applications II," DNA-IR-82-02, Defense Nuclear Agency, Washington D.C., 1982.

Yeh, K. C. and C. H. Liu, "An investigation of temporal moments of stochastic waves," *Radio Science*, Vol. 12, No. 5, pp. 671-680, September-October 1977.

ACKNOWLEDGEMENTS

The research reported in this paper was supported by the Defense Nuclear Agency under contract DAAL02-86-D-0043. This paper is a condensation of *Nickisch* [1990].

AD-P006 289

THE IMPACT OF THE SOLAR CYCLE ON OVER-THE-HORIZON RADAR SYSTEMS

Jeffrey S. Schleher
BDM International, Inc.; Albuquerque, NM

INTRODUCTION

Skywave Over-the-horizon Backscatter radar systems (OTH-B RS) depend on ionospheric propagation conditions. These change with many different cycles; the diurnal cycle, the seasonal cycle and the solar cycle; and with the geographic location and design of the system. Testing the radars under conditions that represent the range of possible variations is difficult. An operational test of a half year or less, can obtain representative data for all but the solar cycle. Missing geographic ionospheric test conditions can be obtained by using data gathered in other locations and by using historical data. Testing system performance over the solar cycle can be accomplished by modeling and simulation. Several models have been developed to accomplish this task. The drawback to this approach is that modeling requires forecasts for the next solar cycle which the scientific community has great difficulty providing.

This paper reviews the impact of solar cycle ionospheric changes have on OTH-B RS that operate in the mid-latitude ionosphere, and view the polar and the equatorial ionosphere. These effects are related to the solar cycle changes in solar flux and geomagnetic conditions. The lack of success of the solar-terrestrial physics community in forecasting the next solar cycle and the accompanying geomagnetic disturbances is considered. Finally, an approach to circumvent the need for a forecast of the solar flux magnitude and timing in OTH-B RS performance modeling is suggested.

THE IONOSPHERE

The Earth's ionosphere is primarily created by electromagnetic radiation from the sun. Because of atmospheric chemistry and density, different altitudes are affected by different frequencies of radiation. Figure 1 shows the number of free electrons that are created by this radiation during solar minimum and maximum. The OTH-B RS depends on the electron density, the height of the ionosphere and the take off angle of the radar beam to determine the radar range that ionospheric refraction can support. As the numbers of electrons increases with the solar cycle the OTH-B RS operators have an easier time finding frequencies to reach the desired range. At solar minimum, reaching the longer radar ranges is difficult. During solar maximum the D-region also is stronger causing increased absorption of the lower frequency energy during the day time. The increase in daytime absorption is not nearly as great as the increase of available frequencies, providing a wider band of available frequencies at solar maximum. (Ivanov-Kholodny & Mikhailov, 1986)

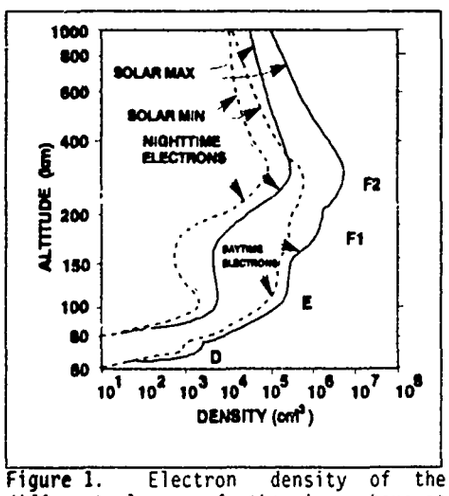


Figure 1. Electron density of the different layers of the ionosphere at solar maximum and minimum.

The increase in available frequencies at solar maximum is not all good. Other users of the HF spectrum also have much improved propagation conditions, as does naturally generated HF noise. The result is that OTH-B RS can "hear" more noise at solar maximum decreasing the signal to noise ratio and increasing the probability that there may be

91-09713

fewer available channels during solar maximum because of the non-interference requirement of the U.S. OTH-B RS.

GEOMAGNETIC DISTURBANCES

The degree of disturbance in the Earth's magnetic field also affects the way the ionosphere is able to support operating frequencies. Figure 2 shows the impact of a geomagnetic disturbance at mid-latitudes. Note that strong geomagnetic field variations will restrict substantially the available frequencies particularly during the critical sunrise transition period. (Townsend et. al., 1982)

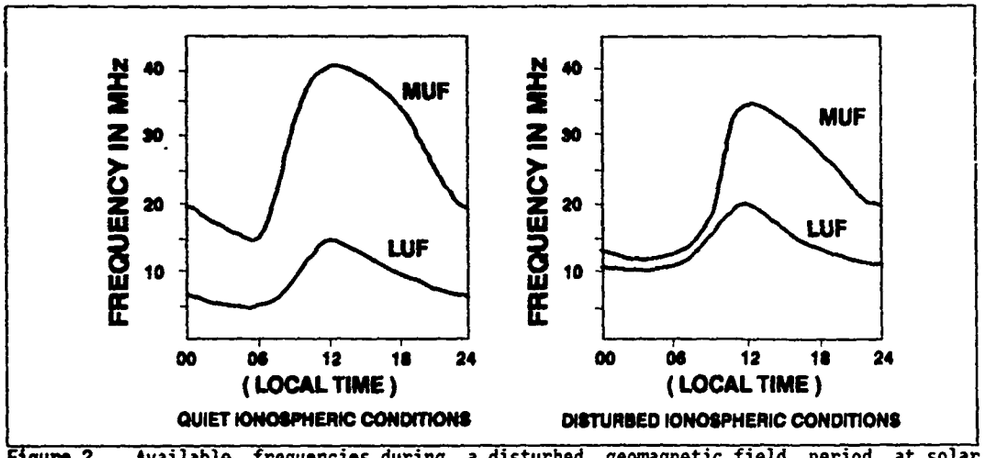


Figure 2. Available frequencies during a disturbed geomagnetic field period at solar maximum and solar minimum on a mid-latitude communications circuit.

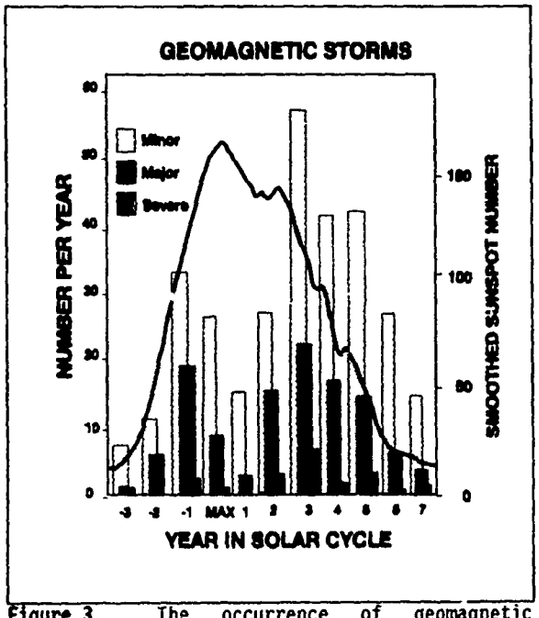


Figure 3. The occurrence of geomagnetic disturbances through out the solar cycle.

These Geomagnetic disturbances tend to occur on the down side of the solar cycle as shown in figure 3. The events of 1989, however, show that a few geomagnetic disturbances of record size can occur on the rising side, figure 4. At solar maximum, large and frequent geomagnetic disturbances generally result from solar activity. The position of the Earth in relation to the solar position of the event controls how strong the resulting geomagnetic disturbance will be.

The occurrence and strength of geomagnetic disturbance control the location of the auroral zone with regard to the ionospheric reflection points. As the level of disturbance increases the auroral zone moves south. OTH-B RS have difficulty operating in this zone, because of absorption and noise. Geomagnetic disturbances pushed the auroral zone down as far as the middle of Air Force OTH-B East Coast (ECRS) and Alaskan radar systems coverage during the first part of 1989. Figure 5 shows the southerly penetration of the auroral zone into the OTH-B ECRS coverage. This figure is based on the Q index (Feldstein, 1967) and is derived from the visible aurora.

Geomagnetic disturbances have three critical effects on the coverage of the OTH-B RS. First, as the aurora zone moves south any sectors with ionospheric refraction points in or near the zone will have reduced frequency availability and in general will be unavailable. The exception is when auroral sporadic-E supports coverage. This occurs consistently only on north looking paths during strong events. Second, to the south of the auroral zone the available frequencies will be restricted with increased noise. In addition, the number of clear- free channels will decrease due to competition with other users who are also affected. Under these conditions multipath tracks will increase as will azimuthal location errors. The latter is the result of electron density tilts along mid-latitude paths. Finally, when geomagnetic activity is low OTH-B RS looking into equatorial zones will experience increased noise as a result of equatorial disturbances.

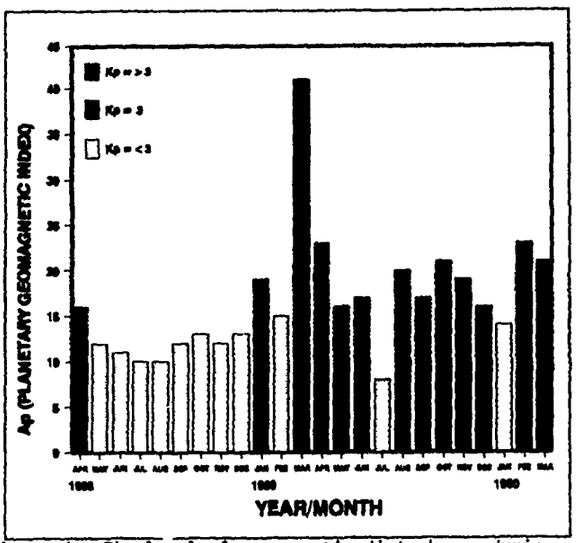


Figure 4. The level of geomagnetic disturbance during 1988-1989 as measured by the Air Force (SESC) A_p geomagnetic index.

EPISODIC DISTURBANCES

Many short term geophysical effects have dramatic effects on OTH-B RS. The occurrence of these effects are also solar cycle related. Major effects are sudden ionospheric disturbances (SID), sporadic E (E_s), Spread-F, Equatorial noise and Travelling Ionospheric Disturbances (TID). (Trizna & Headrick, 1981)

The most dramatic episodic event is the solar flare. Solar flares were originally observed as very bright optical events on the surface of the sun. While dramatic, the important aspect of these events to the OTH-B RS is their X-ray output, which raises the lowest usable frequency on the sunlit paths. This sudden ionospheric disturbance (SID) ensures that at the peak of a large flare, the lowest usable frequency is above the HF band. These peaks last for a few minutes to an hour. Figure 6 shows the relationship of the of solar flare occurrence to the solar cycle while figure 7 shows the relationship of X-ray events to the solar cycle. While the occurrence of optical flares track the solar cycle the large X-ray flares clearly reach a maximum on the down side of the cycle.

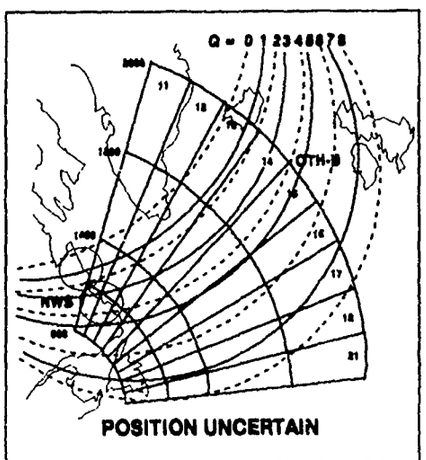


Figure 5. The position of the auroral zone for different levels of geomagnetic disturbances indicated by the index Q. The position for the higher numbers are uncertain.

Sporadic E is a dense layer of electrons in the E-region between 100-120 km, which does not seem to be related to normal E-region layers. E_s is a serious problem for OTH-B RS because the strength, location, and time of occurrence change rapidly. In addition, the virtual height of the layer is independent of frequency and can be partially penetrated by the beam. E_s may cover so small a area that only a part of the beam is reflected. The net effect of E_s is to increase clutter and intermittently provide real targets at the wrong ranges. E_s occurrence follows different rules in mid-latitudes than it does in auroral latitudes.

Auroral E_s is related to the occurrence of geomagnetic disturbances and therefor the solar cycle. Mid-latitude E_s is less associated with the solar cycle, as the operating frequencies rise the impact of mid-latitude E_s on OTH-B RS decreases.

Spread-F is an ionospheric phenomena that manifests itself on vertical ionosondes as energy returned from many altitudes and frequencies, rather than the altitude band that normally refracts radar energy at a given frequency. On OTH-B radar mid-latitude paths, this phenomena is observed as increased noise because the radar energy is reaching the ocean, and any target, by many paths. When spread-F is occurring these radars are generating there own noise. Some spread-F also creates an increase in Doppler spread.

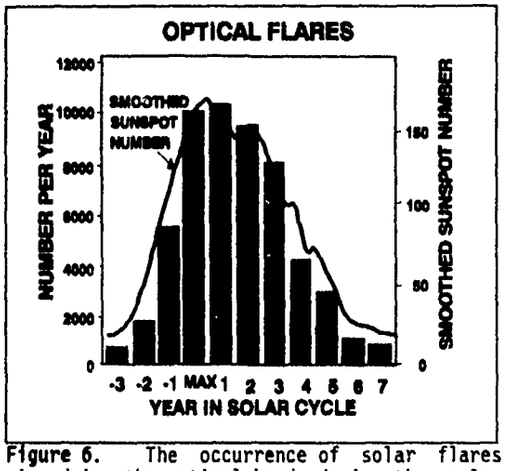


Figure 6. The occurrence of solar flares viewed in the optical band during the solar cycle versus the sun spot number.

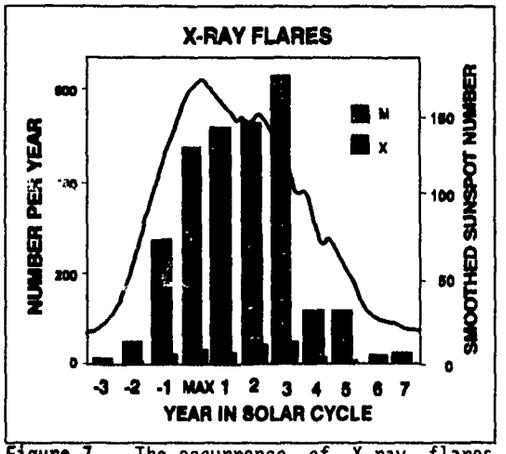


Figure 7. The occurrence of X-ray flares during the solar cycle versus the sun spot number.

In the mid-latitudes, Spread-F tends to decrease with increasing sun spot numbers, and is not very strong at any time. At equatorial latitudes spread-F increases with increasing solar flux, but decreases with disturbances in the geomagnetic field. Strong Spread-F in the equatorial zones is maximum in the hours after local sunset. At solar maximum the increase in the propagating range of the Air Force OTH-B ECRS reaches the equatorial zones. At this time the Spread-F becomes a strong self generated noise source. This manifestation of equitorial noise is observed on the oblique ionosondes as diffuse energy returns at 4000-5000 miles. It is probable this effect will be noted on other OTH-B RS viewing the equatorial regions.

Traveling ionospheric disturbances (TIDs) result from gravity waves in the upper atmosphere. As these waves move past an ionospheric reflection point the electron density decreases and then increases at the reflection height causing radar range errors with a duration of tens of minutes. Electron density changes on the order of a dwell period and spreading of the Doppler from the Earth's surface and any targets will cause a significant increase in Doppler signal noise.

TIDs occur more frequently at solar maximum and virtually all TIDs with large changes in electron density occur at solar maximum. TIDs have been noted on the Jindalee OTH-B Radar. (Anderson and Lees, 1987) The significance of this effect for OTH-B RS is that this noise source frequently occurs on mid-latitude paths during solar maximum and is often overlooked by the operator.

IMPLICATIONS FOR TESTING OTH-B RADAR SYSTEMS

The changes in continuous and episodic ionospheric events with 11 year solar cycle variations in solar flux and geomagnetic activities pose serious problems for operational OTH-B RS testing. Clearly if a reasonable test period is used it is impossible to test an OTH-B radar in the variety of ionospheric conditions that are expected over the life of the system. Testing under these conditions is necessary because of the demonstrated importance of the

ionosphere to OTH-B RS. The solution is to use an ionospheric model that can simulate the radar performance under the un-testable conditions. Several of these models have been developed (Elkins, 1988 and Szuszczewicz, 1988). The major problem with the models is that they both require an input of solar flux and geomagnetic activity index, and the quality of long term forecasting of these parameters is not good.

SOLAR CYCLE FORECASTING

The forecasting of the critical solar and geomagnetic indices has been improving during the latest solar cycle, but have a long way to go before forecasts within 10% of the actual magnitude of the solar maximum are realized. Figures 8 and 9 show the distribution of collected forecasts for solar cycles 21 and 22 with the cycle 21 maximum and the current cycle 22 observation for comparison. (Brown 1984) Figure 9 also contains forecasts that post date the Meudon workshop, from which the basic figure was taken. No forecast of solar cycle 23 were offered at the 1989 Solar Terrestrial Predictions Workshop.

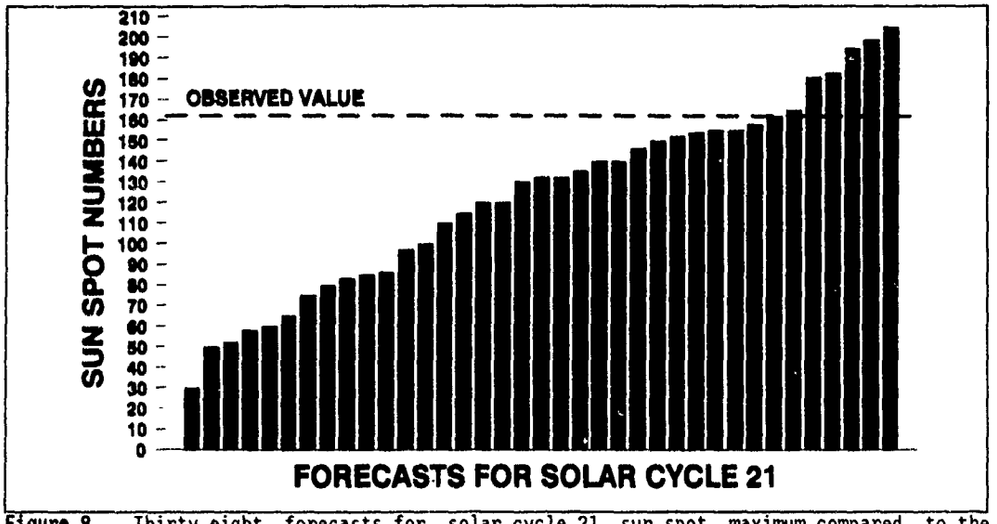


Figure 8. Thirty eight forecasts for solar cycle 21 sun spot maximum compared to the observed maximum.

These two figures represent forecasts made prior to or in the first few months of 1984. A dynamo theory prediction for Solar Cycle 22 (Schatten, 1984) was provided at the same meeting, and is presented in figure 10. This prediction for the declining sun spot number of solar cycle 21 and the forecast for cycle 22, is updated using a current Space Environmental Laboratory forecast. (SEL, 1990) The monthly average observations to date are included for comparison.

GEOMAGNETIC DISTURBANCE FORECASTING

Geomagnetic disturbances are also hard to forecast over the solar cycle. These disturbances are separated into two categories, sudden commencement and gradually commencing disturbances. Sudden commencements storms are related to the shock waves associated with solar flares. On the average stronger solar flares result in stronger geomagnetic disturbance events. The correlation between sun spot numbers and this type of geomagnetic disturbance is .85. (Mayaud, 1975)

Gradually commencing geomagnetic disturbances tend to reoccur with the 27 day rotation period of the sun. These events are the result of interaction between the solar wind and the geomagnetic field. The events can be very strong and comprise the bulk of all geomagnetic disturbances. Gradual commencement disturbances reach a maximum during the declining portion of the sun spot cycle. There appears to be a slow increase in the number and magnitude of the gradual commencing disturbances. Substantial errors occur in the prediction these

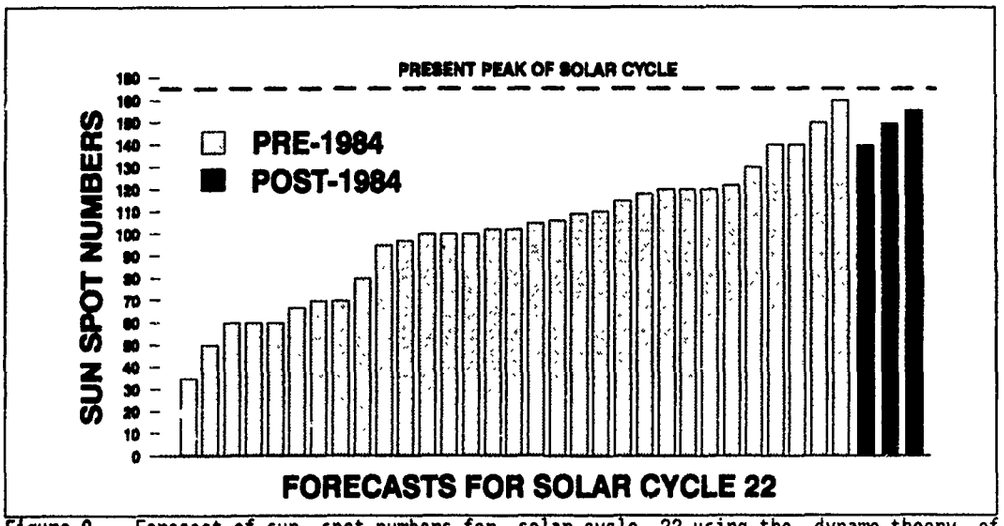


Figure 9. Forecast of sun spot numbers for solar cycle 22 using the dynamo theory of prediction, with the current smoothed monthly value. The first 31 were presented at the Meudon Workshop while the last three were made in 1984, 1985 and 1987 respectively.

disturbances over a solar cycle. Climatological geomagnetic disturbance forecasting is considerably less certain than sun spot forecasting as shown by the data for 1980 when a smoothed yearly sun spot number of 154 would normally indicate strong disturbances. This year, however, was one of the quietest geomagnetic periods on record. (Feynman and Gu, 1986) This dip in geomagnetic indices also occurred during several earlier cycles.

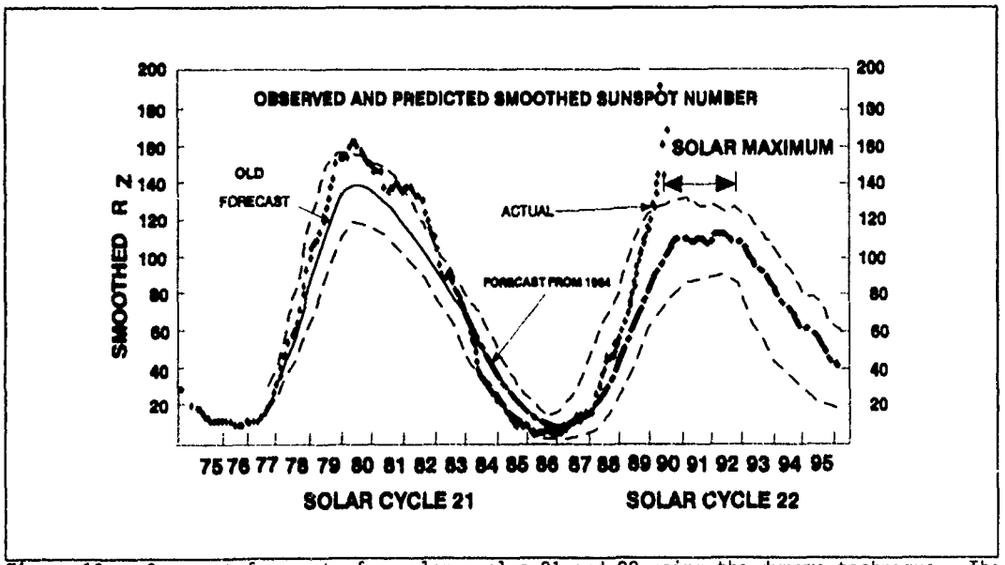


Figure 10. Sun spot forecasts for solar cycles 21 and 22 using the dynamo technique. The actual observations and the 20% confidence limits are also provided.

The forecast of the frequency of occurrence of geomagnetic disturbances in general show a bimodal distribution. The first peak is near solar maximum and is dominated by the sudden commencement disturbances associated with flares. The second peak occurs well into the declining phase and is dominated by the gradual commencement, reoccurring storms. The larger

peak appears to alternate with the odd numbered cycles the peak near solar maximum is larger. The reverse appears true for even numbered cycles. (Hedeman and Dodson-Prince 1984) The leveling off of the observed increase in the number of disturbances during the last two solar cycles lead to a predicted decrease in the number of disturbances in cycle 22. This certainly is not true to date.

FORECASTING SOLAR FLUX VALUES FOR OTH-B RADAR SYSTEM TESTING

This paper has demonstrated both that testing OTH-B radar systems requires consideration of the solar cycle and that forecasting this cycle is very difficult. An alternative approach takes a statistical view of the solar flux values during a solar cycle and applies Monte Carlo techniques to determine which flux values to use for a particular run of the model. This approach has the advantage that no forecast of the next solar cycle flux peak and timing is required. Any error in the peak value of the flux for cycle 23 will be minimized by the relatively short duration of these high numbers. Figure 11 shows a frequency distribution derived from the daily sunspot values from the last five solar cycles. The choice of five cycles is derived from the recognition that solar cycles in this half of the century have been increasing in strength, indicating a distribution containing many lower values are not desired. A number of recent papers that show the long term history of solar activity is similar to the 400 years of recorded data support this conservative approach (Withbroe, 1989)

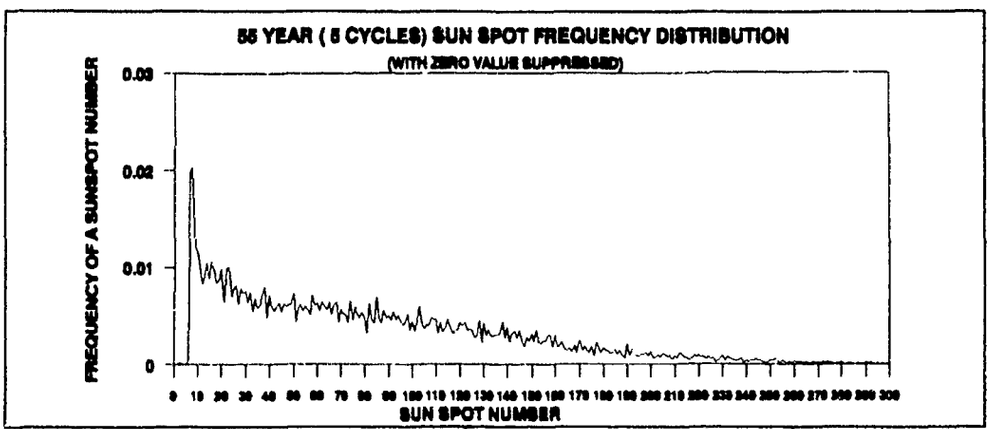


Figure 11. The frequency of occurrence of sun spot numbers during a solar cycle. These values were computed from the last five solar cycles. Sun Spot number zero has a frequency of 6%.

SUMMARY

This paper has shown that ionospheric variation is a critical factor in the operation of OTH-B RS. The range of flux values over the solar cycle causes great variation in the ionosphere. Ionospheric modeling is required for the testing of an OTH-B RS. The major problem in applying these models lies in forecasting the solar cycle. The paper recommends a sun spot frequency of occurrence distribution as a solution to this problem and provides one suggested distribution.

REFERENCES

Anderson S. J. and M.L. Lees, 1987; " High-Resolution Synoptic Scale Measurements of Ionospheric Motions with the Jindalee Radar" in The Effect of the Ionosphere on Communication, Navigation, and Surveillance Systems, ed Goodman, Naval Research Lab, p169-176.

Brown, G.N. 1984; Working Group "A" Report: Long Term Solar Predictions, Solar-Terrestrial Predictions Workshop at Meudon, FR; pp1-7.

Elkins, T.J. 1988; OTH Radar Performance Estimation, MITRE Working Paper, WP-27739.

Feldstein, Y.I. and G.V. Starkov, 1967; "Dynamics of Auroral Belt and Polar Geomagnetic Disturbances," Planet. Space Sci., 15. p209.

Feynman and Gu, 1986; Prediction of Geomagnetic Activity on the Scale of One to Ten Years, Rev. Geophy., 24, 3, pp 650-666.

Hedeman and Dodson-Prince, 1984; Indications of Probable Levels of Sunspot Numbers and Geomagnetic Disturbances in Solar Cycle 22, Solar-Terrestrial Predictions Workshop at Meudon, FR; pp97-98.

Ivanov-Kholdny, G.S. and A. V. Mikhailov, 1986; The Prediction of Ionospheric Conditions, D. Reidel Pub. Co., Boston, pp95-97.

Mayaud, P.N.,1975; Analysis of Storm Sudden Commencement for the Years 1868-1967, JGR, 80, 111.

Schatten, 1984; Solar-Terrestrial Predictions Workshop at Meudon, FR; Jun 84, pp92-96.

Space Environmental Laboratory, 1990 ; Preliminary Report and Forecast of Solar Geophysical Data, U.S. Department of Commerce.

Szuszczewicz, 1988; OTH-B-PROP: A COMPUTER MODEL FOR THE TEST AND EVALUATION OF OTH-B PROPAGATION CHARACTERISTICS IN THE GLOBAL-SCALE IONOSPHERE, a briefing to The Air Force Operational Test and Evaluation Center, 9 August 1988, SAIC, McLean, VA 22102.

Townsend,R. E., R. W. Cannata, R. D. Prochaska, G.E. Rattray, and J.C. Holbrook; 1981, Source Book of Solar-Geophysical Environment, AFGWC/TN-82/002

Trizna, D. B. and J. M. Headrick, 1981, "Ionospheric Effects HF Over-the-horizon Radar", in Effects of the Ionosphere on Radio Wave Systems, ed. Goodman, Naval Research Laboratory, pp262-272.

Withbroe, 1989, Solar Activity Cycle: History and Predictions, J. Spacecraft and Rockets, v26#6, pp.394-402.

AD-P006 290

91-09712

VHF/UHF RADAR SCINTILLATION EFFECTS OBSERVED BY ALTAIR

Dennis L. Knepp and Harry L. F. Houpis
Mission Research Corporation
Carmel, California

ABSTRACT

During the Defense Nuclear Agency (DNA) PEAK (Propagation Effects Assessment - Kwajalein) experiment in August 1988, the ALTAIR VHF/UHF wide bandwidth radar was used to track spherical satellites in low-earth orbit. The purpose of the experiment was to obtain radar data during the most severe propagation disturbances available naturally. The PEAK experiment has been quite successful, giving many measurements of strong scintillation as well as the first measurements of frequency-selective fading on propagating radar pulses. This high level of scintillation severity is statistically similar to that likely to be encountered over large spatial and temporal extent if the ionosphere were disturbed by a high-altitude nuclear detonation.

This paper describes the measurements that were obtained during the track of several spherical U.S. and Soviet calibration satellites. The use of spherical satellites eliminates the problem of target glint that causes fluctuation in the received signal due to interference from multiple target scatterers. Unique measurements were obtained during PEAK of severe frequency-selective scintillation during several passes of Cosmos 1427, a 2-meter diameter sphere.

In this paper the experimental results are used to demonstrate an enhancement in the average received power that is observed during severe scintillation. The observed statistics of the enhancement are compared to analytic calculations using the Nakagami-*m* distribution with very good agreement. This previously unreported enhancement is predictable on the basis of the first-order amplitude statistics for two-way radar propagation in a monostatic propagation geometry as well as by a more thorough analysis using reciprocity; both analytic approaches are presented here. The enhancement is important for both ground and space-based radars that have to operate during scintillation, since as much as a 3 decibel increase in the target (and clutter) signal-to-noise ratio is possible, depending on the severity of the scintillation and on the radar transmit/receive geometry.

INTRODUCTION

It is well-known that a propagating signal in the frequency range from VHF to L-band can experience disturbances due to naturally occurring F-region ionization. Disturbances due to mean or average ionization consisting of dispersion, refraction, Faraday rotation, and phase shift have been the subject of research work over the last several decades and are well understood [*Lawrence, et al.*, 1964]. Disturbances due to smaller scale electron-density irregularities that produce scintillation of propagating waves are the subject of current research [*Basu, et al.*, 1987].

Scintillation consists of the rapid variation of signal phase, amplitude, angle-of-arrival, time-of-arrival and other signal properties. Intense signal scintillation is often observed over satellite links and in radar measurements [*Towle*, 1980] through the ambient ionosphere at VHF and UHF. Strong equatorial scintillation is even observed at frequencies as high as C-band, but rarely [*Franke, et al.*, 1984]. The regions of scintillation consist of a 20° wide belt of latitude centered around the geomagnetic equator and the polar regions above 55° – 60° latitude. Within the equatorial scintillation belt, Rayleigh scintillation is often experienced at frequencies in the VHF to UHF range during the late evening and early morning hours from about 2000-0200 local time.

The DNA PEAK experiment was developed with the major goal of using the severely disturbed ionospheric

91

propagation channel to measure and record worst-case propagation effects. The primary experiment utilized ALTAIR, a high-performance long range VHF/UHF tracking radar. The ALTAIR VHF signal is of prime interest since the lower frequency experiences the greatest degree of scintillation. The goal of the DNA program is to measure radar signal propagation effects, determine the irregularity structure, measure the radar performance degradation, and then to relate the performance degradation to the measured propagation disturbances and to the *in-situ* irregularity structure. Once these relationships are established and verified from the experimental data, work can be directed towards the development of mitigation techniques to improve radar performance during signal disturbances.

ALTAIR WAVEFORM DESCRIPTION

ALTAIR utilizes a circular antenna with a diameter of 150 feet. The 3 dB beamwidths are 2.8 degrees at VHF and 1.1 degrees at UHF. In the PEAK experiment, independent range tracks were maintained at both frequencies, but the UHF angle information was used to steer the antenna. For the satellite tracking portion of the DNA PEAK experiment, the VHF and UHF waveforms chosen corresponded to the maximum bandwidth available. This selection gives an LFM waveform with a chirping bandwidth of 7.1 MHz at VHF and 18 MHz at UHF but limits the maximum pulse repetition frequency to 372 Hz. The actual PRF is determined by the radar control system to avoid eclipsing of pulses (i.e., to avoid reception of a pulse when the radar is transmitting).

The VHF range sample (or range gate) spacing is 30 meters, corresponding to a temporal sampling interval of 20 μsec. However, the minimum Nyquist rate for a 7 MHz chirp corresponds to a maximum sampling interval of about 14 μsec. Thus the recorded dechirped pulses are somewhat undersampled, which limits our measurements of pulse distortion. At UHF the pulses are undersampled by a greater proportion (15 m spacing), but since no pulse distortion due to scattering was observed this is of no concern.

The following signal parameters have been found to be important in the determination of the effects of fading on radar and communications system performance.

Scintillation Index (S_4). The level of amplitude scintillation is often measured by the S_4 scintillation index defined as the normalized standard deviation of the RCS obtained from a constant-RCS target.

$$S_4^2 = \frac{\langle P^2 \rangle - \langle P \rangle^2}{\langle P \rangle^2} \tag{1}$$

where P is the apparent radar RCS (the fluctuations in P are due to scintillation) and the angle brackets denote a stochastic average. Values of S_4 on a one-way path generally range from a minimum of zero signifying no scintillation to a maximum of unity indicating worst-case Rayleigh fading where the quadrature components of the received (one-way) signal are uncorrelated Gaussian variates.

With the ALTAIR data collected during PEAK it is possible to measure the signal-to-noise ratio (SNR) and correct the measured S_4 for noise effects using the expression

$$S_4^2(\text{scint}) = S_4^2(\text{measured})\left(1 + \frac{1}{SNR}\right)^2 - \frac{2}{SNR} - \frac{1}{SNR^2} \tag{2}$$

This expression is obtained by calculating the value of S_4 due to a random desired signal in the presence of white Gaussian noise. To measure a value of S_4, 1000 or 1500 values of sampled RCS measurements (corresponding to about 3 seconds of data during the satellite pass) were averaged.

Signal Decorrelation Time (τ_0). The signal decorrelation time, τ_0 is an appropriate quantity to describe the fading rate (the fading rate is $1/\tau_0$) of the received signal. The quantity τ_0 is defined by the expression

$$\langle E(t+\tau)E^*(t)\rangle = |E_0|^2 \exp\{-\tau^2/\tau_0^2\} \tag{3}$$

which is the correlation function of the complex received signal voltage. The actual value of τ_0 is a function of radar geometry and of the irregularity structure and intensity of the disturbed ionospheric channel. Large values of τ_0 correspond to slow fading conditions and small values correspond to fast fading. Both types of fading can degrade radar performance.

To measure τ_0 from the recorded I and Q quadrature data, the signal phase variations due to target motion must first be removed. That is, range variations and temporal variation in the gross total electron content (TEC) encountered as the line-of-sight traverses the ionosphere give deterministic phase variations not associated with scintillation.

Fortunately, simultaneous ALTAIR VHF and UHF pulses are designed for accurate TEC measurement through the use of range measurements at two frequencies. Once geometric range and TEC are measured pulse-by-pulse, they are smoothed in 10-second segments through polynomial fitting, and the phase changes due to range and TEC are separately removed from the recorded quadrature data. The resulting fluctuations in the complex received signal are due only to propagation effects, and τ_0 is easily measured through the use of FFT techniques. To obtain τ_0, the value of time lag corresponding to the $1/e$ point of the magnitude of the autocorrelation function is utilized.

Channel Coherence Bandwidth (f_0). The channel coherence bandwidth is a measure of the maximum bandwidth available in the propagation channel over which it is possible to transmit a signal without undesired pulse distortion. If the signal spectral components are separated in frequency by an amount greater than the coherence bandwidth, different spectral components will undergo uncorrelated fading. This distortion in the received signal spectrum causes the received time domain signal to display undesired time sidelobes. The terminology "frequency-selective fading" is synonymous with the increased importance of multipath wherein the signal propagates over many transmission paths, interfering coherently with less-timely versions of itself. The value of f_0 is quite important to many aspects of radar design.

The equation for the rms time-delay spread is the defining expression for f_0

$$\sigma_\tau^2 = \frac{\sum P(\tau_i)\tau_i^2}{\sum P(\tau_i)} - \langle\tau\rangle^2, \text{ where } \langle\tau\rangle = \sum P(\tau_i)\tau_i / \sum P(\tau_i) \tag{4}$$

In this equation $P(\tau_i)$ is the sampled received signal power at the output of the matched filter. Now the received signal can be written as the convolution of three components, that due to the dechirped signal itself (with no ionospheric effects); a second component due to the dispersive spreading caused by mean ionization along the propagation path; and a third component due to the frequency-selective scintillation. It is easy to show that

$$\sigma_\tau^2 = \sigma_{\text{pulse}}^2 + \sigma_{\text{tec}}^2 + \sigma_{\text{scint}}^2 \tag{5}$$

The quantity f_0 is defined as $1/2\pi\sigma_{\text{scint}}$; it represents only that portion of the signal spread due to scintillation or multipath.

During the PEAK experiment, $P(\tau_i)$ was recorded at VHF using range-gate samples separated by 30 meters, corresponding to a temporal spacing of 20μ sec. As discussed above, this temporal spacing is greater than the Nyquist interval; consequently measurements of moments of the pulse must be treated with caution.

In our measurements of the pulse spread, the recorded pulse shapes were collected into groups of 100, then the average pulse shape was obtained by noncoherent summation. The moments of this average pulse were obtained using Equation 4. We have not yet removed the σ_{tec}^2 term, so our current measurements of f_0 include the effects of dispersion.

ALTAIR MULTIPATH DATA

Figures 1-2 show a sequence of plots of the recorded received VHF pulse shapes (power versus range) from Cosmos 1427. Cosmos 1427 is a 2-meter diameter calibration sphere in a near polar orbit at an altitude of about 420 km. Figure 1, taken from satellite pass PEAK 4H, is an example of 50 sequential pulses recorded during a near-overhead pass with resultant high SNR (about 35 dB) during the period shown. Each pulse shape shows received power (plotted in dB above a milliwatt) versus range-gate sample. The first pulse at ALTAIR TAG time (time measured from the start of track) of 670.03 seconds is plotted at the bottom of the figure; subsequent pulses are offset above and to the right of the first.

Figure 2 gives an example of the most severely disturbed radar pulse data recorded during ionospheric fading conditions. This example shows 50 sequential pulses recorded using Cosmos 1427 during PEAK 9C over a period of about 140 msec starting at ALTAIR TAG time 157.2 msec. The temporal effect of fading is evident in the rather smooth variation of the peak power with time. The severe effects of ionospheric multipath are evident through the appearance of multiple peaks in the individual pulse shapes. Such pulse distortion will degrade many aspects of radar performance including detection, tracking, imaging, etc.

Figure 3 gives measurements of S_4, τ_0, and f_0 at VHF as a function of time during PEAK 9C. It is seen that saturated scintillation (S_4 (two-way)=$\sqrt{5}$) is experienced for much of the time duration between 135 and 220 seconds. Minimum values of τ_0 are around 10 msec; minimum values of f_0 are around 0.5 MHz.

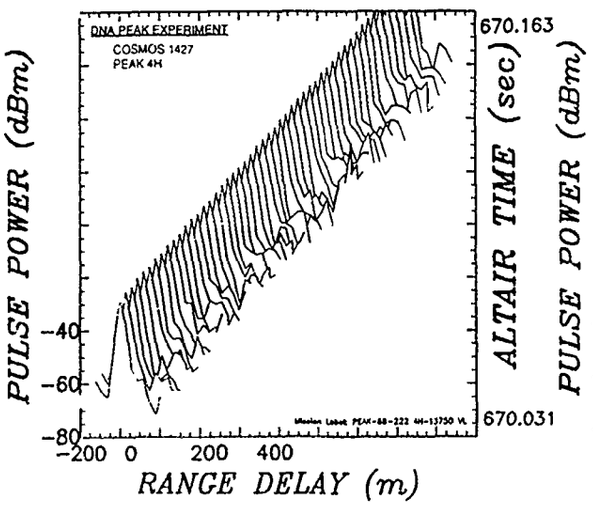


Figure 1. ALTAIR VHF pulse shapes during undisturbed conditions.

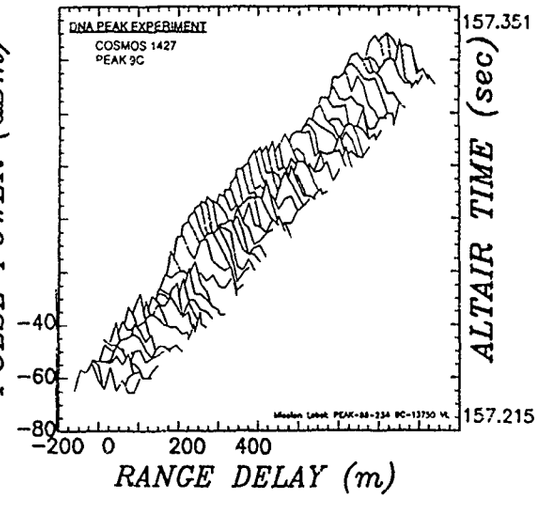


Figure 2. ALTAIR VHF pulse shapes during severe multipath.

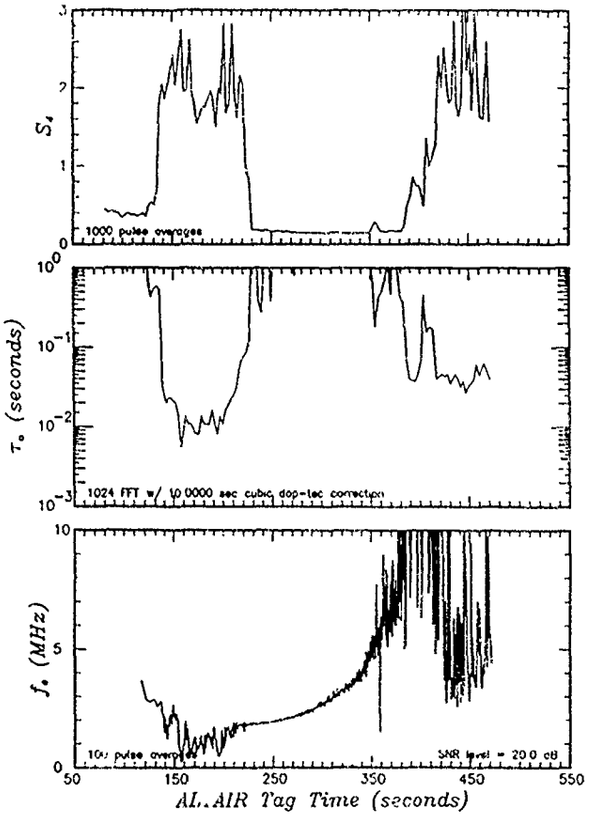


Figure 3. Statistical signal parameters measured during severe scintillation.

Note that the measurements of S_4 exceed in several cases the maximum value $\sqrt{5} = 2.24$, valid for saturated scintillation. This is the result of focusing caused by large scale irregularities "close to" the radar location. Equivalently, the Fresnel zone size is larger than the freezing scales of the irregularities.

THEORY OF POWER ENHANCEMENT

In the following a simple expression is obtained for the two-frequency, two-position, two-time mutual coherence function (MCF) for the case of two-way propagation from a radar antenna to a point target and back to the antenna. The result is given in terms of products of the one-way MCFs for propagation from the radar target to the radar transmit and receive locations and is valid for saturated scintillation where the quadrature components of the signal, after one-way propagation, are uncorrelated Gaussian variates. This result explains the enhancement as caused by correlation of the wave that propagates from the radar location to the target with the wave that propagates from the target to the radar.

Consider the problem illustrated in Figure 4 of a wave transmitted at time t at a center radian frequency of ω with time dependence $\exp(i\omega t)$ from an aperture antenna located in the ρ_x-plane at $\bar{\rho}_T$. *Fante* [1985], gives the received field using the Huygens-Fresnel Principle under two important assumptions. First, the transmitted signal must be narrow-band about the center radian frequency ω, and second, the time rate of change of the index-of-refraction fluctuations and of the transmitted temporal envelope must be slow in comparison to the radian frequency ω. These conditions are generally satisfied in modern pulsed radars. In this case the field received after propagation from the antenna to a location in the plane of the scattering layer, $\bar{\rho}_\xi$ is

$$E_{\rm inc}\left(\bar{\rho}_\xi, \omega, t + \frac{Z_t}{c}\right) = \int d\rho_x^2 A(\bar{\rho}_x - \bar{\rho}_T) G_1(\bar{\rho}_x \rightarrow \bar{\rho}_\xi; \omega; t) E_0(\bar{\rho}_x, \omega) \tag{6}$$

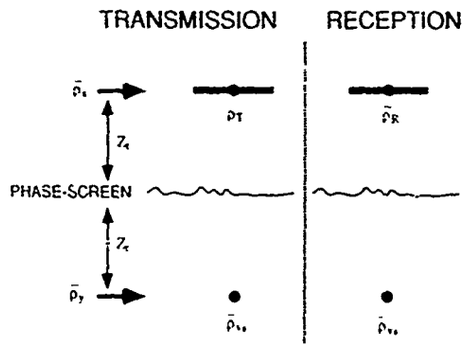


Figure 4. Two-way scattering geometry.

From the figure Z_t is the distance along the direction of propagation from the transmitter plane to a scattering layer represented here as a thin phase-screen, although the thin phase-screen assumption is not required. The quantity Z_r is the distance from the scattering layer to the target location, $E_0(\bar{\rho}_x, \omega)$ is the frequency-domain representation of the signal transmitted from location $\bar{\rho}_x$ and $A(\bar{\rho}_x - \bar{\rho}_T)$ is the aperture weighting function for the transmit antenna located at position $\bar{\rho}_T$. For example, the choice $A(\bar{\rho}) =$ constant for $|\bar{\rho}| \leq \rho_0$ and zero otherwise is appropriate for a circular antenna. The time dependence t in the propagator G_1, given by Equation 7, signifies that the transmission occurs at time t. The signal arrives at the scattering layer at a later time, $t + Z_t/c$ as included in the expression for $E_{\rm inc}$. The propagator is given as

$$G_1(\bar{\rho}_x \rightarrow \bar{\rho}_\xi; \omega; t) = \frac{k}{i2\pi Z_t} \exp\left\{-ikZ_t - \frac{ik}{2Z_t}(\bar{\rho}_\xi - \bar{\rho}_x)^2\right\} \tag{7}$$

Now the signal in the target plane may be written by propagating in the same manner from the scattering layer down to the target. The result can be expressed as

$$E_{\rm tar}(\bar{\rho}_y, \omega, t + \tau) = \int d\rho_\xi^2 E_{\rm inc}\left(\bar{\rho}_\xi, \omega, t - \frac{Z_t}{c}\right) G_2\left(\bar{\rho}_\xi \rightarrow \bar{\rho}_y; \omega, t + \frac{Z_t}{c}\right) \tag{8}$$

where the propagator, G_2 accounts for propagation of a wave starting from the layer at time $t + Z_t/c$, and includes the phase term ϕ which gives the scattering layer contribution to the phase in the Huygens-Fresnel principle.

$$G_2\left(\vec{\rho}_\xi \to \vec{\rho}_y; \omega; t + \frac{Z_t}{c}\right) = \frac{k}{i2\pi Z_r} \exp\left\{-ikZ_r - \frac{ik}{2Z_r}(\vec{\rho}_\xi - \vec{\rho}_y)^2\right\} \exp\left\{i\phi\left(\vec{\rho}_\xi, \omega, t + \frac{Z_t}{c}\right)\right\} \quad (9)$$

For simplification, the above three equations can be combined as

$$E_{\text{tar}}(\vec{\rho}_y, \omega, t + \tau) = \int d\rho_x^2 A(\vec{\rho}_x - \vec{\rho}_T) G_{\text{down}}(\vec{\rho}_x \to \vec{\rho}_y; \omega; t) E_0(\vec{\rho}_x, \omega) \quad (10)$$

which represents downward propagation of a wave from the ρ_x-plane starting at time t, to the ρ_y-plane or target plane, arriving at time $t + \tau$ where $\tau = (Z_t + Z_r)/c$. The quantity G_{down} is a time-variable (varying with t) transfer function for a sinusoid of radian frequency ω.

Now let the target be a point scatterer with unity reflectivity and located at $\vec{\rho}_{y_0}$. Then the field received at time $t + 2\tau$ after upward propagation from the target back to the radar location is given as

$$E_{\text{rcv}}(\vec{\rho}_R, \omega, t + 2\tau) = \int d\rho_x^2 B(\vec{\rho}_x - \vec{\rho}_R) G_{\text{up}}(\vec{\rho}_{y_0} \to \vec{\rho}_x; \omega; t + \tau) E_{\text{tar}}(\vec{\rho}_{y_0}, \omega, t + \tau) \quad (11)$$

Equations 10-11 can be combined to yield a simplified expression

$$\begin{aligned} E_{\text{rcv}}(\vec{\rho}_R, \omega, t + 2\tau) &= \int d\rho_x'^2 d\rho_x''^2 A(\vec{\rho}_x' - \vec{\rho}_T) B(\vec{\rho}_x'' - \vec{\rho}_R) \\ &\quad G_{\text{up}}(\vec{\rho}_{y_0} \to \vec{\rho}_x''; \omega; t + \tau) G_{\text{down}}(\vec{\rho}_x' \to \vec{\rho}_{y_0}; \omega; t)\ E_0(\vec{\rho}_x', \omega) \end{aligned} \quad (12)$$

where the quantity $B(\vec{\rho}_x'' - \vec{\rho}_R)$ specifies the aperture weighting function for the receive antenna which is located with its phase-center at $\vec{\rho}_R$ at time $t + 2\tau$. Now simplify by letting both transmit and receive apertures be point antennas so $A(\vec{\rho}_x' - \vec{\rho}_T) = \delta(\vec{\rho}_x' - \vec{\rho}_T)$ and similarly for B so that the expression above for the received field after two-way propagation contains no explicit integrals. Then the expression for the desired autocorrelation function after two-way propagation is obtained as

$$\begin{aligned} \Gamma_2(\vec{\rho}_{R1} - \vec{\rho}_{R2}; \omega_1 - \omega_2; \Delta) &= \langle E_{\text{rcv}}(\vec{\rho}_{R1}, \omega_1, t + 2\tau + \Delta) E_{\text{rcv}}^*(\vec{\rho}_{R2}, \omega_2, t + 2\tau) \rangle \\ &= \langle G_{\text{up}}(\vec{\rho}_{y_0} \to \vec{\rho}_{R1}; \omega_1; t + \tau + \Delta) G_{\text{down}}(\vec{\rho}_{T1} \to \vec{\rho}_{y_0}; \omega_1; t + \Delta) \\ &\quad G_{\text{up}}^*(\vec{\rho}_{y_0} \to \vec{\rho}_{R2}; \omega_2; t + \tau) G_{\text{down}}^*(\vec{\rho}_{T2} \to \vec{\rho}_{y_0}; \omega_2; t) \rangle\ E_0(\vec{\rho}_{T1}, \omega_1) E_0^*(\vec{\rho}_{T2}, \omega_2) \end{aligned} \quad (13)$$

where the angle brackets denote stochastic average. $E_{\text{rcv}}(\vec{\rho}_{R1})$ is the voltage received at $\vec{\rho}_{R1}$, due to a transmitter at $\vec{\rho}_{T1}$; $E_{\text{rcv}}(\vec{\rho}_{R2})$ is the voltage received at $\vec{\rho}_{R2}$ due to a transmitter at $\vec{\rho}_{T2}$. To proceed further, reciprocity must be invoked to relate upward and down-going waves through the expression

$$G_{\text{down}}(\vec{\rho}_x \to \vec{\rho}_y; \omega; t) = G_{\text{up}}\left(\vec{\rho}_y \to \vec{\rho}_x; \omega; t + \frac{Z_t - Z_r}{c}\right) \quad (14)$$

The above equation states reciprocity as follows. In propagating downward, starting at time t, the wave arrives at the layer at time $t + Z_t/c$; in upward propagation starting at time $t + (Z_t - Z_r)/c$, the wave arrives at the screen at time $t + (Z_t - Z_r)/c + Z_r/c = t + Z_t/c$, i.e., the same time. Thus the scattering layer is in the same position for upward as it is for downward propagation, so reciprocity holds.

Application of reciprocity allows one to replace the downward-going transfer functions in Equation 13 with their equivalent upward-going transfer functions.

$$\begin{aligned} \Gamma_2 &= \left\langle G_{\text{up}}(\vec{\rho}_{y_0} \to \vec{\rho}_{R1}; \omega_1; t + \tau + \Delta) G_{\text{up}}\left(\vec{\rho}_{y_0} \to \vec{\rho}_{T1}; \omega_1; t + \Delta + \frac{Z_t - Z_r}{c}\right)\right. \\ &\quad \left. G_{\text{up}}^*(\vec{\rho}_{y_0} \to \vec{\rho}_{R2}; \omega_2; t + \tau) G_{\text{up}}^*\left(\vec{\rho}_{y_0} \to \vec{\rho}_{T2}; \omega_2; t + \frac{Z_t - Z_r}{c}\right)\right\rangle E_0(\vec{\rho}_{T1}, \omega_1) E_0^*(\vec{\rho}_{T2}, \omega_2) \end{aligned} \quad (15)$$

Now, if scintillation is saturated on a one-way propagation path so that G_{up} is a complex Gaussian variate, then one can write the fourth-order moment using the well-known expression $\langle x_1 x_2 x_3 x_4 \rangle = \langle x_1 x_2 \rangle \langle x_3 x_4 \rangle + \langle x_1 x_3 \rangle \langle x_2 x_4 \rangle + \langle x_1 x_4 \rangle \langle x_2 x_3 \rangle$.

In the application of this expression for the fourth-order moments, the term $\langle G_{\text{up}} G_{\text{up}} \rangle \langle G_{\text{up}}^* G_{\text{up}}^* \rangle$ is found to be of the form $(\langle Re^2 \rangle - \langle Im^2 \rangle)^2$ where the variances of real and imaginary parts are equal, giving zero. The other two terms are written in Equation 16 as the products of MCFs valid for the upward path (Γ_{up}) from the target to the various transmitter and receiver locations. Γ_{20} is the two-way MCF in the absence of a scattering layer.

$$\begin{aligned}\Gamma_2 &= \Gamma_{20}\{\Gamma_{up}(\bar{\rho}_{R1}-\bar{\rho}_{R2};\omega_1-\omega_2;\Delta)\Gamma_{up}(\bar{\rho}_{T1}-\bar{\rho}_{T2};\omega_1-\omega_2;\Delta)\\ &+ \Gamma_{up}\left(\bar{\rho}_{R1}-\bar{\rho}_{T2};\omega_1-\omega_2;\Delta-\frac{2Z_r}{c}\right)\Gamma_{up}\left(\bar{\rho}_{T1}-\bar{\rho}_{R2};\omega_1-\omega_2;\Delta+\frac{2Z_r}{c}\right)\}\\ &\quad E_0(\bar{\rho}_{T1},\omega_1)E_0^*(\bar{\rho}_{T2},\omega_2)\end{aligned} \tag{16}$$

The quantity $\bar{\rho}_{T1}$ is the transmitter location at time $t+\Delta$; $\bar{\rho}_{T2}$ is the transmitter location at time t; $\bar{\rho}_{R1}$ and $\bar{\rho}_{R2}$ are the receiver locations at times $t+2\tau+\Delta$ and $t+2\tau$, respectively.

The quantity $2Z_r/c$ is the time required for the wave to propagate from the scattering layer, to the target, and back to the scattering layer. This quantity then represents the temporal interval of importance to the correlation of down-going and up-going waves.

Expressions for Γ_{up} and Γ_{20} for this geometry are given in *Knepp* [1983; 1985] as

$$\Gamma_{up}(\bar{\rho},\omega_d,t) = (1+i\omega_d/\omega_{coh})^{-1}\exp\left\{-\sigma_\phi^2\omega_d^2/2\omega_0^2\right\}\exp\left\{\frac{-\rho^2/\ell_0^2}{1+i\omega_d/\omega_{coh}}\right\}\exp\left\{-t^2/\tau_c^2\right\} \tag{17}$$

$$\Gamma_{20} = \left(\frac{1}{Z_r+Z_t}\right)^4\exp\left\{\frac{-ik_1(\rho_{R1}^2+\rho_{T1}^2)+ik_2(\rho_{R2}^2+\rho_{T2}^2)}{2(Z_r+Z_t)}\right\} \tag{18}$$

where ω_{coh} is the coherence bandwidth, σ_ϕ is the phase-screen standard deviation, ℓ_0 is the signal decorrelation distance, and $\exp\{-t^2/\tau_c^2\}$ has been added to reflect temporal variation in the scattering layer. The quantities k_1 and k_2 are the wavenumbers at ω_1 and ω_2.

If transmitter and receiver are far away from one another, the second term in Equation 16 is zero, since each Γ_{up} is separately zero. If transmitter and receiver are collocated and $2Z_r/c$ is small, the second term is equal to the first term and thereby represents the enhancement. The enhancement depends on the relative positions of transmitter and receiver at two different times. It is therefore important in the evaluation of the performance of space-based radar, particularly displaced phase center array (DPCA) wherein the transmit and receive apertures have complicated relative motion.

For the case of no scattering, the resulting expression for the MCF is $\Gamma_2(\text{no scint}) = \Gamma_{20}E_0(\bar{\rho}_{T1},\omega_1)E_0^*(\bar{\rho}_{T2},\omega_2)$. For the case of a monostatic radar geometry with small values of the quantity $2Z_r/c$, so that $\Gamma_{up}=1$ in all cases in Equation 16, one obtains $\Gamma_2(\text{monostatic radar}) = 2\Gamma_2(\text{no scintillation})$ or a 3 dB increase in average received power above that measured with no scintillation.

Measurements of the Enhancement. Fairly straightforward analysis enables one to check several aspects of this power enhancement theory. At its simplest the enhancement can be derived using reciprocity which states that, if the propagation environment is unchanging, a signal with a single frequency component traverses the same path in traveling from the radar to the target and back. In that case the voltage after two-way propagation is proportional to the square of the one-way voltage. Under the assumption of Nakagami-m statistics for the one-way power, the probability density function (pdf) of received RCS is given by

$$p_1(S)dS = \frac{mS^{m-1}}{\Gamma(m)\langle S\rangle^m}\exp\left\{\frac{-mS}{\langle S\rangle}\right\}dS,\ S\geq 0,\ (\text{one}-\text{way}) \tag{19}$$

If reciprocity holds, the radar measures the two-way RCS, $Q=S^2$ which then has the pdf given by

$$p_2(Q)dQ = \frac{m^mQ^{m/2-1}}{2\Gamma(m)\langle S\rangle^m}\exp\left\{\frac{-m\sqrt{Q}}{\langle S\rangle}\right\}dQ,\ Q\geq 0,\ (\text{two}-\text{way}) \tag{20}$$

From Equation 20 the mean received two-way power can be computed in terms of the square of the one-way power as $\langle Q\rangle = \int_0^\infty p_2(Q)QdQ = \frac{(m+1)}{m}\langle S\rangle^2$. In the case of a constant signal with no fading, S_4 is zero, m is infinity and one obtains $\langle Q(\text{no scint})\rangle = \langle S\rangle^2$ which can be combined with the above expression to yield a relationship between average RCS and m.

$$\frac{\langle\text{RCS}\rangle}{\langle\text{RCS (no scint)}\rangle} = \frac{m+1}{m} \tag{21}$$

ALTAIR VHF Data Analysis. Figure 5 gives an example of ALTAIR UHF data in which the backscatter enhancement is quite evident. This data was obtained from the track of the Soviet sphere, Cosmos 1427, on the evening of August 21, 1988. This particular satellite pass is referred to as PEAK 9C, data from which is shown in Figures 1-2.

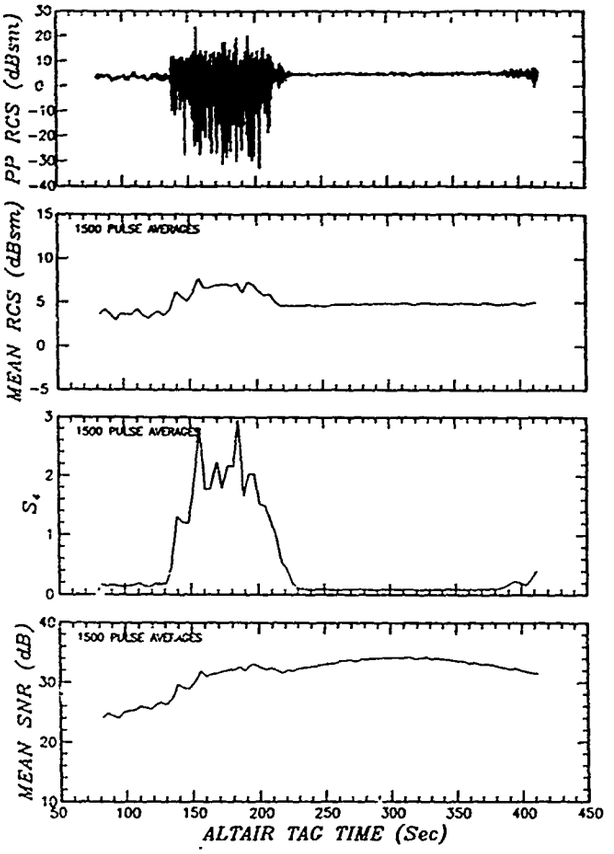


Figure 5. ALTAIR UHF data taken during PEAK 9C.

The top frame of Figure 5 shows the measured UHF radar cross section in decibels below a square meter as a function of time from the start of radar track. In this frame about one-tenth of the recorded data is shown, at a sample spacing of 25 msec. The radar data sampling interval varies but is generally around 2.7 msec. Notice the abrupt onset of a patch of severe scintillation starting at a time of 135 seconds and ending at 220 seconds. The elevation at 140 seconds is 7.5° and is increasing slowly until about 300 seconds when the satellite begins to descend. This high elevation angle rules out ground multipath as the source of the signal fluctuations.

The lower three frames in Figure 5 show computed quantities obtained by averaging the data obtained over 1500 consecutive received pulses. The second frame gives the average RCS. The average is obtained by summing 1500 values of power, dividing the sum by 1500, and then converting to decibels. The third frame gives the values of S_4 computed on each 1500 second segment of data. The bottom frame shows the average signal-to-noise ratio (SNR) computed by averaging the ratio of the power in the track gate to the power in the last recorded range gate. In the case of PEAK 9C, the track gate corresponds to sample number 4, the last recorded sample is number 9. Note that our SNR measurements could be corrupted by range sidelobes, but are not here since the range sidelobes are 35 dB below the peak and, therefore, below the noise, for the most part. Unfortunately, we do not have available for analysis the radar's actual noise power measurement.

The enhancement in received RCS during scintillation is clearly visible in the second frame of Figure 5, which gives the average RCS as a function of TAG time. If one uses the period between 250 to 300 seconds to measure the mean RCS without scintillation, increases in RCS due to scintillation approaching 3 dB are apparent in the period 140-210 seconds.

In Figure 6 the UHF measurements of S_4 and average RCS are combined in a format to check the enhancement theory given above. By using the ALTAIR data shown in Figure 5 it is possible to measure all the values in

Equation 21 to check this relationship. Take the average RCS during undisturbed propagation conditions as the average of the received RCS over the interval 250-300 seconds. Then consider the other measurements of average power and S_4 with (and without) scintillation during the time period 130-380 seconds. These measurements of S_4 are first corrected for noise using Equation 2 and then used to obtain the m-parameter through the expression $S_4^2(\text{two-way}) = (4m^2 + 10m + 6)/(m(m+1)^2)$ [*Knepp and Reinking*, 1989].

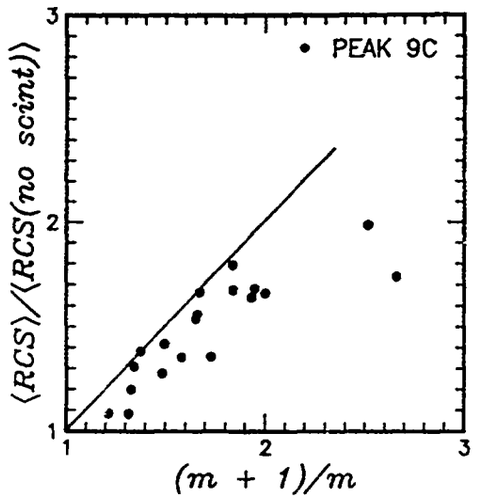


Figure 6. Comparison of ALTAIR measurements to theory, PEAK 9C.

The resulting data is then presented in the form of a scatter plot shown in Figure 6 where the ratio of the average RCS measurements with and without scintillation are plotted as a function of the ratio $(m+1)/m$. A straight line with unity slope corresponds to the enhancement theory described above. The agreement is quite good between measured data and theory for this single pass of Cosmos 1427 near ALTAIR.

Figures 7 and 8 show similar ALTAIR data at UHF and VHF, respectively, for three additional passes of Cosmos 1427 of (PEAK 2G, 7C, 6F) and a pass of a smaller U.S. calibration sphere, LCS-4 (PEAK 6G). The sphere LCS-4 is too small (1.13 m diameter) and in too high an orbit to have a sufficient SNR per pulse at VHF to be useful for this study.

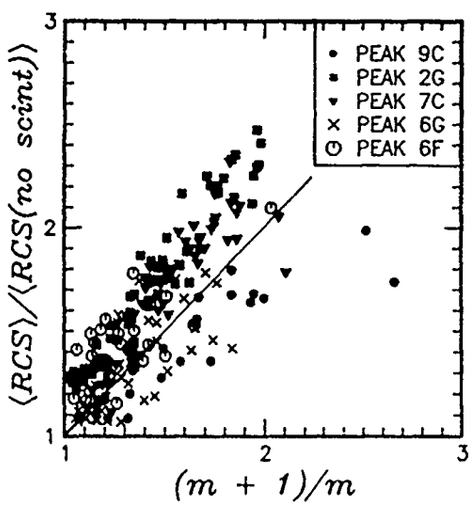


Figure 7. Comparison of ALTAIR UHF measurements to theory.

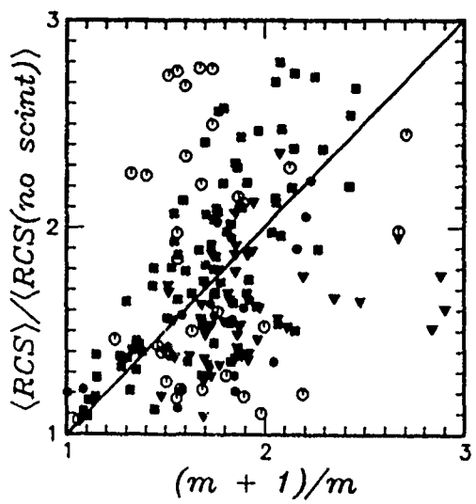


Figure 8. Comparison of ALTAIR VHF measurements to theory.

There is general agreement between the data collected in Figures 7 and 8 and the theory. The UHF data shows quite good agreement with the theory; the VHF data exhibits much more erratic behavior, most likely due to dispersive pulse spreading that causes loss in peak power not accounted for in the scattering theory, and possibly due to the reduced SNR available at VHF.

CONCLUSIONS

This paper has presented ALTAIR data illustrating the effects of severe ionospheric multipath on the received VHF wideband chirp waveform. The theory of a previously undiscovered backscatter enhancement due to scintillation is presented with convincing data as partial proof.

ACKNOWLEDGEMENT

The authors are indebted to Dr. Robert Stagat of Mission Research Corporation, Mr. Warren Dickinson of the U.S. Army Strategic Defense Command, and Dr. Leon Wittwer of the Defense Nuclear Agency for useful discussions regarding this work. The work was supported under DNA Contract No. DNA001-87-C-0169 and NRL Contract No. N00014-87-C-2336.

REFERENCES

Basu, S. E. MacKenzie, S. Basu, E. Costa, P. F. Fougere, H. C. Carlson and H. E. Whitney, "250 MHz/GHz Scintillation Parameters in the Equatorial, Polar and Auroral Environments," *IEEE Journal on Selected Areas in Communications*, Vol. SAC-5, No. 2, 1987.

Fante, R. L., "Wave Propagation in Random Media: A Systems Approach," in *Progress in Optics XXII*, E. Wolf, Ed., Elsevier, 1985.

Franke, S. J., C. H. Liu and D. J. Fang, "Multifrequency Study of Ionospheric Scintillation at Ascension Island," *Radio Science*, Vol. 19, pp. 695-706, May-June 1984.

Knepp, D. L and J. T. Reinking, "Ionospheric Environment and Effects on Space-based Radar Detection," in *Space-Based Radar Handbook*, L. J. Cantanfio, Ed., Artech House, 1989.

Knepp, D. L., "Aperture Antenna Effects After Propagation Through Strongly Disturbed Random Media," *IEEE Trans. Antennas and Propagation*, Vol. AP-33, No. 10, pp. 1074-1084, October 1985.

Knepp, D. L., "Analytic Solution for the Two-Frequency Mutual Coherence Function for Spherical Wave Propagation," *Radio Science*, Vol. 18, No. 4, pp. 535-549, July-August 1983.

Kumagai, H., "Spatial correlations in intense ionospheric scintillations: Comparison between numerical computation and observation," *Radio Science*, Vol. 22, No. 3, pp. 439-448, May-June 1987.

Lawrence, R. S., C. G. Little and H. J. A. Chivers, "A Survey of Ionospheric Effects Upon Earth-Space Radio Propagation," *Proceedings of the IEEE*, pp. 4-27, January 1964.

Towle, D M., "VHF and UHF Radar Observations of Equatorial *F* Region Ionospheric Irregularities and Background Densities," *Radio Science*, Vol. 15, No. 1, pp. 71-86, January-February 1980.

AD-P006 291

AN IONOSPHERIC PROPAGATION PREDICTION METHOD FOR LOW- AND MID-LATITUDES

Harry L. F. Houpis and L. J. Nickisch
Mission Research Corporation
Carmel, California

ABSTRACT

An ionospheric propagation prediction method is proposed that is more responsive than current methods to the actual propagation channel of a radio wave. These other methods typically use an "average" propagation channel. The range of conditions that comprise this "average", when considered separately, can potentially lead to quite different predictions of the propagation channel relative to the "average" alone. Thus, to the extent that a real-time trend in the propagation channel can be established or that correlations with other observables can be determined, better predictions can be obtained.

The proposed method obtains a real-time trend by employing a nonlinear regression analysis technique coupled together with an appropriately parameterized (empirical) ionosphere model and a ray-tracing algorithm. Parameter indices are then generated and a future trend is established. To test the effectiveness of this approach, measurements of group path length and elevation to several orbiting satellites from a dual-frequency (UHF/VHF) pulsed radar located near the magnetic equator are utilized. This data provides the needed statistics for the regression analysis.

The proposed method can also be used to obtain a better understanding of ionosphere-magnetosphere coupling, and in particular, to investigate the influence of geomagnetic activity on the low- and mid-latitude ionosphere. This is accomplished by comparing our method's parameter indices to solar and geomagnetic indices to see if correlations exist. To the extent that such a correlation does exist, a more accurate wave propagation prediction will result, which in turn will provide a more reliable radar tracking performance. Our present results indicate that a correlation exists between geomagnetic activity (as measured by the Kp three-hourly range index) and the parameter indices, with the correlation falling off as the equator is approached from 25° N geographic latitude.

INTRODUCTION

The terrestrial ionosphere is often defined as that part of the atmosphere where free electrons and ions are of sufficient concentration to affect the propagation direction and speed of radio waves. This definition has come from many decades of research aided with radars, ionosondes, rockets, artificial satellites, and the space shuttle. Theoretical studies and numerical simulations have also assisted. What has emerged in our understanding is an ionosphere with four major ionization zones, the D, E, F1, and F2 regions, which as a whole range in altitude from about 50 km to roughly 3000 km above the Earth's surface. At the lower altitudes, the ionosphere interacts through various collisional and photochemical processes with the neutral atmosphere or thermosphere, while at higher altitudes, the ionosphere meshes with a variety of magnetospheric regions such as the plasmasphere and the polar cusp. Consequently, with the variable solar corpuscular and radiative fluxes impinging upon the Earth's plasma environment and hence acting as a driver for the assorted physical and chemical processes, the ionosphere displays structures that are themselves highly variable in time and space. This variability in turn is manifested in the performance of telecommunication or radar systems which use signals that are propagated through the ionosphere (henceforth referred to as radio systems).

When evaluating the desirability of a propagation channel or attempting to improve the performance of a radio system by accounting for ionospheric effects, two distinct but interrelated sets of predictive equations are needed: one for the ionosphere and another for the transmitter/receiver system itself. The ionosphere propagation prediction takes the form of a method that can give as output such quantities as the critical frequencies of the ionospheric layers (designated as foD, foE, foF1, and foF2), the altitude of the layer's maximum electron density (hmD, hmE, hmF1, and hmF2), the maximum usable frequency (MUF), the lowest usable frequency (LUF), and the ray path of the radio wave. The method is also used to make predictions of the ionospheric structure. The transmitter/receiver system prediction then uses these results in conjunction with other criteria to evaluate the quality or desirability of a communication link or propagation channel. In this paper, we present preliminary evidence for the utility of improving ionospheric propagation prediction accuracy in real-time by employing a regression technique together with correlations of ionospheric properties with measurable parameters (e.g., geomagnetic indices).

91-09711

Various ionospheric propagation prediction methods (henceforth referred to as IPPMs) have been developed over the years with the emphasis placed on either long-term (usually defined for periods greater than 1 month) or short-term predictions. The model chosen is dependent on what is to be accomplished. For radio systems, where precise predictions are required, only the short-term approach has provided the desired accuracy. However, to date, even this accuracy is limited by the model input parameters (e.g., see Ivanov-Kholodny and Mikhailov, 1986). For example, typical input parameters are the solar radiative flux, $F_{10.7}$, and the geomagnetic activity index, K_p. In practice, the input of these time-dependent quantities does not follow their real-time values, but is updated at finite intervals (usually a day or so, although time periods as low as an hour may be obtained for a few parameters in some rare instances). Furthermore, some input parameters, such as $F_{10.7}$ and K_p, are global in extent and hence do not readily reflect local conditions. Consequently, both short-term and long-term IPPMs inevitably suffer temporal and spatial inaccuracies due to their uncertain input, regardless of how physically accurate the model itself is. Thus, these IPPMs are best suited to solar, geomagnetic and ionospheric activity levels that are quiet over an extended period of time. This is usually only accomplished at mid-latitudes. For other conditions, particularly when irregular transient activity exists, predictions become less accurate. However, for short-term predictions, our proposed IPPM will substantially retain its fidelity even under these conditions.

Inherent in any IPPM is the underlying ionosphere model itself. This model may assume any one of three forms: theoretical, statistical, or empirical (see review by Rush, 1986). A theoretical model reflects our current understanding of the physical and chemical processes involved in the evolution of the ionosphere. Presently, the more sophisticated models require complex numerical techniques and supercomputer computation speeds (e.g., see Schunk, 1987, and Fuller-Rowell, 1987). Statistical models, on the other hand, assume nothing about the physics or chemistry involved. Any such model is merely a statistical compendium of data for a predetermined span of space and (Universal) time, which can be given in either tabular or functional form, and is quite often dependent on certain input parameters, for example , season and solar activity (Davies, 1980, gives a summary of these models). Finally, empirical models combine various features of both the theoretical and statistical approaches (the International Reference Ionosphere is an well-known example). Invariably, a basic theoretical model is devised, being careful not to introduce any processes or equations which are either computationally slow or fairly uncertain in nature. The "missing" information is then introduced by a variety of statistical parameters in functional form (dependent on solar activity, etc.). Which of the three aforementioned model formats one uses naturally depends on the purpose and/or application at hand.

The ionosphere, of course, does not exist in isolation. Increasingly over the last decade, more attention is being directed at the fact that the thermosphere, the ionosphere, and the magnetosphere are dynamically linked; any model which isolates one of these regions by not coupling to the other two runs the risk of drawing erroneous conclusions when compared to data. A number of efforts focusing on this important problem are presently under way. Three major examples are: (1) The CEDAR (Coupling, Energetics, and Dynamics of Atmospheric Regions) program of the Atmospheric Sciences Division at the National Science Foundation is an international collaboration directed at understanding the global coupling of the thermosphere and ionosphere (Smith et al, 1989). (2) GEM (Geospace Environment Modeling), a recently formulated program at the National Science Foundation, deals primarily with the global structure of the magnetosphere, but includes campaigns pertaining to ionospheric coupling (Roederer, 1988). (3) SCOSTEP (Scientific Committee on Solar-Terrestrial Physics), through its STEP (Solar-Terrestrial Energy Program) Steering Committee, has set up five international working groups with their main emphasis being coupling processes (Rostoker, 1989; see also, Walt, 1988).

The method we offer seeks to apply these recognized connections (a brief discussion of this method was recently presented by Houpis and Nickisch, 1989). If successful, it will improve the capabilities of short-term IPPMs. In particular, the proposed method addresses the problem of the ionosphere model input parameter uncertainty by answering the basic question: "What is the most likely set of input parameters that will reproduce the observed ionospheric conditions?". Briefly, this is accomplished by using a standard technique in nonlinear regression analysis together with an appropriately parameterized (empirical) ionosphere model and a ray-tracing algorithm. (Data that is processed to test the method comes from a dual-frequency pulsed radar located near the equator). The "new" input parameters, determined by the regression analysis, may then be related to the "standard" input parameters as normally used in the ionosphere model. Subsequently, a trend in the "new" input parameters is established for future short-term prediction.

The relation between the "new" and "standard" parameters will also point to improvements in the underlying ionosphere model. For example, using a "reference" input parameter as the "zero level", the index of the corresponding "new" input parameter, calculated over some temporal and spatial interval, is generated (see Mayaud, 1980, for a complete discussion of index generation). This index may then be compared to solar and/or geomagnetic indices as well as *in situ* measurements of the interplanetary magnetic field (IMF). These comparisons may then point to an ionosphere-magnetosphere coupling that could be incorporated in the initial ionosphere model.

Modifying the ionosphere model in this manner serves two purposes. The primary purpose follows the recent emphasis in the space physics community (as discussed above) to better understand the coupling between the thermosphere, the ionosphere, and the magnetosphere. Naturally, this improved understanding should reflect in a better IPPM for radio systems. Hence, the other purpose is the improvement of radio system performance. In this paper, we give our initial progress towards meeting these two objectives.

RADAR SITE AND DATA

The data set in the possession of Mission Research Corporation was acquired from the ALTAIR radar (9° 23' 43" N, 167° 28' 45" E) on Roi-Namur Island of the Kwajalein Atoll in the Marshall Islands during the PEAK-88 (Propagation Effects

Assessment - Kwajalein) campaign held for 14 noncontiguous days in August, 1988 (see Knepp et al, 1988). The radar is a simultaneous dual-pulsed radar with center frequencies of 155.5 MHz (VHF) and 422.0 MHz (UHF).

The available radar data was obtained by tracking satellites of opportunity (66 separate tracks including those of HiLat, Polar Bear and San Marcos) with geographic latitudes ranging from -2.5° to 25° and with altitudes typically between 400 and 1000 km. Since the PEAK-88 campaign was specifically designed to maximize the probability of obtaining scintillated return signals, satellites were only tracked during a four-hour time span around midnight (local time). For a single satellite track in this time span, the number of pulses received by the radar is typically 350 per second over a time interval of about 500 seconds (a total of about 175,000 pulses for both UHF and VHF). Furthermore, for a given radar pulse "tag time" (UT is also available), the raw data obtained for both the UHF and VHF signals include: group path length (range), azimuth, elevation, and signal amplitude and relative phase.

METHOD DEVELOPMENT AND RADAR PROPAGATION PREDICTION

Given the measured group path length to a satellite, r_m (half the round trip travel time of a radar pulse multiplied by the speed of light, c), it is related to first order to the geometric path length, r_g (the actual distance from the radar to the satellite along the pulse ray path), and the total electron content along the geometric path (TEC), N_T, by the simple relation:

$$r_m = r_g + \frac{c^2 r_e}{2\pi f^2} N_T \tag{1}$$

where r_e is the classical electron radius (2.818×10^{-15} m) and f is the center frequency (in Hz) of the transmitted pulse. The standard method of solving for r_g and N_T is to use Equation (1) twice, once for the UHF pulse and then again for the simultaneous VHF pulse. It is implicitly assumed that the dispersion between the two pulses together with their different indices of refraction are small enough to allow r_g and N_T to be the same for both frequencies. This assumption is quite reasonable for most applications.

However, additional information can be discerned if we use the fact that the two ray paths, and hence r_g and N_T, are not exactly the same. This quite naturally leads to the use of a ray-tracing algorithm. Assuming that the ray path lies in the plane defined by the radar, satellite, and Earth sub-satellite point (i.e., we ignore ray path differences in azimuth), we have the following equations (Croft and Hoogasian, 1968; Hill, 1979):

$$r_m = \int_{r_o}^{r} \frac{r\,dr}{\sqrt{r^2\mu^2 - r_o^2\cos^2\beta_o}} \tag{2}$$

$$D = \int_{r_o}^{r} \frac{r_o^2 \cos\beta\,dr}{r\sqrt{r^2\mu^2 - r_o^2\cos^2\beta_o}} \tag{3}$$

$$r_g = \int_{r_o}^{r} \frac{r\mu\,dr}{\sqrt{r^2\mu^2 - r_o^2\cos^2\beta_o}} \tag{4}$$

where β_o is the elevation angle of the return signal at the radar (we note that refraction due to the neutral atmosphere is accounted for in the measured values of β_o and r_m by using the ALTAIR site thermospheric refraction model), r is the Earth-centered distance to the satellite, r_o is the radius of the Earth, D is the distance to the sub-satellite point measured along the Earth's surface from the radar, and μ is the index of refraction given by:

$$\mu^2 = 1 - 80.62 n_e / f^2 \tag{5}$$

and n_e is the electron density (in m^{-3}).

To "close" the set of Equations (2)-(5) requires two additional conditions. One is a model of the electron density (hence, an ionosphere model). The other is that the two ray paths converge at the satellite. This last condition is equivalent to the ground range for both frequencies being the same at the satellite Earth-centered distance, r:

$$D(r,\mathrm{UHF}) = D(r,\mathrm{VHF}) \tag{6}$$

Strictly speaking, the equations above are not closed; in fact, they are over-specified. Given the measured group path length and elevation for a particular frequency pulse, and an ionosphere model, with its standard input parameters which are fixed, Equation (2) will give a unique value of the satellite distance r regardless of the measured values for the other frequency pulse. Thus, it is extremely unlikely that the two ray paths will cross at the satellite. However, recognizing that measured data have errors associated with them and that the ionosphere model input parameters have a natural uncertainty about them as discussed in the Introduction, we allow both data and parameters to vary in a nonlinear regression analysis technique, forcing the ray paths to meet at the satellite.

More specifically, we use a nonlinear regression algorithm (NLREG) which employs any one of a number of variations of the standard Levenberg-Marquart method (e.g., see Press et al, 1987). Basically, this algorithm will fit, by least squares, a given function of a set of parameters and one or more independent variables to a set of data by repeated linear regressions. The data and parameters referred to here are not necessarily the same as those referred to above. We henceforth distinguish between the two by specifically referring to the former as regression data and parameters. The distinction is necessary since there is much leeway in choosing regression data and parameters.

Our experience using NLREG with a single quasi-parabolic ionospheric layer (which provides our electron density relation for Equation (5)) and the ALTAIR data has resulted in the break down of the regression data and parameters given in Table 1. The first two regression data are the actual measured data; the third regression data is set to zero and reflects our knowledge that the two actual ray paths must have intercepted the satellite; and the last regression data is calculated from the measured data and equation (1) used twice, and gives us an idea of the state of the ionosphere. We are implicitly assuming here that the total electron content experienced by the two different pulses along their different ray paths will both be within a standard deviation of this N_T (experience shows that this is a reasonable assumption). As for the regression parameters, f_c, h_m, and y_m (henceforth referred to as ionosphere model regression parameters (IMRPs)) are simply the ionosphere model parameters, while A is used in our theoretical calculation of D. β_o fits more naturally as a regression parameter than as a regression data point since there is no theoretical value for it; we merely assume it is statistically the same as the measured value. Figure 1 shows the results of using NLREG for every 2000th pulse pair (approximately every 5.7 seconds) along a 450 second satellite track. Note the correlation between the model layer parameters and the measured (at every 5th pulse) total electron content.

Table 1

Regression Data	Regression Parameters
r_m(UHF)	β_o(UHF)
r_m(VHF)	β_o(VHF)
D(UHF) - D(VHF)	Layer Critical Frequency, f_c
N_T	Altitude of Maximum Electron Density, h_m
	Layer Semi-Thickness, y_m
	Satellite Altitude, A

The usefulness of the above method to ionospheric propagation prediction is clear. If for example a quasi-parabolic model is being used to assess a particular propagation path, f_c, h_m, and y_m are input with "typical" values. However, increased accuracy can be obtained if the received data itself can provide these input parameters through the nonlinear regression technique described above. Short-term predictions may then be accomplished by extrapolating these derived input parameters into the near future. We point out that NLREG is extremely fast for this quasi-parabolic case, accomplishing its task in a CPU time far less than the real-time needed to acquire the data itself. Thus, predictions can always keep pace with a real-time schedule.

Besides a quasi-parabolic model, which merely represents an equivalent ionosphere by its f_c, h_m, and y_m parameters, we have also investigated a more realistic ionosphere model (preliminary findings were given by Houpis and Nickisch, 1989). In particular, we have used a modified version of FAIM (Fully Analytic Ionosphere Model) developed by Anderson et al (1989). FAIM is an empirical model, based on earlier models by Chiu (1975) and Anderson et al (1987), that is computationally fast for short-term predictions, and is constructed on a well-founded basis in physics. The model is geared primarily to low- and middle-latitudes (which is compatible with our ALTAIR data) and provides as output the electron density as a function of a variety of input parameters (e.g., local and annual time, geographic latitude, altitude, monthly smoothed Zurich sunspot number, and geomagnetic latitude, longitude, and dip angle). FAIM's only drawback is that it may underestimate the electron density for altitudes greater than about 500-600 km for very high sun spot numbers. Under these conditions, we use a simple power law drop-off rather than the exponential decrease of FAIM. We will not present here our analysis using the FAIM model. We merely point out that certain input parameters, such as the Zurich sunspot number, can be treated as an IMRP and thus provide a more accurate propagation prediction. Details of our results will be presented in a subsequent paper.

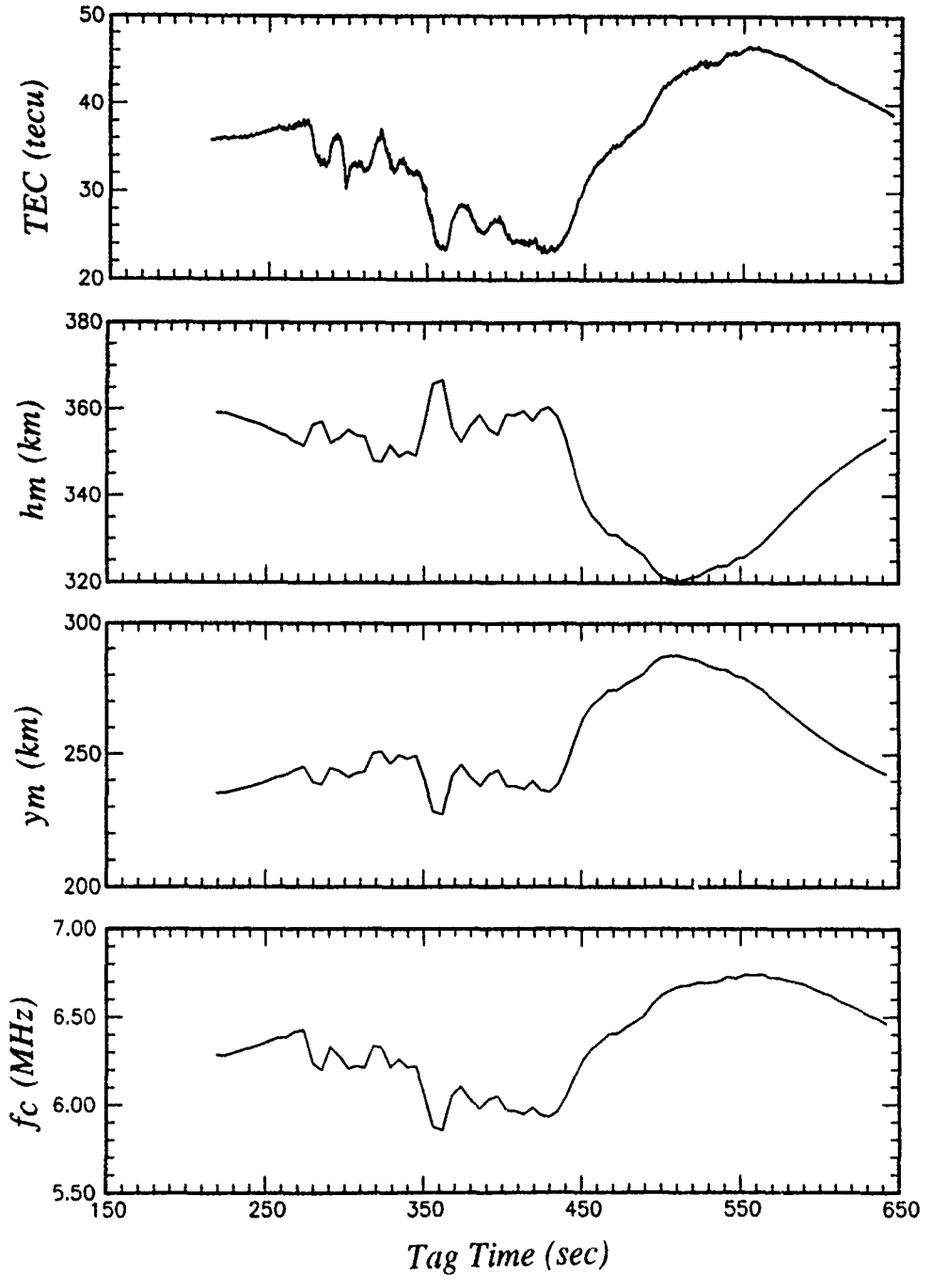


FIGURE 1 - DERIVED IPPM INPUT PARAMETERS AND MEASURED TEC (IN UNITS OF 10^{16} M^{-2}) VS ALTAIR TAG TIME FOR COSMOS SATELLITE 1427 ON AUGUST 7,1988 (PEAK-88 TRACK 2G-13750).

IONOSPHERE-MAGNETOSPHERE COUPLING

As discussed in the Introduction, the unsteady solar corpuscular and radiative fluxes impinging upon the Earth's plasma environment gives the ionosphere its characteristic variability. How exactly this is accomplished has been the subject of research for a number of years. In particular, the question of neutral wind and electrodynamic coupling between the polar ionosphere and the low-latitude ionosphere, especially during the transient phases of a magnetic storm, is a problem of current interest. In fact, there is mounting evidence that an electrodynamic as well as a neutral dynamic and chemical coupling exists For instance, it is widely acknowledged that during the early phases of a magnetic storm, electric fields may penetrate to low latitudes due to the region 2 current sheet no longer providing adequate shielding. Traveling ionospheric disturbances (TIDs) may also be launched from the polar region during these times and their effects can be seen at low-latitudes. (See Forbes, 1989, for a brief discussion of the evidence for these processes; a more detailed discussion can be found in Kelley, 1989). With this evidence as incentive, we have undertaken an investigation of the ALTAIR radar data set to see if an ionosphere-magnetosphere coupling effect can be detected.

To accomplish this task, we have taken the following approach. A subset of the ALTAIR radar tracks is subjected to the nonlinear regression analysis technique described in the last section; in the case presented here, 13 tracks were chosen from 9 separate nights. The resulting IMRPs are then binned according to latitude and the day in August, 1988. We do not bin with respect to longitude or local time, but leave this to a future investigation. For a given IMRP, we calculate its "theoretic" index for each latitude-day bin. This theoretic index is derived from the IMRP in a manner similar to geomagnetic indices: A "zero-level" is defined for all the latitude-day bins. The theoretic index of a given latitude-day bin is then the "deviation" of the IMRP from the zero-level. For geomagnetic indices, the deviation is often defined as the maximum magnitude difference between the zero-level and the measured quantity. For our investigation, we have chosen as the deviation, the difference between the average value of the IMRP in the latitude-day bin and the zero-level. This definition of the theoretic index should be more responsive to large-scale ionospheric changes as opposed to small-scale fluctuations. In fact, the definition we have adopted here is very similar to that used by Forbes (1989) to detect a correlation between the Kp three-hourly index and the equatorward penetration of electric fields and winds.

Figures 2-4 give the results for the IMRPs f_c, h_m, and y_m, respectively, by utilizing every 2000th pulse pair along a given satellite trajectory. The latitude-day bins referred to above are represented by the latitude designation and the abscissa of each figure panel, while the left ordinate axis gives the theoretic index value (referred to as Δ_{fc}, Δ_{hm}, or Δ_{ym}). The latitude bins are 5° wide and are centered about the displayed number. An asterisk locates the calculated IMRP index with a maximum of 9 values on each panel; less than 9 values indicates that a certain number of satellite tracks did not intercept the panel's latitude bin. The zero-level used for the f_c, h_m, and y_m indices are 6.0 MHz, 375 km, and 250 km, respectively. (Presently, these levels are rather ad hoc; a more appropriate level would be tied to geomagnetically quiet conditions, but this is not easily defined for all the panels). The August, 1988, days are given in UT with August 1 at 12 noon defined as 1.5. The solid curve in each panel represents the Kp three-hourly index with the right ordinate giving its maximum span from 0 to 9.

From the Kp variation shown in Figures 2-4, it is evident that August, 1988, was largely characterized by quiet to moderate geomagnetic activity. Nevertheless, assuming that the effects of a geomagnetic disturbance may require a time lag of as much as a few hours at low latitudes, a mild correlation appears to exist between the theoretic indices and the Kp index. In addition, the correlation appears stronger at higher latitudes. Overall, Δ_{fc} changes over a span of $\pm$ 2 MHz, Δ_{hm} over $\pm$ 125 km, and Δ_{ym} over $\pm$ 100 km, although these spans are somewhat smaller for any given latitude bin. Forbes (1989) finds a similar qualitative effect (see his Figures 1-3). We note, however, that Forbes' Δ_{hm} excursion is about one-half the value we obtain. This is easily understood as a difference between our use of a quasi-parabolic ionosphere model and Forbes' use of ionosonde data from a real exponential ionosphere (i.e., a quasi-parabolic layer must overcompensate in thickness relative to a Chapman-like layer to acquire the same TEC).

DISCUSSION AND CONCLUSION

Although the method discussed in this paper provides an improved means of predicting radio propagation channels and establishes a useful tool in the study of ionosphere-magnetosphere coupling, a few problems still need to be addressed. The most serious problem is the unpredictability of the nonlinear regression algorithm itself. We have found that under severely scintillated conditions or conditions where the satellite displays rapidly varying radar glint, NLREG is unable to settle into a unique solution. Perhaps preprocessing of the radar data to filter out the scintillation and glint will solve this problem. However, the question of uniqueness will still remain. We are presently looking into other regression schemes.

Of equal concern is a more quantitative assessment of the utility of our method. For instance, the position of a particular satellite in space as a function of time is a direct result of our method. A separate determination of this functional relationship, either from an accurate ephemeris or a simultaneous high frequency (L band and above) truth track, should provide an accurate appraisal of the method's predictive capabilities. In a similar vein, the input parameter extrapolation technique discussed in the text above can be applied and then compared to the coincident IMRPs to determine how much time the extrapolation can be expected to remain accurate. Finally, the establishment of a theoretic/geomagnetic index correlation can best be checked under severe geomagnetic activity (e.g., during the highly active March, 1989, time frame). A strong correlation under such conditions lends more credence to the mild correlation that we have qualitatively claimed in this paper. All these matters are the subjects of a future study.

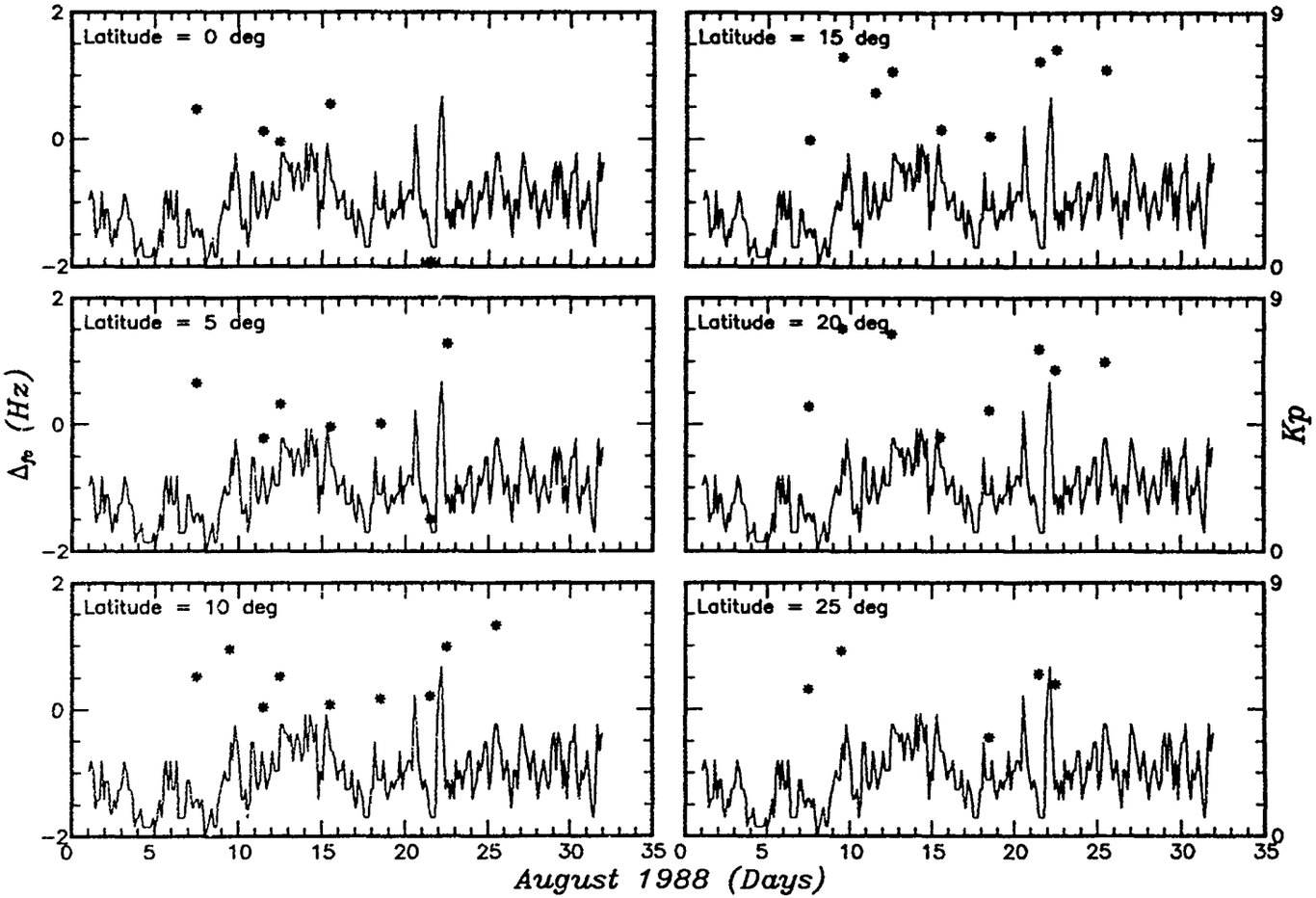


FIGURE 2 - CRITICAL FREQUENCY INDEX (ASTERISKS) AND THREE-HOURLY KP GEOMAGNETIC INDEX (SOLID CURVE) FOR DAYS IN AUGUST 1988

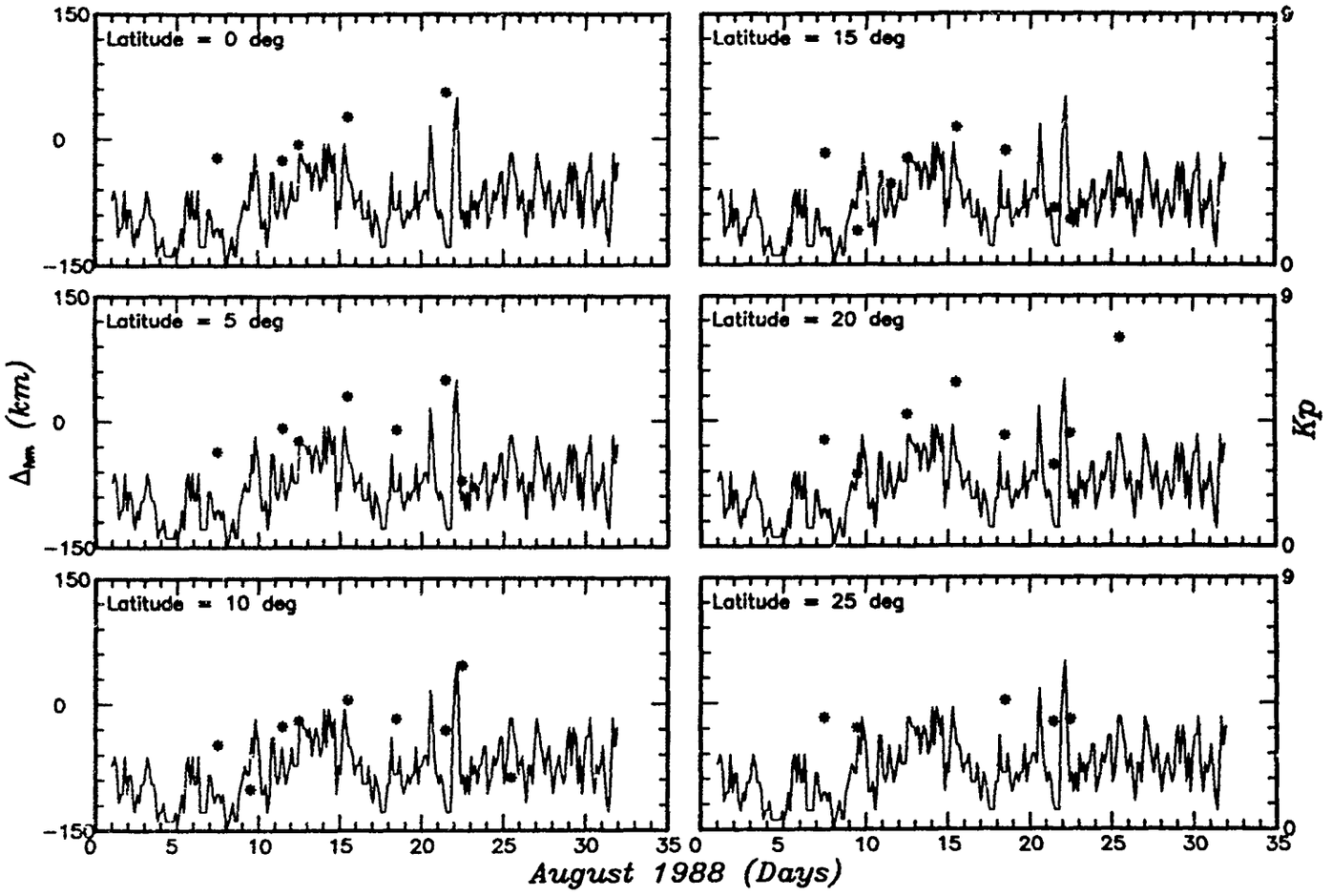


FIGURE 3 - LAYER HEIGHT INDEX (ASTERISKS) AND THREE-HOURLY KP GEOMAGNETIC INDEX (SOLID CURVE) FOR DAYS IN AUGUST 1988

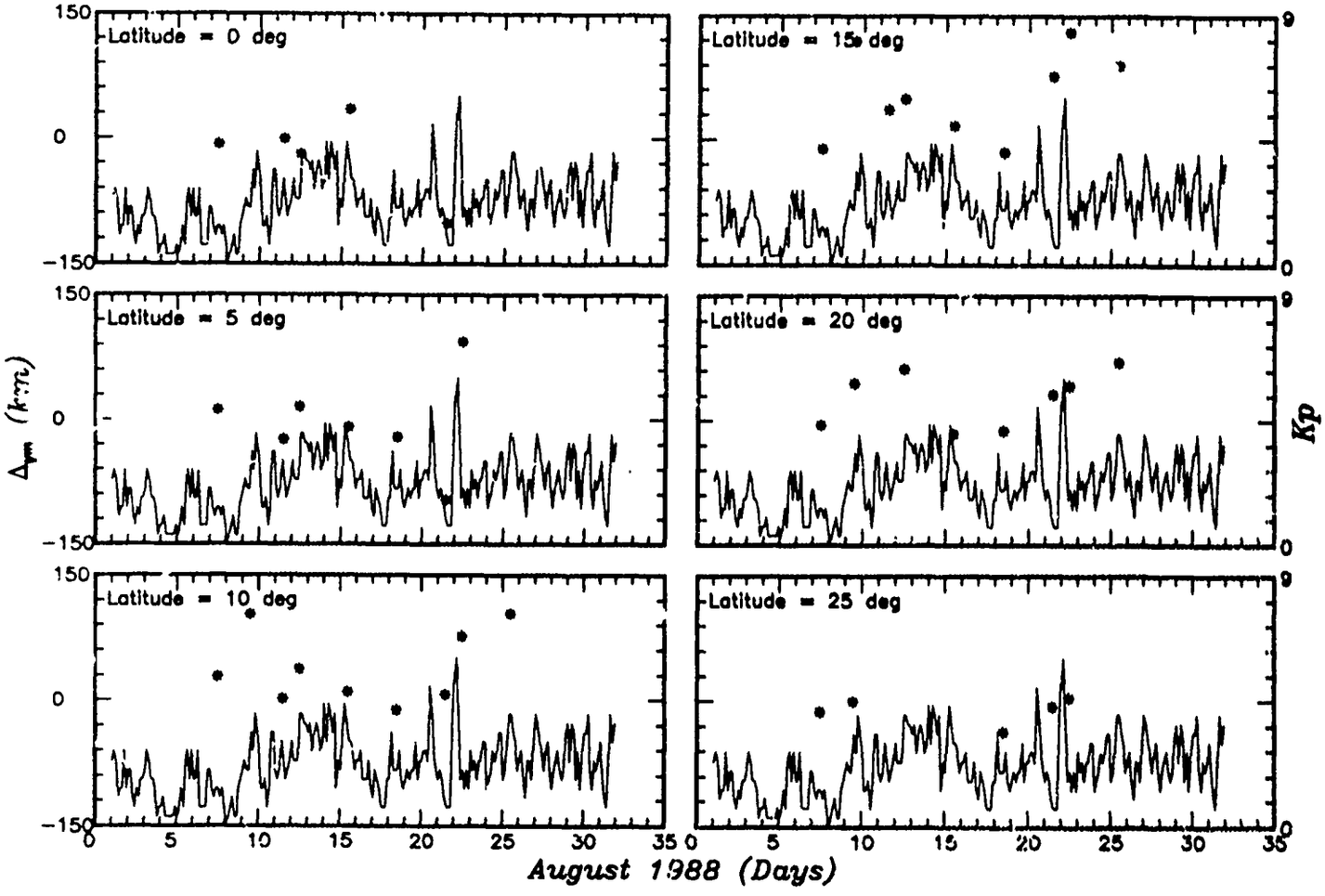


FIGURE 4 - LAYER SEMI-THICKNESS INDEX (ASTERISKS) AND THREE-HOURLY KP GEOMAGNETIC INDEX (SOLID CURVE) FOR DAYS IN AUGUST 1988

A minor problem to be dealt with is the binning of the derived IMRPs for comparison purposes with solar and geomagnetic indices. In this paper, we used only geographic latitude and the day of the month. Geomagnetic latitude is more appropriate and the separation of data into pre- and post- midnight local time sectors may be advantageous. Binning by geomagnetic longitude may also be explored.

Finally, the method can be made more versatile by employing a ray-tracing algorithm that includes azimuthal deviations (this is particularly important for low frequency propagation), a more realistic ionosphere model (e.g., the FAIM model mentioned earlier is more representive of the actual ionospheric structure), and an extension of the regression technique to include thermospheric refraction.

REFERENCES

Anderson, D. N., M. Mendillo and B. Herniter, "A Semi-Empirical Low-Latitude Ionospheric Model", *Radio Sci.*, Vol. 22, pp. 292-306, March-April, 1987.

Anderson, D. N., J. M. Forbes and M. Codrescu, "A Fully Analytic, Low- and Middle-Latitude Ionospheric Model", *J. Geophys. Res.*, Vol. 94, pp. 1520-1524, February, 1989.

Chiu, Y. T., "An Improved Phenomenological Model of Ionospheric Density", *J. Atmos. Terr. Phys.*, Vol. 37, pp. 1563-1570, 1975.

Croft, T. A. and H. Hoogasian, "Exact Ray Calculations in a Quasi-Parabolic Ionosphere", *Radio Sci.*, Vol. 3, pp. 69-74, 1968.

Davies, K., "Recent Progress in Satellite Radio Beacon Studies with Particular Emphasis on the ATS-6 Radio Beacon Experiment", *Space Sci. Rev.*, Vol. 25, pp. 357-430, 1980.

Forbes, J. M., "Evidence for the Equatorward Penetration of Electric Fields, Winds, and Compositional Effects in the Asian/Pacific Sector During the September 17-24, ETS Interval", *J. Geophys. Res.*, Vol. 94, pp. 16999-17007, December, 1989.

Fuller-Rowell, T. J., "Numerical Investigations of Thermospheric/Ionospheric Coupling in the Polar Regions", *Physica Scripta*, Vol. T18, pp, 229-239, 1987.

Hill, J. R., "Exact Ray Paths in a Multi-Segment Quasi-Parabolic Ionosphere, *Radio Sci.*, Vol. 5. pp. 855-861, 1979.

Houpis, H. L. F. and L. J. Nickisch, "An Ionospheric Propagation Prediction Model for Low- and Mid-Latitudes - Applications to Radar Tracking Performance and Ionosphere-Magnetosphere Coupling", *Eos. Trans. AGU*, Vol. 70, No. 43. p. 1248, October, 1989.

Ivanov-Kholodny and A. V. Mikhailov, *The Prediction of Ionospheric Conditions*, D. Reidel Publ. Co., Boston, 1986.

Kelley, M. C., The Earth's Ionosphere: Plasma Physics and Electrodynamics, Academic Press, New York, 1989.

Knepp, D. L., L. J. Nickisch and R. T. Tsunoda, "Kwajalein 1988 Propagation Effects Experiment", MRC Report MRC/MRY-R-010, January, 1988.

Mayaud, P. N., Derivation, *Meaning, and Use of Geomagnetic Indices*, Geophysical Monograph 22, American Geophysical Union, Washington, D. C., 1980.

Press, W. H., B. P. Flannery, S. A. Teukolsky and W. T. Vetterling, *Numerical Recipes - The Art of Scientific Computing*, Cambridge University Press, New York, pp. 521-528, 1987.

Roederer, J. G., GEM: Geospace Environment Modeling, *Eos, Trans. AGU*, Vol. 69, No. 33, p. 786, August, 1988.

Rostoker, G., "Solar-Terrestrial Energy Program 1990-1995 - A Study of Energy Transfer Mechanisms in the Solar-Terrestrial System, *Eos, Trans. AGU*, Vol. 70, No. 14, p. 209, April, 1989.

Rush, C. M., "Ionospheric Radio Propagation Models and Predictions - A Mini-Review", *IEEE Trans. Antennas and Prop.*, Vol. AP-34, No. 9, Sept., 1986.

Schunk, R. W., "Interactions between the Polar Ionosphere and Thermosphere", *Physica Scripta*, Vol. T18, pp, 256-275, 1987.

Smith, R. W., J. W Meriwether, Jr., G. Hernandez, D. Rees, V. Wickwar, O. de la Beaujardiere and T. L. Killeen, "Mapping the Wind in the Polar Thermosphere - A Case Study within the CEDAR Program", *Eos, Trans. AGU*, Vol. 70, No. 11, p. 161, March, 1989.

Walt, M., "STP Activities - United States", STP Newsletter 88-1, distributed by World Data Center - A for STP, pp. 7-8, 1988.

AD-P006 292

EFFECTS OF IONOSPHERIC SCINTILLATION ON DPCA CLUTTER REJECTION PERFORMANCE

J. Todd Reinking, Trent N. Moody and Dennis L. Knepp
Mission Research Corporation
Carmel, California

ABSTRACT

Detection of moving targets by Space-Based Radar (SBR) will be significantly affected by the SBR platform motion. This motion will both shift and spread the clutter spectrum associated with stationary returns, degrading the SBR performance. A technique for compensating for spectrum spreading effects due to motion is to include a displaced phase center antenna (DPCA) as part of the clutter rejection system. Calculations taking into account ionospheric scintillation of the clutter improvement factor (CIF) for such a system have been performed previously. However, to date these calculations have assumed a gaussian correlation function for the ionosphere scintillation component of the radar return envelope.

The effects of ionospheric scintillations are obtained employing current models for ionospheric electron density irregularities, and propagating signals through these irregularities taking into account the detailed geometry of the relevant two-way propagation paths. These CIF calculations then allow DPCA performance for an SBR to be evaluated as a function of both scintillation levels. It is shown that ionospheric scintillation can significantly effect the CIF for a DPCA system, and impose critical performance constraints.

INTRODUCTION

Detection of moving targets from a space based radar (SBR) is constrained not only by the usual clutter rejection problems from a moving platform, but also by signal variations induced by electron density irregularities in the ionosphere. These rapid variations in signal amplitude, phase, and angle-of-arrival are called scintillations and are often observed over satellite links through the ambient ionosphere at VHF and UHF. Since even small fluctuations in received power can cause degraded system performance, the effect of scintillation must be considered in the design of an SBR operating through a disturbed ionospheric propagation channel.

Because the return signal from the earth's surface (clutter) may be very large compared to the target return, it is necessary to insure the clutter does not mask the moving targets of interest. Conventional clutter mitigation is often achieved with MTI filters in cascade with a doppler filter bank. The intended effect of the filters is to attenuate the low frequency clutter, while passing target doppler returns.

Scintillation and platform motion both spread the clutter power spectral density (PSD). In addition, scintillation spreads the target PSD while platform motion may translate the clutter PSD away from zero doppler. As the clutter PSD is shifted to higher frequencies, the performance of radar employing conventional MTI degrades; therefore, it is desirable to compensate for the effects of scintillation and motion.

The mean clutter doppler can be shifted to zero doppler (i.e., the center of the clutter rejection filter notch) using phase locking techniques such as TACCAR [*Skolnik*, 1970]. Spectral spreading due to platform motion can be compensated for using a Displaced Phase Center Antenna (DPCA). The DPCA has the effect of making the radar appear stationary for at least a pair of pulses, increasing the effectiveness of clutter suppression filters. Since the severity of scintillation is a function of platform motion relative to the ionosphere, it is possible to design DPCA techniques that are effective in mitigating both doppler effects due to platform motion and scintillation.

91-09710

The effects of ionospheric scintillation are obtained employing current models for the ionospheric electron density irregularities, and propagating signals through these irregularities, taking into account the detailed geometry of the relevant two-way propagation paths. The clutter improvement factor (CIF) calculations then allow DPCA performance for a SBR to be evaluated as a function of the scintillation level.

The goal of this paper will be to derive an expression for the CIF of radar using the DPCA technique. The resulting expression will be evaluated at the output of each doppler bin for varying degrees of scintillation and clutter.

DPCA OPERATION

The DPCA operation is depicted in Figure 1. This figure shows the antenna positions during transmit and reception of a frame of pulses, taken here to consist of three pulses.

The left side of Figure 1 shows the antenna position during pulse transmission. The right side shows the antenna position on pulse reception. Note that only a fraction of the available aperture is used on reception. The first pulse in a frame is received on the leading antenna section; each successive pulse is received on the next section. Note that these sections are shown to be nonoverlapping for ease of illustration only. The sections can overlap as long as the proper separation between phase centers is maintained.

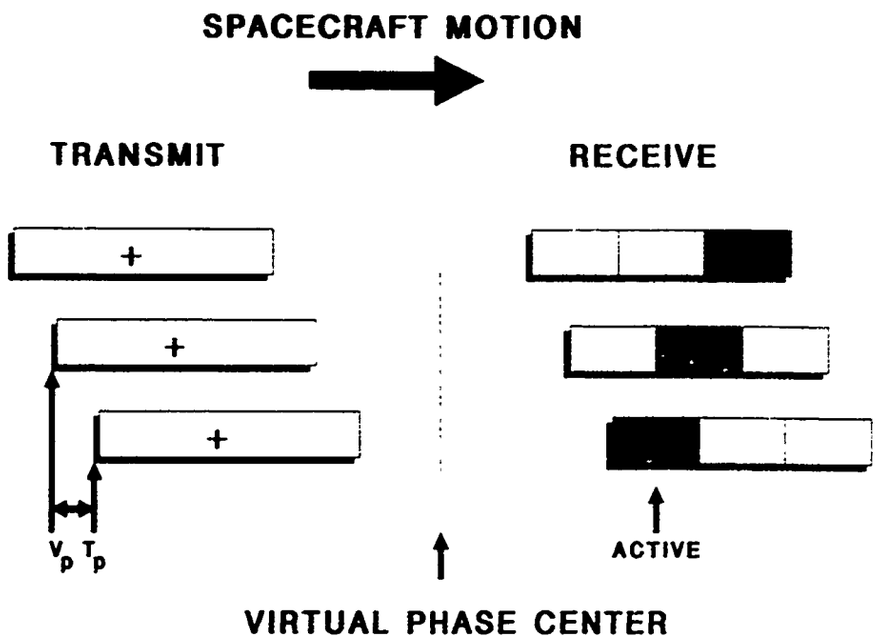


Figure 1. DPCA Transmission and Reception.

Each transmit and receive antenna position forms a bistatic radar. For an SBR the range to the clutter surface will be large enough to treat the bistatic radar formed for each pulse as a monostatic radar with a virtual phase center located midway between the transmit and receive position for a given pulse.

Let δ be the distance traveled by a fixed point on the SBR during the interval between pulses. If the separation between phase centers of adjacent sections of the receive antenna is 2δ then the virtual phase center remains at a fixed position relative to the ground for a frame of pulses. A frame of pulses may now be processed to reject clutter without the effects of motion. However, because each pulse in a frame is received on a different two-way path and the SBR-ground velocity and SBR-ionosphere velocity differ, there will still be some decorrelation due to propagation path irregularities. Of course, intrinsic motion of the clutter contributes additional decorrelation. All these effects are considered in the ensuing analysis.

Note the number of pulses used for DPCA processing need not be constrained to three. Two or more pulses could be used, although consideration of the impact on antenna design and ambiguous doppler for a fixed coherent integration time (to name just a few considerations) place an upper limit on the number of pulses that could practically be considered.

After all K pulses in a single frame are received, the pulses are weighted and summed using binomial weights given by :

$$a_n = (-1)^n \binom{K-1}{n} \tag{1}$$

RADAR RETURN CHARACTERISTICS

The received clutter signal will be modeled as a complex function of time, t, and angle off the antenna bore-sight, θ.

$$r_c(t,\theta) = g_2(\theta)\, a(t)\, s(t)\, \exp\{i\phi_d(t,\theta)\} \tag{2}$$

where :

- $|g_2(\theta)|^2 = G_2(\theta)$ is the two-way antenna power pattern
- $a(t)$ is the envelope fluctuation due to intrinsic clutter
- $s(t)$ is the envelope fluctuation due to variations in the two-way propagation gain
- $\phi_d(t,\theta)$ is the received signal phase after TACCAR correction

The doppler shift of the clutter return from this clutter patch is :

$$f_d = -2\, V_n/\lambda \tag{3}$$

where V_n is the radial velocity of an individual clutter scatterer relative to the SBR, and λ is the transmission wavelength.

If it is assumed the antenna bore-sight is perpendicular to the velocity, then the doppler spread due to platform motion will be worst case, while the mean doppler shift will be zero, assuming the clutter is spatially homogeneous. These assumptions are made through out the remainder of the paper. The geometry considered is shown below.

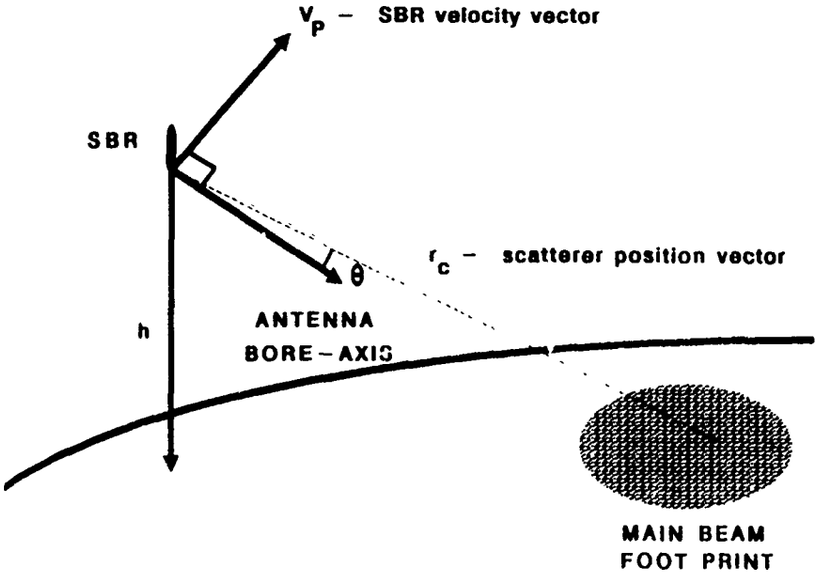


Figure 2. SBR - Clutter Geometry.

For antenna with small beamwidth, only r_c close to the antenna bore-sight need be considered (i.e., only small θ and ϕ need be considered). In which case V_n and f_d are approximately:

$$V_n \approx -V_p\theta \tag{4}$$

$$f_d \approx -2V_p\theta/\lambda \tag{5}$$

As long as V_n remains nearly constant from time t_0 to t, the clutter return phase advance due to platform motion from a point scatterer at r_c is :

$$\phi_d(t,\theta) \approx -4\pi V_p\theta(t-t_0)/\lambda \tag{6}$$

Recall the received clutter signal envelope in Equation 2 is composed of three components. The first is the dependence of the two-way antenna gain on the angle θ defined earlier. Only the mainlobe of the antenna will be considered here. A gaussian antenna pattern is assumed with two-way power gain given by:

$$G_2(\theta) = |g_2(\theta)|^2 = \frac{\sqrt{2}}{\theta_0\sqrt{\pi}}\exp\{-(\theta/\theta_0)^2\} \tag{7}$$

The second component, $a(t)$, is random and represents fluctuations in the envelope due to intrinsic clutter motion. It is assumed the autocorrelation of $a(t)$ is a gaussian function

$$\begin{aligned} R_{aa}(\tau) &= \langle a(t)a^*(t+\tau)\rangle \\ &= \exp\{-(\tau/\tau_c)^2\} \end{aligned} \tag{8}$$

The third term is also random and represents the envelope fluctuations due to two-way propagation through the ionosphere.

SCINTILLATION STATISTICS

The general form for the three-dimensional spatial power spectrum of ionospheric refractive index fluctuations used here was developed by *Wittwer* [1985] and is given by:

$$\Phi_\mu(\vec{k}) = \frac{C_\mu}{[1+L_r^2k_r^2+L_s^2k_s^2+L_t^2k_t^2]^n\,[1+\epsilon^2(L_r^2k_r^2+L_s^2k_s^2+L_t^2k_t^2)]^{p-n}} \tag{9}$$

The constant C_μ controls the magnitude of ionospheric fluctuations, and its exact form will not be important in what follows here. Φ_μ describes a "two-component" (anisotropic) power law spectrum with the L_i's corresponding to the largest inhomogeneities occurring in the medium (the spectrum is flat for $\sum_{i=1}^{3} L_i^2K_i^2 \ll 1$), and the ϵL_i's corresponding to the scales at which the spectrum makes the transition from a power law with index n, to a power law with index p, these are referred to as the "freezing scales". For $\epsilon = 1$, Φ_μ reduces to a single-component spectrum with a power law index of p, and for (spatial) scales smaller than the L_i's this spectrum is identical (up to a factor of $\frac{1}{\epsilon^{2\alpha}}$) to the two-component model for scales smaller than the freezing scales.

In particular L_t is the outer scale along the earth's magnetic field lines, while L_r and L_s are the outer scales transverse to the field lines. *In situ* measurements indicate the anisotopy ratio,

$$a \equiv \frac{L_t}{L_r} \approx 10 \tag{10}$$

and

$$10 \text{ km} \leq L_r, L_s \leq 10^2 \text{ km}. \tag{11}$$

Parameters for this two-component power law are provided by the WBMOD database of ionospheric parameters [*Secan and Fremouw*, 1988]. *In-situ* irregularity strength parameters [*Secan and Fremouw*, 1988] are used to compute C_μ, and for the results shown here, $p = 2$, $\epsilon = 1.0$, $L_r = L_s = 30$ km, and $a = 10$. The ionospheric layer was given a thickness of 200 km centered at a height of 350 km, the Zurich sunspot number was set to 200, the ionospheric Kp index was 3.33, and the WBMOD database was sampled for day 240 at 2200 hours local time (these parameters were chosen to ensure significant levels of scintillation).

Signal correlation functions are then provided by the ionospheric scintillation code WWMOD which computes the mutual coherence function for a one-way path based on numerical evaluation of analytical expressions which are appropriate for the two-component spectrum discussed above. Two-way signal correlations for two pulses emitted at two different times and spatial positions, and similarly received at two different times and spatial positions, are obtained from the one-way coherence values via the relation [*Knepp*, 1990]

$$\begin{aligned}\Gamma(\text{two-way path}) &= \Gamma(\Delta T, \vec{R}_{t2} - \vec{R}_{t1}) \times \Gamma(\Delta T, \vec{R}_{r2} - \vec{R}_{r1}) \\ &+ \Gamma(\Delta T - \frac{2Z_{ti}}{c}, \vec{R}_{t2} - \vec{R}_{r1}) \times \Gamma(\Delta T + \frac{2Z_{ti}}{c}, \vec{R}_{r2} - \vec{R}_{t1}) \end{aligned} \quad (12)$$

where the Γ's on the right-hand-side of this equation correspond to one-way path coherence values. Z_{ti} is the line of sight distance between the target and (the middle of) the ionospheric layer, and c is the speed of light. T corresponds to the time interval for the correlation, $\vec{R}$ corresponds to a spatial position, while the subscripts $t1$, $t2$, $r1$, $r2$ label transmission one, transmission two, reception one, and reception two, respectively. Note that the one-way coherence values are only functions of temporal and spatial separations, and not functions of absolute times or positions, indicating that the two-way coherence function is not dependent on the angle of arrival of the received signal.

CLUTTER IMPROVEMENT FACTOR

The measure frequently used to asses the performance of a clutter rejection technique is the clutter improvement factor (CIF). The CIF considered here is modified slightly to also account for receiver noise. Specifically, the CIF is defined for the output of the k^{th} radar doppler bin as, I_k, where:

$$I_k = \frac{P_T^{\text{out}}(k)/[P_C^{\text{out}}(k) + P_N^{\text{out}}(k)]}{P_T^{\text{in}}/[P_C^{\text{in}} + P_N^{\text{in}}]} \quad (13)$$

- P_T^{in} - Input target power
- P_C^{in} - Input clutter power
- P_N^{in} - Input receiver noise power
- P_T^{out} - Target power in output of k^{th} doppler bin
- P_C^{out} - Clutter power in output of k^{th} doppler bin
- P_N^{out} - Noise power in output of k^{th} doppler bin

The CIF for the k^{th} doppler bin is the ratio of the processed target power to the sum of the processed clutter power and processed receiver noise divided by the ratio of the input target power to the sum of the input clutter power and input receiver noise. With this definition, an increase in CIF indicates improved system performance.

By determining the CIF at the output of each doppler bin and comparing the result with a minimum acceptable CIF required to maintain an acceptable detection probability, a so called minimum detectable velocity (MDV) can be determined.

It is convenient to refer to a noise power-to-clutter power ratio as opposed to the absolute level of noise power. Without loss of generality, the denominator of Equation 13 is selected as unity. To do so, choose $P_T^{\text{in}} = 1$ and $[P_C^{\text{in}} + P_N^{\text{in}}] = 1$. Define E_w as:

$$E_w = P_N^{\text{in}}/P_C^{\text{in}} \quad (14)$$

The clutter + noise power at the input, P_{CN}^{in}, is then :

$$P_{CN}^{\text{in}} = P_C^{\text{in}} + P_N^{\text{in}} = P_C^{\text{in}}(1 + E_w) \quad (15)$$

Since it is required $P_{CN}^{\text{in}} = 1$, and E_w is generally very small:

$$P_C^{\text{in}} = 1/(1 + E_w) \approx 1 \quad (16)$$

and:

$$P_N^{\text{in}} = E_w/(1 + E_w) \approx E_w \quad (17)$$

The output target and clutter + noise power are found from the input statistics by first determining the correlation of the compensated stage output, R_{zz}, then finding the power at the output of the uncompensated stage for each doppler bin using R_{zz}.

The uncompensated stage is simply a linear transversal filter. Possible designs include an M-stage MTI filter in cascade with an unweighted N-point DFT (Discrete Fourier Transform) filter bank, or simply a weighted DFT filter bank. The former is considered here with binomial MTI filter weights.

Let the transversal filter operate on L pulses. When $\mathbf{x}$ is a vector (the term vector is used to imply a column vector) of the most recent L outputs of the compensated stage, the transversal filter output is:

$$\mathbf{y}_k = \mathbf{w}_k^T \mathbf{x} \tag{18}$$

and the expected output power for the k^{th} doppler bin is:

$$\langle \mathbf{y}_k \mathbf{y}_k^* \rangle = \mathbf{w}_k^T \mathbf{C}_{zz} \mathbf{w}_k^* \tag{19}$$

where $\mathbf{C}_{zz}$ is the correlation matrix of $\mathbf{x}$.

Sampled clutter returns are identified by a frame index, n, and channel index, m. Below let $\Delta m = m - m'$, $\Delta n = n - n'$, and $\tau = [\Delta m + 3\Delta n]T_p$, then form three-pulse DPCA, the sampled clutter return correlation is:

$$R_{cc}(\Delta m, \Delta n) = R_{ss}(\tau) R_{ss}(\Delta m, \Delta n) \int_{-\pi/2}^{\pi/2} G_2(\theta) \exp\{i\, 4\pi V_p \theta 3 \Delta n T_p / \lambda d\theta\} \tag{20}$$

The last term is the decorrelation due to motion and is independent of Δm due to the use of DPCA.

Define $\mathbf{r}(n)$ as vector where the i^{th} element corresponds to the sum of clutter and noise in the i^{th} DPCA channel. Model the noise as zero mean, white noise with variance E_w. If $\mathbf{a}$ is the filter weight vector with elements given by Equation 1, then the motion compensated output at the end of a frame due to clutter and receiver noise is:

$$x(n) = \mathbf{a}^T \mathbf{r}(n) \tag{21}$$

The correlation of $x(n)$ is therefore :

$$\begin{aligned} R_{xx}(\Delta n) &= \mathbf{a}^T < \mathbf{r}(n) \mathbf{r}^H(n + \Delta n) > \mathbf{a} \\ &= \mathbf{a}^T (\mathbf{C}(\Delta n) + E_w \delta(\Delta n) \mathbf{I}) \mathbf{a} \end{aligned} \tag{22}$$

where the elements of $\mathbf{C}$ are obtained from Equation 20, superscript H represents the Hermitian transpose, k is the scalar Kronecker delta function, and I is the identity matrix.

Determination of the target power, $P_T^{out}(k)$, is done in a manner similar to determining the clutter power. For results obtained here the target return is modeled as a complex exponential at doppler frequency f_d:

$$r_t(t) = s(t) \exp\{i 2\pi f_d t\} \tag{23}$$

The target signal correlation is:

$$R_{tt}(\Delta m, \Delta n) = R_{ss}(\Delta m, \Delta n) \exp\{i 2\pi f_d n \tau\} \tag{24}$$

RESULTS

The clutter improvement factor described above has been evaluated for a pulse-doppler SBR operating over a propagation channel with varying degrees of scintillation. Clutter mitigation is examined for two cases. The first using a three-pulse DPCA followed by a doppler filter bank (DFT). The second using a two-pulse DPCA, again followed by a doppler filter bank. The assumed radar parameters are: Transmit PRF – 577 Hz; T_{ci} – 222 ms; Frequency – UHF and Beamwidth – 2.4°; where T_{ci} is the coherent integration time.

Given this set of parameters the DFT length is 43 for the three-pulse DPCA system, and 64 for the two-pulse DPCA system. (Recall the data rate after the DPCA is reduced by the number of pulses in a frame.) Note the beamwidth given is the one way half-power beamwidth on transmit, θ_{3dB}^1. The effective two-way half-power beamwidth, θ_{3dB}^2, is calculated using:

$$\theta^2_{3dB} = \frac{1}{\sqrt{(1/\theta^1_{3dB})^2 + (1/\mu\theta^1_{3dB})^2}} \tag{25}$$

where μ is the fraction of the transmit aperture used on receive. For this study μ is assumed to be 2/3 for the two-pulse DPCA and 1/3 for the three-pulse DPCA.

The variance of the velocity spectrum of the clutter is chosen as 1 m/s, typical of sea-clutter. This results in a clutter decorrelation time of 168 ms. Results presented here assume E_w is zero.

To calculate the propagation channel coherence, the SBR altitude is selected as 1000 km, with a velocity of 6730 m/s. The coherence values used to calculate the CIF where typical of those predicted along a path between the point at -1.1° of latitude, 166.3° of longitude and the point at 9.0° of latitude, 163.7° of longitude while viewing a point on the ground at 9.4° of latitude and 167.5° of longitude (Kwajalein) for the conditions noted previously. Three sets of coherence values are considered for the given cumulative probabilities. An example of interpreting these probabilities is given. Under the given ionosphere conditions and transmitter-receiver geometry only 5% of the time will signal decorrelation be worse than that predicted for coherence values from the 5% curve.

Work currently in progress has resulted in the calculated CIF's shown in Figures 3 and 4 for two-pulse DPCA and three-pulse DPCA systems, respectively. Each curve shows the CIF for conditions of no scintillation and the various levels of scintillation indicated. Note, to produce the figures the improvement factor is calculated at the center of each of the doppler filters. These calculated values are then connected by straight lines to form the two figures. Therefore, the CIF for a given target will be lower if the target doppler falls away from the center of the doppler bin.

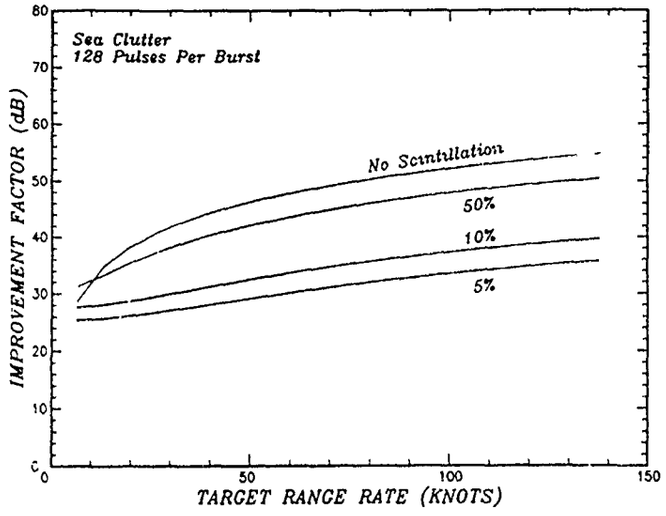


Figure 3. Clutter Improvement Factor for Two-pulse DPCA.

A typical value of clutter attenuation required for satisfactory performance against sea clutter is 35 dB for the radar considered. Both systems have an MDV below 20 knots for ambient conditions, but for the worst fading conditions considered the three-pulse system is unable to obtain the required clutter rejection while the two-pulse system has an MDV of about 100 knots. For coherence values from the 50% distribution point, the MDV is about 20 knots for the two-pulse system and 80 knots for the three-pulse system.

Notice that under ambient conditions, the clutter rejection offered by the three-pulse system is significantly greater than that available from the two-pulse system. Under even moderate fading conditions, however, the opposite is true. The loss in performance of the DPCA system in scintillating conditions is due to each pulse being received on a separate two-way path. The poorer performance of the three-pulse system relative to the two-pulse system is due primarily to the increased separation between transmit and receive phase centers of pulses received using the three-pulse system, which results in increased signal decorrelation relative to the two-pulse system.

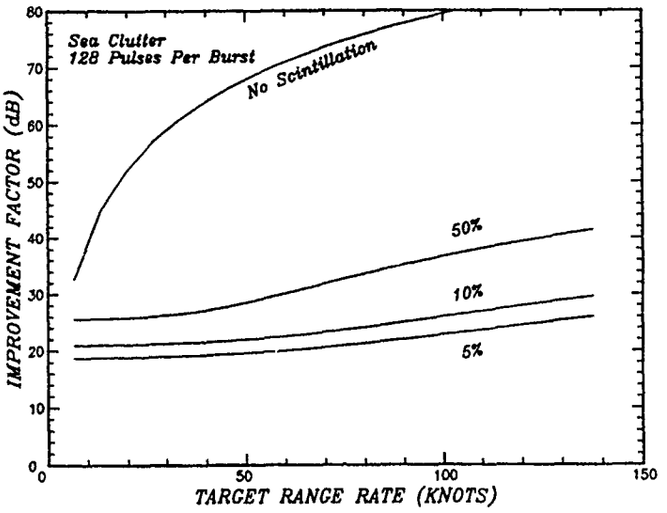


Figure 4. Clutter Improvement Factor for Three-Pulse DPCA.

CONCLUSIONS

A method was developed for obtaining the clutter improvement factor for an SBR employing a DPCA system, taking into account the effects of scintillation due to irregularities in the propagation path. The CIF was found for a two-pulse and three-pulse DPCA operating against sea clutter. The results indicated a significant increase in the MDV under severe fading conditions, with the three-pulse system unable to obtain the minimum acceptable clutter rejection. Counter to intuition, it was found that the two-pulse system was superior to the three-pulse system under even moderate fading conditions.

Obvious methods for increasing the achievable clutter rejection include increasing the PRF and decreasing the antenna beamwidth. The latter has the dual effect of increasing the signal coherence while decreasing the area of illuminated clutter.

Some additional clutter rejection may be achieved by using a weighted DFT in the case of the two-pulse system. This strategy would not be directly effective in the case of the three-pulse system, however, because the DPCA processor essentially "whitens" the clutter spectrum for the conditions considered.

In addition to considering the effects of varying parameters of the clutter rejections system just considered, future investigations should consider alternate DPCA strategies, including those that use auxiliary antenna patterns for phase correction to compensate for motion effects [*Andrews*, 1977]. A DPCA scheme that uses the same bistatic path for N pulses is also possible and should be investigated. Such a scheme could provide enhanced clutter rejection under severe fading conditions, at the expense of using only a fraction of the available antenna aperture on transmit. The inclusion of receiver noise will be necessary when comparing the relative performance of each of these systems.

REFERENCES

Andrews, G.A. and K. Gerlach, "SBR Clutter and Interference", Chapter 11 in Space-Based Radar Handbook, L.J. Cantafio (ed.), Artech House, 1989.

Andrews, G.A., "An Airborne Radar Doppler Processing Philosophy," Naval Research Laboratory Rep. 8073, January 1977. Reprinted in Lewis, B.L., F.F. Kretschmer, and W.W. Shelton, "Aspects of Radar Signal Processing", Artech House, Norwood, MA, 1986, pp. 465-499.

Knepp, D.L. and H.L.F. Houpis, "VHF/UHF Radar Scintillation Effects Observed by ALTAIR", prepared for the IES '90 Proceedings, February 1990.

Knepp, D.L. and R.A. Dana, The Impact of strong Scintillation on Space Based Radar Design - Clutter Rejection, MRC-R-787, Mission Research Corporation, October 1983.

Knepp, D.L. and R.A. Dana, "Impact of the Propagation Environment on Space Based Radar Design", AIAA 20th Aerospace Sciences Meeting, 1982.

Secan and Fremouw, "Development of Advanced Ionospheric Models in Support of Space-Based Radar Design Studies", NWRA-CR-R022, Subcontract No. SC-RD04-87- 0003, 1988.

Skolnik, M. I., Introduction to Radar Systems, Second Edition, McGraw-Hill, 1980.

Skolnik, M. I., (ed.), Radar Handbook, McGraw-Hill, 1970.

Wittwer, L. A., "Radio Wave Propagation in Structural Ionization for Satellite Applications III", Defense Nuclear Agency, 1985.

AD-P006 293

UHF RADAR BACKSCATTER ENHANCEMENTS FROM THE TOPSIDE IONOSPHERE

K. M. Groves, M. C. Lee
Plasma Fusion Center
Massachusetts Institute of Technology
Cambridge, Massachusetts 02139

J. C. Foster
M.I.T. Haystack Observatory
Westford, Massachusetts 01886

91-09709

ABSTRACT

Intense ($\simeq$20 dB) enhancements of the incoherent radar backscatter spectrum from the topside ionosphere have been observed with the Millstone Hill UHF radar; enhancements occurring at the local ion acoustic frequency causing large asymmetries in the measured ion line may be produced by current-driven instabilities, (Foster, et. al, *GRL, 15*, 160, 1988). In addition to current-driven processes, we propose the generation of intense ion acoustic turbulence by the nonlinear interaction of Langmuir waves in the ionosphere. Qualitatively, an enhanced Langmuir wave spectrum is produced by the action of suprathermal electrons associated with field-aligned currents. Energy from the enhanced Langmuir waves can then be coupled into ion acoustic modes through nonlinear processes. The resulting amplification of the ion modes can be detected by incoherent scatter radars. Analytical and numerical calculations describing the above processes quantitatively will be presented. This newly observed phenomenon is associated with intriguing ionospheric plasma processes. The induced plasma modes having large radar cross-sections give rise to strong backscatter radar echoes. Hence, in the past, they were probably misidentified to be satellite radar returns. From the point of view of space surveillance, it is desirable to identify potential geophysical processes leading to false satellite signals.

1. Introduction

Large enhancements of the incoherent backscatter from the topside ionosphere are observed frequently with the Millstone Hill UHF radar. One class of these enhancements is characterized by large spectral asymmetries in the so called "thermal ion line" normally measured by the incoherent scatter radar. These enhanced returns had been attributed to hard targets intersecting the radar beam; recent studies, however, suggest that some of the enhancements may be caused by current-driven ionospheric plasma processes [*Foster et al.*, 1988]. The processes described by Kindel and Kennel [1971] and Rosenbluth and Rostoker [1962], for example, can explain the ion acoustic wave power enhancements (~20 dB), but relatively large current densities (~100 μamps/m^2) are required to produce the magnitudes of the observed effects.

While currents of this intensity in the topside ionosphere have occasionally been reported [*Bythrow et al.*, 1984; *J. D. Winningham, personal communication*, 1987], the frequency of such events is unknown. We, therefore, have performed a preliminary theoretical investigation of a new mechanism which can enhance radar backscatter in conjunction with known current driven processes, thereby reducing the threshold amplitudes of the current densities they require. The proposed scenario is as follows. Energetic electrons associated with intense current filaments in the ionosphere excite a suprathermal spectrum of Langmuir waves as they stream through the background plasma. Subsequently the Langmuir waves can become nonlinearly coupled via the ponderomotive force to induce low frequency density perturbations driving ion acoustic waves in the ionospheric plasmas.

91

In Section 2 we briefly describe new multi-diagnostic observations which provide additional evidence supporting the existence of "false satellite signals" in the topside ionosphere. A preliminary theory of ion acoustic wave enhancement through the nonlinear interaction of suprathermal electron-induced Langmuir waves is outlined in Section 3. Section 4 contains a quantitative analysis of the ion acoustic wave enhancement based on satellite measurements of energetic electron fluxes. The results and conclusions are summarized in Section 5.

2. Observations

Beginning in April, 1989, a series of multi-diagnostic experiments, known as MICAD, involving the MIllstone Hill UHF radar, part of the CANOPUS system of instruments (BARS radar, magnetometer array, meridian photometers) and the DE-1 satellite was conducted to provide additional information about ion acoustic radar backscatter enhancements due to ionospheric processes. In these experiments, as illustrated in Figure 1, the Millstone Hill radar was used to track the geomagnetic field line above the DE-1 satellite as it flew through perigee within the radar's field of view; CANOPUS instruments were utilized as diagnostics of the background ionosphere. Ultimately, the goal was to simultaneously detect the signature of large, field aligned ionospheric currents with remote sensors (radars, etc.) and in situ instruments aboard DE-1.

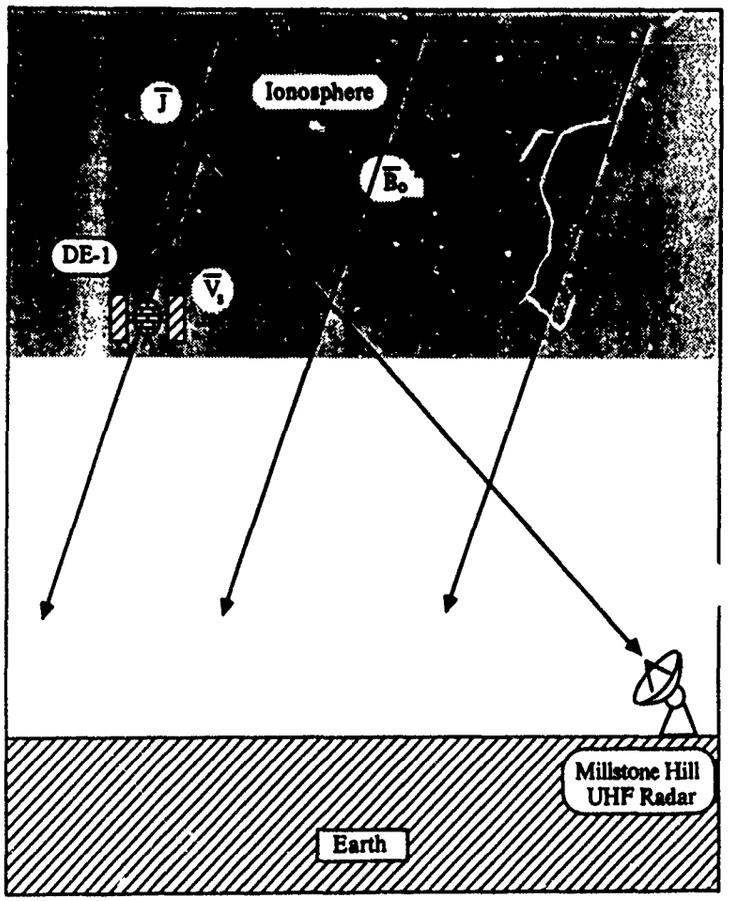


Figure 1. Cartoon illustrating the concept of MICAD experiments. The Millstone Hill UHF radar tracks the geomagnetic field line above the DE-1 satellite as it passes through perigee near Millstone Hill. Intense field aligned currents should be detected by both sensors simultaneously.

During one such MICAD experiment on 22 May, 1989, the Millstone Radar detected a 7 dB enhanced backscatter return at the ion acoustic frequency at 23:58 UT; the DE-1 satellite passed through the same region 30 seconds later and recorded a geomagnetic fluctuation corresponding to a field aligned current density ≥ 50 μamps/m^2 (see Figure 2). Because the radar was actually tracking about 30 seconds ahead of the satellite's position, however, the measurements cannot be classified as simultaneous. An examination of data from two consecutive passes of the DMSP-F8 satellite (the first 42 minutes prior to, and the second 56 minutes after the DE-1 overflight in the same longitude band) reveal that the region producing the enhanced radar backscatter was characterized by localized intense, soft (E$\leq$ 1 keV) electron precipitation features. Comparison of the measured electron and ion fluxes show that current density spikes in excess of 60 μamps/m^2 were detected on both passes. Data from the first pass formed the basis for deriving the suprathermal contribution to the electron distribution function. Finally, particle flux spectrograms from a northern polar pass of the NOAA-10 satellite exhibit a sharp local maximum in the low energy electron flux within a degree in latitude of, and only 8 minutes prior to, the DE-1 satellite and Millstone Hill radar observations. These in situ data establish the presence of discrete low energy particle precipitation events carrying large current densities in the region of enhanced UHF radar backscatter.

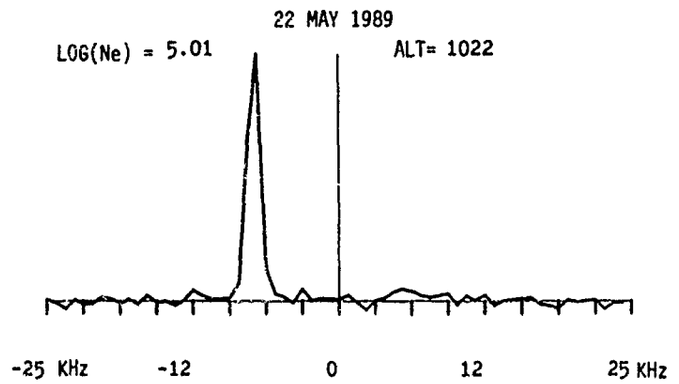


Figure 2. Asymmetric 7 dB enhancement of a normal incoherent scatter power spectrum measured by the Millstone Hill UHF radar tracking ahead of the DE-1 satellite at 23:58 UT, 22 May, 1989.

3. Theory

Because of large Landau damping by bulk electrons, a plasma at or near thermal equilibrium exhibits only very weak electron plasma oscillations (Langmuir waves). If the electron distribution function has a nonthermal, high energy tail, however, it is possible for the high energy electrons to resonantly excite intense Langmuir waves with large phase velocities which are relatively undamped. The spectrum of enhanced Langmuir waves in the presence of suprathermal electrons can be calculated by evaluating the normalized velocity dependent plasma temperature T_p [*Yngvesson and Perkins*, 1968]:

$$\frac{T_p(v_\phi)}{T} = \frac{f_m(v_\phi) + f_p(v_\phi) + \chi}{f_m(v_\phi) - KT\frac{d}{dE_\phi} f_p(v_\phi) + \chi} \tag{1}$$

where f_m and f_p represent, respectively, the Maxwellian and photoelectron contributions to the one-dimensional velocity distribution. They are given by

$$f_m = n\left(\frac{m}{2\pi KT}\right)^{1/2} \sum_{n=-\infty}^{\infty} \frac{e^{-b\sin^2\theta}}{\cos\theta} I_n(b\sin^2\theta) e^{-\left[\frac{(v-n)^2}{2b\cos^2\theta}\right]} \tag{2}$$

$$f_p = \int_{E_\phi}^{\infty} \left(\frac{m}{8E}\right)^{1/2} \rho(E)\, dE \tag{3}$$

The quantity χ in (1) describes the excitation and damping of plasma waves by electron-ion collisions and has the following expression:

$$\chi = \frac{32\pi}{3\lambda^3}\left(\frac{m}{2\pi KT}\right)^{1/2} \ln[4\pi nD^3] \tag{4}$$

The parameters in (2)–(4) are defined by $y = \omega_p/\omega_c$, $b = k^2a^2/2$, and $a = \omega_c^{-1}\sqrt{2KT/m}$; n, T, k, θ, denote the electron density, electron temperature, wave number associated with the backscatter geometry, and the angle between the radar beam and the earth's magnetic field, respectively. The density of particles per unit volume-energy is given by $\rho(E)$; this parameter is derived from satellite data for the quantitative analysis presented in Section 3. When $\rho(E)$ is sufficiently large at high energies, a spectrum of enhanced Langmuir waves will be excited near the plasma frequency. These Langmuir waves can become a potential source for generating ion acoustic waves through nonlinear coupling via the ponderomotive force. The ponderomotive force induces low frequency ($\omega = \omega_1 - \omega_2$) density perturbations, δn, with the form

$$\delta_n = n_o \frac{\vec{F}\cdot(\vec{k}_1 - \vec{k}_2)\sigma_e\sigma_i}{e^2(\omega_1 - \omega_2)(\sigma_e + \sigma_i)} \tag{5}$$

where $\vec{F}$ represents the ponderomotive force, $(\omega_1, \vec{k}_1)$ and $(\omega_2, \vec{k}_2)$ represent two high frequency Langmuir waves, and σ_e and σ_i are the electrical conductivities for electrons and ions, respectively. The scattering cross-section of these low-frequency density perturbations can be expressed as

$$\sigma_s(\omega_o + \omega)d\omega = \frac{\sigma_T < |\delta_n(k,\omega)|^2 >}{2\pi V_L t_L} d\omega \tag{6}$$

where $<|\delta_n(k,\omega)|^2>$ is the power spectrum of the density perturbations, σ_T is the Thomson differential cross-section, V_L is a very large volume and t_L is a very long time interval. If the enhanced Langmuir waves are characterized by a spectral density $W(\vec{k}_1)$, the expression (7) can be written as

$$\sigma_s(\omega_o + \omega) = \gamma k^{-1} \int\int W(\vec{k}_1)W(\vec{k}_1 - \vec{k}_2)\,d\vec{k}_1 \tag{7}$$

where γ is a parameter to be determined by the characteristics of ionospheric plasmas and the detailed nonlinear mode coupling process.

4. Analysis

The velocity dependent temperature, T_p, expressed by (1) represents the enhancement of the Langmuir wave power relative to the weak thermal level power as a function of phase velocity (v_ϕ) and aspect angle (θ); its value depends on three primary parameters, f_m, f_p and χ. Of these three, f_m and χ are determined by the background ionospheric parameters, given by $T = 2000°$ K, $\omega_p/2\pi = 2.8$ MHz, and $\omega_c/2\pi = 1.1$ MHz. As previously stated, f_m represents the Maxwellian contribution to the one-dimensional electron distribution function and vanishes for velocities much larger than the electron thermal velocity ($v_\phi \gg v_t h$).

At such large velocities, the ratio in (1) is usually dominated by the electron-ion collision term, χ, and T_p remains equal to the usual electron temperature, T. If, however, the photoelectron (suprathermal) contribution to the distribution function, f_p, is sufficiently large and its derivative sufficiently small at a given energy, T_p may be many times greater than T, indicating the resonant excitation of intense electron plasma waves. These plasma waves may then couple nonlinearly to drive ion acoustic waves as described in the previous section.

The quantities used to derive $f_p(E)$ in this calculation are from electron flux measurements made aboard the DMSP-F8 satellite (31 ev - 31 keV) and the DE-2 satellite (5 eV - 31 eV) (*W. Denig*, personal communication, 1990; *J. D. Winningham*, personal communication, 1987). The resulting values for $\rho(E)$ were then extrapolated linearly down to 1 eV and integrated to give the one-dimensional distribution function along the magnetic field, f_p. Figure 3 shows the derived f_m and f_p as a function of particle velocity used to calculate the spectrum of excited Langmuir waves via equation (1). To estimate the magnitude of the radar cross section enhancement of the ion acoustic fluctuations, a quantity $T_p' \propto W(\vec{k}_1)W(\vec{k}_1 - \vec{k}_2)$, the integrand in (7), was calculated; the results are displayed in Figure 4. KZ and KY represent the ion acoustic wave vectors parallel and perpendicular to the magnetic field, respectively. The plot shows substantial ion acoustic wave enhancement over a wide range of wavenumbers, with the most intense amplification occurring for wavenumbers less than ten. This is not surprising because smaller wavenumbers correspond to higher energy particles which completely dominate the Maxwellian term in (1) at energies much greater

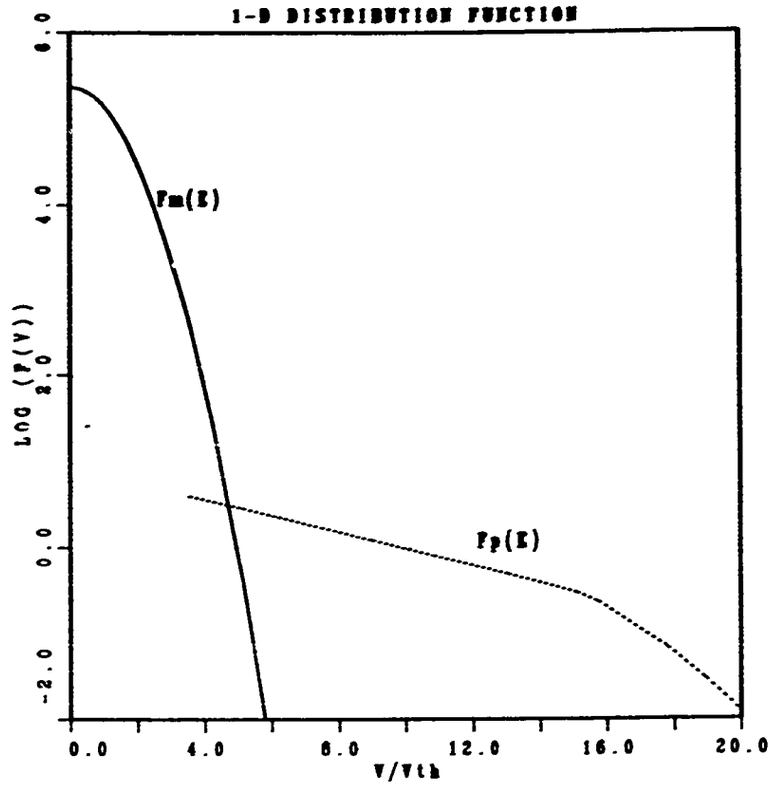


Figure 3. The 1-D electron distribution function used to calculate the enhanced Langmuir wave spectrum. The ordinate is in units of $m^{-4}sec$, while the abscissa is normalized by the thermal velocity, V_{th}.

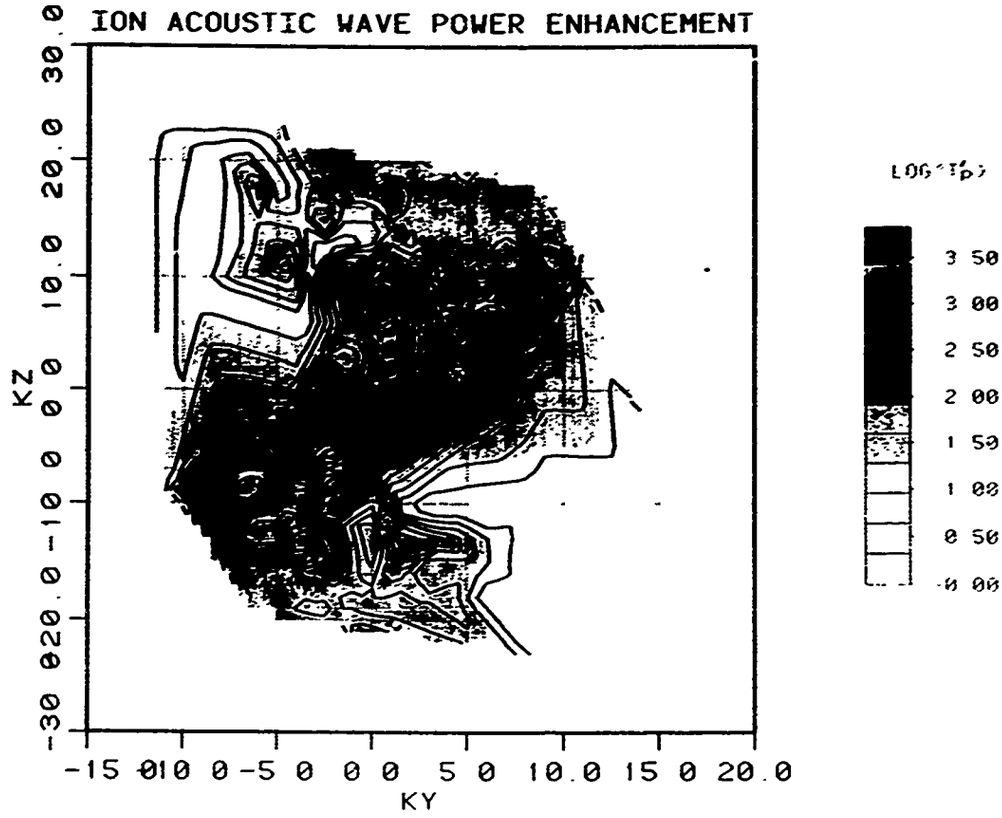


Figure 4. Enhancement of ion acoustic wave modes as a function of parallel (k_z) and perpendicular (k_y) wavenumbers. Actual enhancement of radar cross section can be obtained by integrating over k space.

than the thermal energy, KT. For the backscatter wave number at which Millstone's 440 MHz incoherent scatter radar is sensitive, $k \sim 18.5$ m^{-1}, up to two orders of magnitude enhancement are predicted.

The predicted enhancement of wave modes with large aspect angles relative to the magnetic field was somewhat unexpected because the suprathermal electron distribution used in the calculation was assumed to be field aligned. Indeed, for electron energies below about 10 eV, the excited Langmuir wave spectrum is highly field aligned due to the strong magnetic field dependence of the Landau damping; at higher energies, however, wave phase speeds are sufficiently large to support propagation with relatively weak damping at all aspect angles. Because T_p' is derived from a product of two Langmuir wave modes, it is possible for the field aligned electrons to couple energy through the Langmuir waves into a broad spectrum of ion acoustic modes.

5. Summary

An attempt has been made to provide conclusive observational evidence of geophysically-produced enhanced ion acoustic radar backscatter and to partially explain such backscatter through a theoretical investigation of a nonlinear mode coupling mechanism. Through a series of multiple diagnostic experiments, we have established the association of enhanced backscatter events with large field aligned currents. Suprathermal electrons present in the current filaments excite a broad spectrum of Langmuir waves within the space plasma. Significant enhancement of ion acoustic waves through the nonlinear Langmuir wave coupling is apparently possible when the local electron distribution function has a substantial nonthermal component, as in the presence of intense field aligned currents. This mechanism may be an important effect acting in conjunction with other current driven processes, mentioned previously, in amplifying the UHF radar cross section of ion modes. The calculations presented here represent the enhancement of individual

ion acoustic modes; to determine the net increase in radar cross section for a given acoustic speed, one must integrate the individual modes in k-space as specified by (7). There will be some cancellation in this integration and the net enhancement in cross section will be less than the amplification of some individual modes. A more detailed investigation of these effects is needed.

Acknowledgments Special thanks to Dr. W. K. Peterson of the Lockheed Corporation for his assistance in acquiring DE-1 satellite data and helpful discussions, to Dr. W. Denig of the Geophysics Laboratory for his assistance in providing information and data from the DMSP satellites, and to the entire CANOPUS scientific team for their cooperation and enthusiasm in coordinating joint observations with the Millstone Hill radar. This work was supported by NASA contract NAG5-1055, AFGL contract F19628-90-K-0024, and the AFOSR grant AFOSR-88-0217.

References

Bythrow, P. F., T. A. Potemra, W. B. Hanson, L. J. Zanetti, C.-I. Meng, R. Huffman, F. Rich, and D. Hardy, Earthward directed high-density Birkeland currents observed by HILAT, *J. Geophys. Res.*, 89, 9114–9118, 1984.

Ergun, Robert E., Linear and nonlinear wave processes in the auroral ionosphere, U. C. Berkeley, Ph. D. Dissertation, 1989.

Foster, J. C., C. del Pozo, K. Groves, and J.-P. St. Maurice, Radar observations of the onset of current driven instabilities in the topside ionosphere, *Geophys. Res. Lett.*, 15, 160–163, 1988.

Kindel, J. M , and C. F. Kennel, Topside current instabilities, *J. Geophys. Res.*, 76, 3055–3078, 1971.

Rosenbluth, M. N., and N. Rostoker, Scattering of electromagnetic waves by a non-equilibrium plasma, *Phys. Fluids*, 5, 776–788, 1962.

Yngvesson, K. O., and F. W. Perkins, Radar Thomson scatter studies of photoelectrons in the ionosphere and Landau damping, *J. Geophys. Res.*, 73, 97–109, 1968.

AD-P006 294

IONOSPHERIC EFFECTS RELATED TO THE PERFORMANCE OF AN ORBITAL DEBRIS RADAR SYSTEM

Dennis F. Bishop
Jet Propulsion Laboratory
California Institute of Technology
4800 Oak Grove Drive
Pasadena, CA 91109

ABSTRACT

An orbital debris radar system is designed to detect the presence of small objects in low earth orbit by reflecting radio waves off of the objects. Free electrons and ions contained in the ionosphere cause incoherent scatter of the radar signal. This ionospheric reflection tends to increase the noise at the terrestrial radar receiver. A parameter called the ionospheric scattering cross section per unit volume, which is a function of altitude, is useful for computing the power of the ionospheric reflected signal. The doppler frequency spread of the ionospheric reflected signal is a function of altitude, also. The ionospheric noise of a 9 GHz orbital debris radar receiver is computed using these concepts. Annual and diurnal variations of the noise are included.

1.0 INTRODUCTION

Preliminary design plans for an orbital debris radar system are to place an antenna in a tropical region near the equator with a zenith orientation. The antenna launches an electromagnetic wave in the zenith direction and listens for an echo from objects that are in low earth orbit. However, the return echos are corrupted by noise, including clutter caused by incoherent scatter from free electrons and ions in the ionosphere. This paper will demonstrate a method to compute the additional noise due to reflection of the radar signal from the ionosphere. The approach suggested by Crane [1] is to use the parameters of average cross section per unit volume and the cross section spectral width to describe the incoherent scatter from the ionosphere. Plots of these two parameters as a function of height above earth are provided.

The next section will discuss the cross section per unit volume and doppler frequency spread (spectral width). The following section will discuss how to determine the noise at the radar receiver due to the ionospheric reflection.

2.0 CROSS SECTION PER UNIT VOLUME AND DOPPLER FREQUENCY SPREAD

The cross section per unit volume is a function of electron density and has units of m^2 per m^3 (resultant units of m^{-1}). Figure 1, from Crane [1], contains the cross section per unit volume for altitudes up to 2000 km. The curves are valid for a frequency of 9 GHz which was selected as the most likely frequency for the orbital debris radar system. The left curve applies for summer solstice, sunrise (near minimum of the production of electron density), and sunspot minimum. The right curve applies for equinox, noon (near peak of the production of electron density), and sunspot maximum.

The cross section spectral width is a function of the particle (electrons and ions) velocity distribution. It may also be thought of as a doppler shift of the transmitted frequency. Figure 2, from Crane [1], contains the cross section spectral width for altitudes up to 2000 km. The curves are valid for 9 GHz and the two conditions of electron density indicated above are represented by two different curves.

91 9 4 150

91-09707

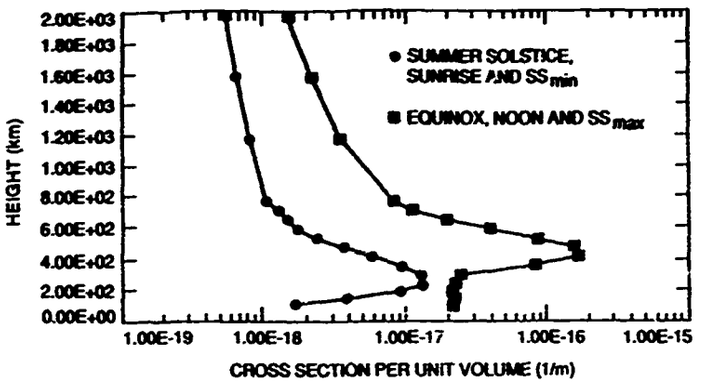


Figure 1. Average Cross Section per Unit Volume - Incoherent Scatter at 9 GHz

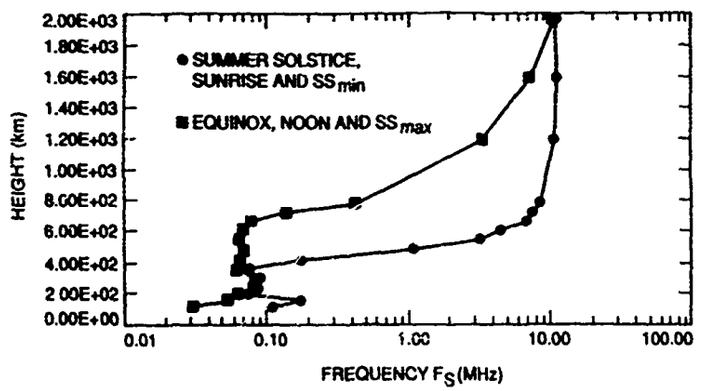


Figure 2. Cross Section Spectral Width - Incoherent Scatter at 9 GHz

To account for solar storms or periods of low solar activity the values in Figures 1 and 2 may differ by a factor of two. The 1979 International Reference Ionosphere [2], was used for generating the curves.

3.0 IONOSPHERIC NOISE AT THE RADAR RECEIVER

The cross section per unit volume may be interpreted as the reflectivity per linear meter along the antenna boresight, Brunstein [3]. Figure 3 provides an illustration of the concept. The radar transmitter power is assumed to be concentrated in the antenna beamwidth. The transmitter illuminates a volume of the ionosphere that is centered around the object in low earth orbit. The height of this volume is one half of the distance that light travels during the radar pulse period. The cross section per unit volume times the total illuminated volume yields the effective cross section of the ionosphere for this particular antenna and pulse height.

$$\beta = \beta_V * V_T \qquad (1)$$

where, β = effective cross section of ionosphere
β_V = cross section per unit volume
V_T = volume of illuminated ionosphere

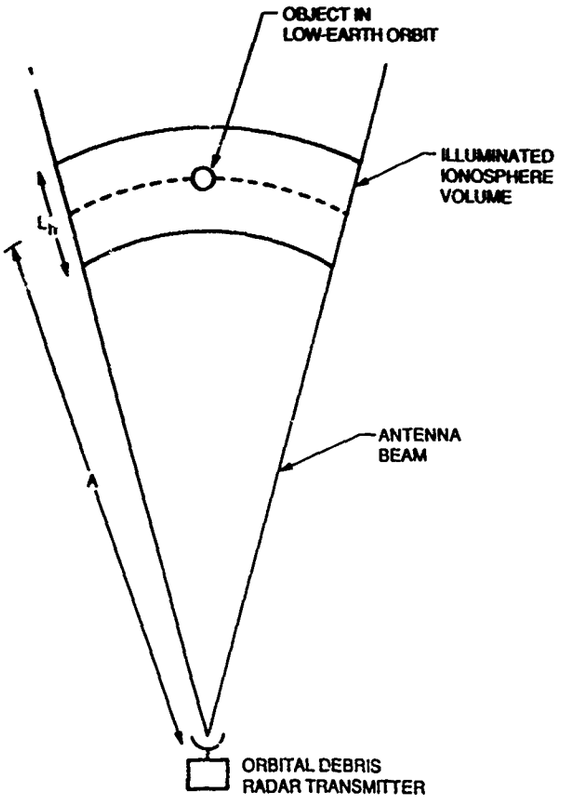


Figure 3. Radar Transmission Diagram

However, the ionospheric scattered power relative to the total transmitter power is the parameter of interest. If the pulse height (distance that light travels during the period of a radar pulse) is small then the illuminated volume is equal to the area of the antenna beam (at the altitude of the object in low earth orbit) times half of the pulse height.

$$V_T = A_T * L_h \tag{2}$$

where, V_T is from equation 1
A_T = cross section area of antenna beam at the altitude of the object in low earth orbit
L_h = one half of the pulse height
= (c*T)/2
where, c = speed of light = $3*10^8$ meters/sec
T = radar pulse width (sec)

From equations 1 and 2, it may be determined that the cross section per unit volume times half the pulse height is equivalent to the effective cross section of the ionosphere divided by the cross section area of the antenna beam at the altitude of the object in low earth orbit.

$$\beta_v * L_h = \beta/A_T \tag{3}$$

where, β and β_V are from equation 1
L_h and A_T are from equation 2

It is assumed that the total transmitter power passes through the area, A_T; and furthermore, that the area is evenly illuminated. Therefore, it is determined that the ionospheric scattered power relative to the transmitter power is equal to the effective cross section of the ionosphere divided by the cross section area of the antenna beam at the altitude of the object in low earth orbit.

$$\beta_v * L_h = \beta/A_T = P_{RI}/P_T \tag{4}$$

where, β_v, L_h, β, and A_T are from equation 3
P_{RI} = power reflected isotropically from the ionosphere
P_T = transmitter power

These equations require many assumptions and are, therefore, approximate. They are useful for a first order calculation of the reflected power from the ionosphere.

It was shown in section 2 that the cross section per unit volume is a function of height. The variation with height is illustrated in Figure 1. Therefore, equation 4 must be modified to calculate the ionospheric reflected power. The fraction of transmitter power that is reflected isotropically at a given altitude is the integral of the scattering cross section per unit volume over half of the pulse length centered around the altitude of the object in low earth orbit.

$$P_r = \int_{A-L_h/2}^{A+L_h/2} \beta_v * dh \tag{5}$$

where, P_r = fraction of transmitter power that is reflected isotropically at a given altitude
β_v and L_h are from equation 3
h = height (meters)
A = altitude of object in low earth orbit

This integral may be approximated with the trapezoidal integration method.

$$P_r = (\beta_{v0}/2 + \beta_{v1} + \ldots + \beta_{vn-1} + \beta_{vn}/2) * L_h/n \tag{6}$$

where, P_r is from equation 5
β_{vn} = n^{th} value of scattering cross section per unit volume from figure 1
L_h is from equation 3
n is the number of steps in the integration

Equation 6 is modified to show the reflected power from each incremental height.

$$P_r = P_1 + P_2 + \ldots + P_n \tag{7}$$

$$= [(\beta_{v0}+\beta_{v1})/2 + (\beta_{v1}+\beta_{v2})/2 + \ldots + (\beta_{vn-1}+\beta_{vn})/2] * L_h/n$$

where, P_r is from equation 5
P_n = power from n^{th} increment
β_{vn}, L_h, and n are from equation 6

Table 1 contains calculations of fractional ionospheric power based on equation 7. The altitude increment is 50 km. For example, if the radar pulse height is between 200 and 250 km then the fraction of transmitter power that is reflected isotropically from the ionosphere at an altitude of 225 km is $1.13*10^{-12}$. The procedure is to add the incremental powers over the total pulse height centered around the altitude of interest. For example, if the radar pulse height is between 450 and 550 km then the fraction of transmitter power that is reflected isotropically from the ionosphere at an altitude of 500 km is $(8.50 + 6.08)*10^{-12}$.

Height (km)	Fractional Reflected Power ($*10^{-12}$)	Bandwidth ($*10^{6}$) (Hz)
200	1.13	.073
250	1.18	.087
300	2.15	.082
350	5.05	.068
400	8.00	.064
450	8.50	.066
500	6.08	.065
550	3.05	.066
600	1.55	.072
650	.900	.101
700	.560	.193
750	.435	.355
800	.375	.525
850	.330	.675
900	.295	.875
950	.265	1.20
1000	.238	1.55
1050	.215	1.95
1100	.198	2.55
1150	.180	3.10
1200	.165	3.45
1250	.155	3.85
1300	.148	4.30
1350	.143	4.75
1400	.135	5.10
1450	.125	5.55
1500	.118	6.10
1550	.113	6.75
1600	.108	7.45
1650	.103	7.95
1700	.098	8.40
1750	.093	8.80
1800	.089	9.40
1850	.086	10.2
1900	.084	11.0
1950	.081	11.7

Table 1. Ionospheric Reflected Power - Equinox, Noon, and Sunspot Maximum

The cross section spectral width (bandwidth) of the reflected power is a function of height. This is illustrated in Figure 2. This data is included with the incremental reflected power in Table 1. The bandwidth information enables the calculation of the incremental spectral density of reflected power at each height. It may be assumed that the spectral density is uniform over the bandwidth, to a first approximation. The values in Table 1 occur during equinox, noon, and sunspot maximum. Table 2 contains similar data during summer solstice, sunrise, and sunspot minimum. The incremental spectral density is computed in equation 8.

$$S_n = P_n/F_{sn} \quad (Hz^{-1}), \text{ for } f < F_{sn}$$
$$= 0 \quad , \text{ elsewhere} \tag{8}$$

where, S_n = incremental spectral density at the n^{th} altitude interval
P_n is from equation 7
F_{sn} = bandwidth of reflected power at n^{th} altitude interval (from Table 1 or 2)

Height (km)	Fractional Reflected Power ($*10^{-12}$)	Bandwidth ($*10^{6}$) (Hz)
200	.550	.077
250	.650	.087
300	.555	.082
350	.400	.107
400	.280	.285
450	.188	.915
500	.135	2.30
550	.103	3.75
600	.085	5.25
650	.073	6.70
700	.063	7.60
750	.058	8.10
800	.055	8.40
850	.053	8.80
900	.049	9.05
950	.047	9.20
1000	.046	9.35
1050	.045	9.70
1100	.043	10.5
1150	.042	11.3
1200	.041	11.6
1250	.040	11.7
1300	.039	11.8
1350	.037	11.9
1400	.036	12.0
1450	.035	12.0
1500	.034	12.0
1550	.033	12.0
1600	.033	12.0
1650	.032	12.0
1700	.031	12.0
1750	.030	12.0
1800	.030	12.0
1850	.029	12.0
1900	.028	12.0
1950	.027	12.0

Table 2. Ionospheric Reflected Power - Summer Solstice, Sunrise, and Sunspot Minimum

The total spectral density is the summation of these incremental parts. These incremental spectral densities will be two-valued (a constant positive value over a frequency interval and zero outside of that interval) functions of frequency and the total spectral density could be complicated, depending on the bandwidths at each interval.

$$S_T(f) = S_1 + S_2 + \ldots + S_n \tag{9}$$

where, $S_T(f)$ = resultant power spectral density of the ionospheric reflected power
S_n is from equation 8

The fractional ionospheric power in the radar receiver bandwidth is computed from equation 10.

$$P_I = \int_0^{B_R} S_T(f)*df \tag{10}$$

where, P_I = fractional ionospheric power in receiver bandwidth
$S_T(f)$ is from equation 9
B_R = radar receiver bandwidth
f = frequency (Hz)

This integral may be approximated with trapezoidal integration in a manner similar to the integral in equation 5.

$$P_I = (S_{T0}/2 + S_{T1} + \ldots + S_{Tn-1} + S_{Tn}/2)*B_R/n \quad (11)$$

where, P_I and B_R are from equation 10

S_{Tn} = n^{th} value of resultant power spectral density from equation 9
n = the number of steps in the integration

As an example, assume that an object is located at an altitude of 525 km. The orbital debris radar transmits a pulse of 1 milli-second in duration. From equation 2, half of the pulse height is 150 km. Therefore, an ionospheric reflection could be expected from 450 to 600 km. The fraction of transmitter power that is reflected isotropically from the ionosphere may be computed from equation 5. The integral is approximated with equation 7. Three intervals are chosen. Table 1 contains the reflected power from each interval during equinox, noon, and sunspot maximum.

$$P_1 = 8.5*10^{-12} \quad (12)$$
$$P_2 = 6.08*10^{-12}$$
$$P_3 = 3.05*10^{-12}$$

The bandwidth of each incremental reflected power can be obtained from Table 1, also. These bandwidths may be used to compute incremental power spectral density in equation 8.

$$F_{s1} = 66 \text{ kHz} \quad (13)$$
$$F_{s2} = 65 \text{ kHz}$$
$$F_{s3} = 66 \text{ kHz}$$

$$S_1 = 1.29*10^{-16} \quad f < 66 \text{ kHz}$$
$$S_2 = 9.35*10^{-17} \quad f < 65 \text{ kHz}$$
$$S_3 = 4.62*10^{-17} \quad f < 66 \text{ kHz}$$

A matched filter receiver for this pulse would have a bandwidth of 1 kHz, Skolnik [4]. The resultant power spectral density of the ionospheric reflected power is computed from equation 9.

$$S_T = 2.687*10^{-16} \ (Hz^{-1}), \ f < 65 \text{ kHz} \quad (14)$$

The fractional ionospheric power in the radar receiver bandwidth is computed with equation 10.

$$P_I = 2.687*10^{-13} \quad (15)$$

In contrast, the target power relative to the transmitter power at an altitude of 525 km may be computed, Skolnik [4].

$$P_{tg} = G*x/(4*\pi*R^2) \quad (16)$$

where, P_{tg} = target power relative to transmitter power at the range indicated
x = target cross section ($meters^2$)
G = numerical value of antenna gain
R = range of target (meters)

For a target with a diameter of 1 cm (cross section of 0.78 cm^2), antenna gain of 54 dBi, and a range of 525 km the target power relative to transmitter power is $5.70*10^{-12}$. Hence the target signal to ionospheric noise (equation 15) ratio is 13.3 dB.

4.0 SUMMARY AND CONCLUSIONS

A method to calculate the incoherent scatter of a radar signal from the ionosphere has been shown. This method uses the concept of the ionospheric scattering cross section per unit volume. The doppler frequency spread of the reflected power is shown, also. Using these concepts, the noise due to the ionosphere at the orbital debris radar receiver is computed. From the example calculation, it is feasible that an orbital debris radar can detect objects with a diameter of 1 cm in the presence of ionospheric reflection at a frequency of 9 GHz.

The research described in this paper was carried out by Jet Propulsion Laboratory, California Institute of Technology, under a contract with the National Aeronautics and Space Administration.

5.0 REFERENCES

1. Crane, R.K., Correspondance with R.K. Crane.

2. International Reference Ionosphere - IRI 79, Report UAG-82, World Data Center A for Solar-Terrestrial Physics, NOAA, Boulder, CO.

3. Brunstein, S., "Atmospheric Attenuation and Noise, and Ionospheric Reflection at 9 GHz", IOM 3346-89-008, Jet Propulsion Laboratory, Pasadena, CA, Jan. 19, 1989 (an internal document).

4. Skolnik, Merrill I., "Introduction to Radar Systems", McGraw-Hill Inc., New York, 1980.

AD-P006 295

UPDATED CLIMATOLOGICAL MODEL PREDICTIONS OF IONOSPHERIC AND HF PROPAGATION PARAMETERS

M.H. Reilly, F.J. Rhoads, J.M. Goodman
Ionospheric Effects Branch, Space Science Division
E.O. Hulburt Center for Space Research
Naval Research Laboratory, Washington, DC 20375-5000

M. Singh
Sachs-Freeman, Inc.
Landover, MD 20785

ABSTRACT

Several climatological models predict ionospheric parameters. A few also predict HF circuit parameters, such as the maximum usable frequency (MUF). Using conventional sunspot number inputs, we compare ICED, RADARC, and IONCAP model prediction performances for the F2 layer critical frequency foF2. Monthly median, 5-day average, and daily sunspot numbers, which are based on the 10.7 cm. solar flux, are used for this purpose, along with the ICED effective sunspot number. Relating prediction points to ionosonde data of comparable corrected geomagnetic latitude, in conjunction with an effective sunspot number, improves prediction accuracy. Near-real-time (NRT) updates of an ionospheric model from observations at a single station is found to dramatically improve predictions of foF2 out to distances of thousands of kilometers, as compared with the use of a global sunspot number. Similar performance for other ionospheric and HF circuit parameters is indicated. For the purpose of NRT predictions of the entire electron density profile the use of two distinct sunspot numbers for the bottomside, coupled with one or more parameters for the topside, is suggested for model updates from a single station.

1. INTRODUCTION

A number of climatological ionospheric models have been developed, which are based on hourly, monthly median values extracted from an archived data base. The data base encompasses a wide range of locations, times of day, seasons, and levels of magnetic and solar activity. These dependencies are included in the model. The ionospheric models are often mated to propagation models for computation of radio performance. However, NRT ionospheric specification from such a model is limited. Substantial fluctuations from a climatological description occur routinely. Further, the input parameters of a typical model, notably the sunspot number, are oversimplified and usually only loosely connected with the ionospheric phenomena of interest. In the past, the input sunspot number was determined from solar observations and suitably averaged over time. A significantly improved sunspot number has been provided by the U.S. Air Force (AFGWC) in recent years for use with the AWS-NOAA ionospheric prediction model ICED [Tascione et al, 1988]. It is a daily, global sunspot number which is determined from a least-squares fit of the ICED model to the last five days of foF2 data from a global network of vertical-incidence sounders. Further improvement is expected from a more strategic use of ionospheric data. For example there is an increasing tendency for AFGWC to supply region-specific sunspot numbers for ICED-model updates, based on foF2 data for the region, as well as sunspot numbers which change several times during the day (e.g., every hour or six hours) and utilize different averaging intervals (e.g., six, twelve or twenty-four hours). A similar approach of using a sunspot number which depends on time and location, as determined from a network of oblique-incidence sounders, has been developed for NRT prediction of shortwave propagation parameters [Reilly and Daehler, 1986]. A review of this and other relevant material has been given by Goodman and Reilly [1988].

In Section 2 we start with foF2 prediction results from three prominent ionospheric models, using conventional, global sunspot numbers. These are the ionospheric models in IONCAP [Lloyd et al, 1978], RADARC [Thomason et al, 1979], and ICED itself. Results will be interpreted in light of model features and limitations of global sunspot numbers. In Section 3 we report the use of alternative approaches for predictions of ionospheric parameters, based on updates and

91 9 4 149

91-09706

observations from a single station. Substantial improvements will be noted over large spatial scales, although some of the techniques have not yet been optimized, and this will bring in the notion of spatial correlation lengths. Implications of this work will be discussed in the final section with a view toward improved HRT prediction of electron density profiles at remote locations.

2. IONOSPHERIC MODELS AND GLOBAL SUNSPOT NUMBERS

Sunspot numbers (SSNs) are often obtained from the daily observed 10.7 cm solar flux (F). We use a well-known relation

$$SSN = 1.08F - 62 \qquad (1)$$

to find daily SSN inputs to the ionospheric models in IONCAP, RADARC, and ICED, in order to assess their performance for foF2 predictions at particular stations for a particular month. Ground-truth data for comparison with predictions has been assembled in a data base of the ionospheric parameters foF2, fmin, foEs, and M(3000) from vertical-ionosondes for up to 45 northern hemisphere stations and up to 7 southern hemisphere stations. Also contained in the data base is total electron content (TEC) data from polarimeters for up to 12 northern hemisphere stations. New data is obtained once per hour usually, and the total period covered by the data base is January 1981 through December 1989. For the sake of illustration, we selected foF2 data for July 1982. The distribution of magnetic and solar activity for this month is quite broad, as will be seen in detail later. The prediction error for a particular hour of a particular day is calculated as a percentage according to

$$\text{error}(\%) = 100(\text{foF2p} - \text{foF2m})/\text{foF2m} , \qquad (2)$$

and statistics of the error are compiled for three stations. The results are shown in Table 1 for Manila, Boulder, and Narssarsuaq, which represent low-, mid-, and high-latitudes, respectively. Corrected geomagnetic latitude (CGLAT) and geographic longitude (LON) are indicated for each station. The first, second, and third prediction error columns give the average prediction error for the month, the percent of all the errors which exceed 25%, and the percent which exceed 50%, respectively. RADARC and IONCAP are based on the same ITS-78 coefficients for foF2, which probably explains their comparable performance in Table 1. RADARC might have been expected fare slightly better, because of its corections for magnetic activity and high latitude effects, but this is not noticeable. ICED contains a more recent version of the foF2 coefficients, known as the URSI coefficients [Rush et al, 1989], and this may account for its somewhat superior performance in Table 1. It is surprising that ICED, with its high latitude corrections, does not do signicicantly better than IONCAP at Narssarsuaq. Definitive comparisons are clouded, however, by the ambiguity of the sunspot number, which introduces a substantial bias error in addition to that introduced by the day-to-day variability about the monthly-median behavior.

A variety of SSNs are tried with IONCAP for the July 1982 Boulder data. The results are shown in Table 2. The SSNs are (1) the monthly mean of the daily SSN from Equation (1), (2) the mean of SSN values from the last five days, (3) the daily SSNs used in Table 1, and (4) the five-day average effective sunspot number (SSNE) supplied by AFGWC for use with the ICED model. Of the various SSNs based on the 10.7 cm flux the monthly mean performs best, but there is a significant improvement with the use of the ICED value SSNE. This results from using ionospheric data for ionospheric predictions.

TABLE 1. foF2 PREDICTION ERRORS FOR JULY 1982, BASED ON DAILY OTTAWA 10.7 CM FLUX VALUES

Location	Model	foF2 Prediction Error Results		
		Average(%)	% > 25%	% > 50%
Manila	IONCAP	27.2	44.2	11.1
CGLAT=6.7N	RADAR-C	25.2	39.9	9.0
LON=121E	ICED	25.7	40.1	10.0
Boulder	IONCAP	27.1	36.9	15.0
CGLAT=49.1N	RADAR-C	28.4	33.4	19.0
LON=105.3W	ICED	24.9	31.6	11.9
Narssarsuaq	IONCAP	17.8	23.2	4.1
CGLAT=67.4N	RADAR-C	19.6	26.4	8.3
LON=45.4W	ICED	17.7	22.3	3.2

TABLE 2. foF2 IONCAP PREDICTION ERRORS AT BOULDER FOR JULY 1982 FOR VARIOUS SUNSPOT NUMBERS

Type of SSN	Average (%)	% > 25%	% > 50%
Monthly Mean (10.7 cm)	18.5	25.1	5.2
Mean from last 5 days (10.7 cm)	25.3	37.2	13.8
Daily (10.7 cm)	27.1	36.9	15.0
APGMC SSNE for ICED - last 5 days	15.2	15.1	2.3

3. NRT UPDATES AND PREDICTION OF THE REMOTE IONOSPHERE

We begin with the definition of "actual" sunspot numbers SSNAC, one for each foF2 determination at each sounder station, which produce agreement of the ionospheric model with an foF2 observation. These values are determined for thirty vertical-incidence sounders in the data base for July 1982 with respect to the ICED model. Geographic and geomagnetic coordinates of these stations are given in Table 3. Another input parameter of the ICED model is the Q index, which is determined from the latest Kp value, according to the relation

$$Q = \begin{cases} = 3Kp & (0 \leq Kp \leq 1) \\ = Kp+2 & (1 < Kp \leq 6) \\ = 8 & (6 < Kp) \end{cases} \qquad (3)$$

used in ICED. This determines auroral boundaries and is involved in high-latitude corrections, but the bulk of magnetic activity and other effects at the sounder location is included in the determination of SSNAC.

Figure 1a shows the 10.7 cm flux, SSNE, and the magnetic activity parameters Kp and Ap for July, 1982. A wide range of solar and magnetic activity levels is exhibited. The SSNAC results for eighteen of the thirty stations are arranged in order of increasing sounder CGLAT from the bottom of the Figure 1d to the top of Figure 1b. There are similarities in the the appearance of the SSNAC results for adjacent stations of comparable CGLAT, despite large longitude separations in some instances. The beginnings of organized behavior, which tends to be be negatively correlated with Kp variations, is not readily apparent at low CGLAT values, but becomes so from Vandenberg (CGLAT = 40.6) on up. For example, the similarities between Vandenberg, Poitiers, and Kiev are striking, despite a longitude separation of 150 degrees. For stations with CGLAT values within about six degrees it is tempting to find and test model update schemes, such that an update determined at one station is capable of successfully predicting ionospheric behavior at the other stations, even though they may be separated from the update point by great distances. Initial results are presented next.

The first update technique involves a linear regression analysis of SSNAC on the variable Kp. We select Dourbes as the control point station and determine linear regression coefficients for each hour of the day from the July Dourbes data. We then use a value of SSN based on these coefficients,

$$SSN \cong S_0 + S_1 K_p \ , \qquad (4)$$

as input to ICED to make predictions of foF2 at Boulder, Wallops Island, Slough, Poitiers, Kiev, Tomsk, and Sverdlovsk. These prediction points are all within six degrees in CGLAT from Dourbes, although longitude differences with Dourbes range up to 110 degrees. At a particular time for prediction, the coefficients for that time, as determined from the Dourbes data, are inserted into Equation (4) to determine SSN for predictions. The only NRT aspect of this technique is the use of the latest K_p value. We shall see that prediction error for foF2 is uniformly improved, relative to that from the SSNE predictions, by 10-28%. This method is somewhat similar to that of Wrenn and Rodger [1989], although their method is not based on the sunspot number in conjunction with an ionospheric climatological model. In the future we expect to find improvement, as they did, by using a time weighted accumulation value for Kp, instead of the latest value. This accounts for inertia in the ionospheric response to magnetic activity perturbations.

TABLE 3. GEOGRAPHIC AND CORRECTED GEOMAGNETIC COORDINATES OF VERTICAL-INCIDENCE SOUNDERS

STATION	LAT.	LON.	CGLAT	CGLON
MANILA	14.60	121.00	6.74	191.58
TAIPEI	25.00	121.20	17.64	192.03
MAUI	20.80	-156.50	21.55	270.27
KOKOFUNJI	35.70	139.50	28.22	210.12
NICOSIA	35.17	33.28	28.45	104.98
ASHKHABAD	37.90	58.30	32.66	129.98
VANDENBERG	34.73	-120.57	40.62	302.10
ALMA ATA	43.25	76.92	37.97	148.72
POITIERS	46.57	0.35	42.19	78.35
BOULDER	40.00	-105.30	49.14	317.93
WALLOPS IS.	37.90	- 75.50	49.83	358.27
KIEV	50.72	30.30	46.10	104.55
IRKUTSK	52.50	104.00	47.00	176.00
MOSCOW	55.50	37.30	51.00	112.01
DOURBES	50.10	4.60	46.25	82.81
SLOUGH	51.13	- 0.57	47.87	79.05
MAGADAN	60.00	151.00	53.21	218.07
OTTAWA	45.40	- 75.90	57.23	358.57
UPPSALA	59.80	17.60	59.26	97.02
SVERDLOVSK	56.70	61.10	51.99	134.06
TOMSK	56.50	84.90	51.50	157.09
GOOSE BAY	53.50	- 60.33	63.08	22.87
YAKUTSK	62.00	129.60	55.71	199.74
COLLEGE	64.90	-147.80	64.85	261.23
CAPE SCHMIDT	68.90	-179.50	64.20	235.02
KIRUNA	67.83	20.42	64.40	104.10
MURMANSK	68.00	33.00	63.88	113.97
NARSSARSUAQ	61.20	- 45.45	67.45	44.45
SALEKHARDT	66.55	66.70	61.59	141.51
DIXON	73.50	80.40	67.84	155.68

The second approach involves a more traditional NRT update at Dourbes, similar to that used for prediction of MUFs on HF sky wave paths [Reilly and Daehler, 1986]. However, instead of using the latest value of SSNAC at Dourbes for predictions, we use an average of the last four SSNAC values. We want to average out local travelling ionospheric disturbance (TID) effects in the determination of SSNAC at Dourbes, which are spatially correlated on a scale of roughly 400 km, in order to improve our predictions out to thousands of kilometers. Since prevailing TID oscillations have time scales on the order of an hour or less, it was felt that the four-hour sliding average was sufficient, although probably not optimum.

The results for all the July data of both update methods at Dourbes, relative to predictions of foF2 from the global sunspot number SSNE, are shown in Figure 2. The use of SSNE with ICED gives prediction errors in the range of about 15%. Use of the linear regression method, based on Equation (4) gives uniform improvement out to several thousand km. The improvement over the SSNE results is in the range of 28% near Dourbes and 10% far from Dourbes. The four-hour sliding average method is indicated as the NRT curve in Figure 2. Evidently there is dramatic improvement near Dourbes, which diminishes for greater separations. These results are strongly reminiscent of those for predictions of total electron content departures from a predicted mean, based on updates at a single polarimeter station [Klobuchar, 1985; Klobuchar and Johanson, 1977]. A spatial correlation distance could be defined relative to an improvement level of 29%, and this was found to occur for midday TEC data at a distance of about 2900 km in an east-west direction, and about 1800 km in a north-south direction. This result depended only weakly on season. The corresponding result for the NRT curve in Figure 2 is a 29% improvement at a ground range of about 2850 km, although significant improvement is noticed out to a range of about 4500 km.

Corresponding results are obtained in Figure 3 for the percentage of occurrence of foF2 prediction error in excess of 25%. A greater improvement of the update methods over SSNE results is noticed here, particularly with the NRT method. This indicates a significant advantage of the NRT method in drastically reducing the incidence of large prediction errors within the aforementioned spatial correlation distance.

About 50% of the ionosonde stations in the data base report M(3000), in addition to foF2. This is the ratio of the MUF on a 3000 km path to foF2, and it is obtained by sliding a transmission curve for a 3000 km path until it is just tangent to the virtual height curve for the F2 layer. It is a very important parameter in the ionospheric model we have been considering, since it largely determines the height and thickness parameters for the F2 layer.

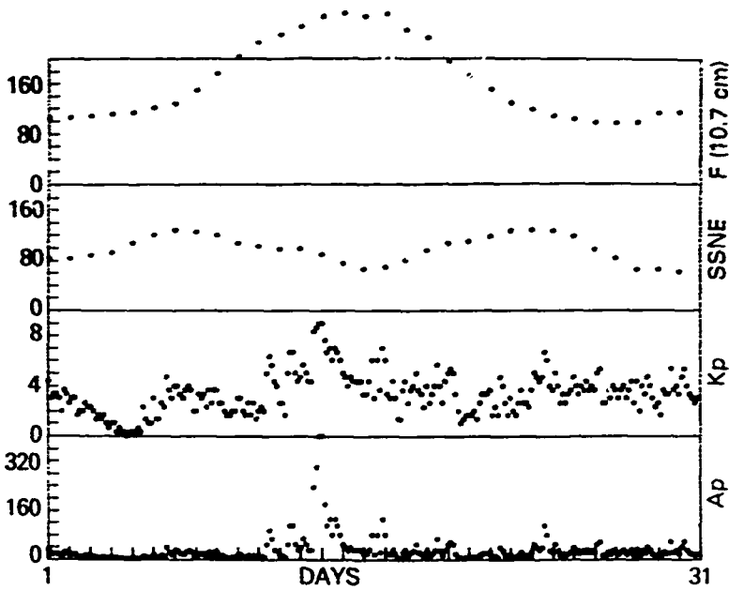
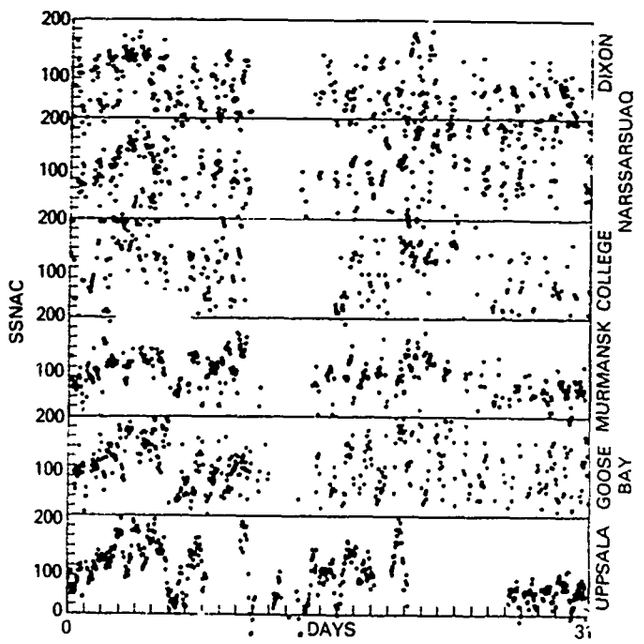


Fig. 1a. Magnetic (K_p and A_p) activity, solar activity (F(10.7)), and AFGWC sunspot number SSNE variation for July, 1982



(1b)

Fig. 1b-1d. SSNAC values for 18 ionosonde stations for July 1982 in order of increasing corrected geomagnetic latitude from bottom to top

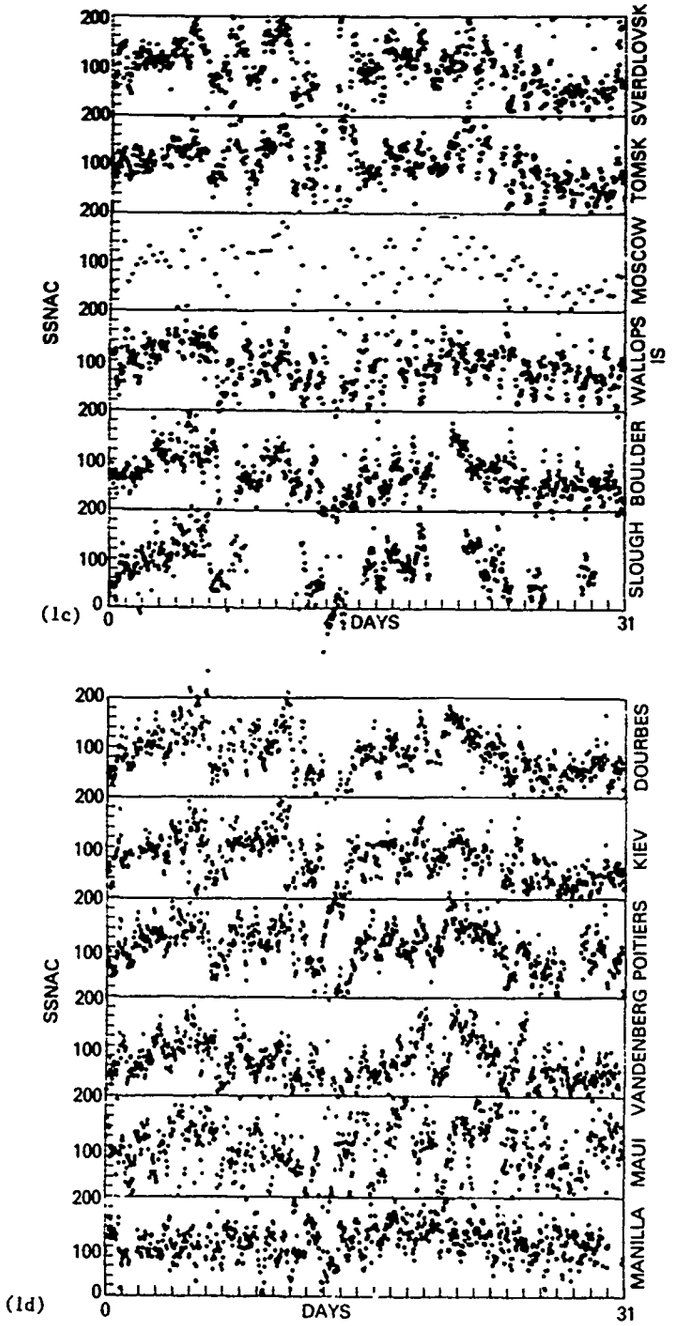


Fig. 1b-1d. SSNAC values for 18 ionosonde stations for July 1982 in order of increasing corrected geomagnetic latitude from bottom to top

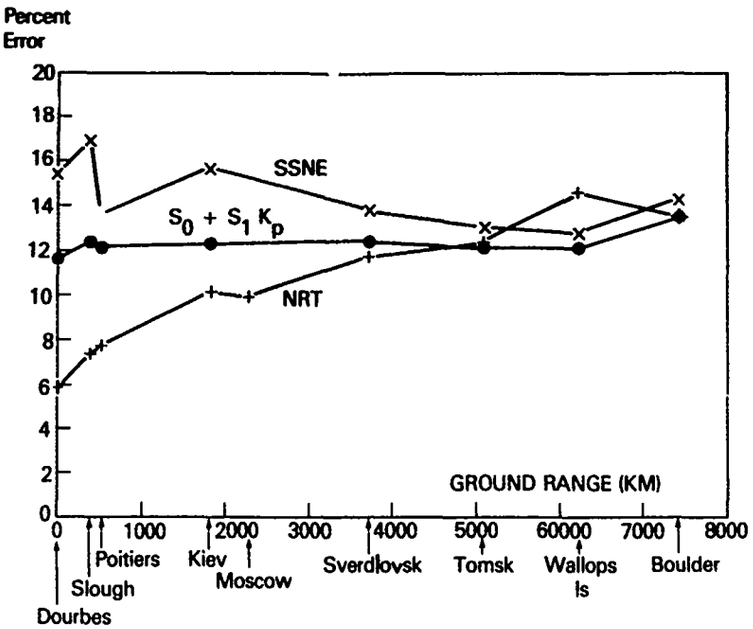


Fig. 2. ICED prediction error for foF2 vs. sunspot number update method

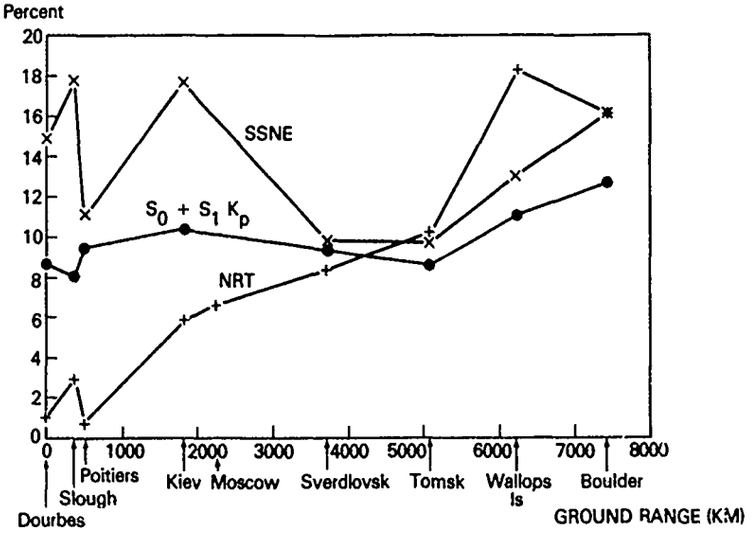


Fig. 3. Percentage of foF2 prediction errors which exceed 25% vs. sunspot number update method

The July 1982 M(3000) prediction error results for various sunspot numbers are shown in Figure 4. The SSNE and NRT(foF2) SSNs are the same as used in Figure 3. Also shown are the results for the daily SSNs inferred from the 10.7 cm flux and for the SSNAC values at each station, which would in fact reduce the prediction error curve for foF2 to zero. However, the prediction error is worst for M(3000) in this case, followed by the NRT(foF2) update at Dourbes. The results are better for predictions based on SSNE and the 10.7 cm flux. Hence, there is no indication of spatially correlated effects for an update based on foF2 measurements, or that it is especially useful for prediction of M(3000). It is gratifying, however, that the prediction error is only in the range of 5-7%. This translates to about an 8-11% error for hmF2, the height at the maximum density of the F2 layer. The best results in Fig. 4 are obtained for the NRT sunspot number based on M(3000) data, again with a similar spatial correlation effect. For the purpose of predicting the entire electron density profile, it therefore seems advisable to determine a separate SSN value, which is based on M(3000) measurements.

If it was feasible to combine TEC and ionosonde data for a single station, there would be potential for great improvement in NRT prediction of the entire electron density profile at remote sites, at least for sites which are not far apart in corrected geomagnetic latitude. The ionosonde data could be used for model updates to specify the bottomside profile, up to the F2 layer peak, and the TEC data could be used to specify one or more parameters in the topside profile model. Other data sources could also be used. To optimize this scheme, an ionosonde control point and an ionospheric pierce point for the TEC data should be within a few hundred kilometers. An improved electron density profile specification would naturally enhance the validity of propagation calculations for radio systems.

A further test of the benefits of NRT updates is provided by the oblique-incidence sounder data base of MUFs assembled in the SSL-BCT experiment during 29 November - 18 December 1982 [Reilly and Daehler, 1986]. We choose the 1500 km Erie, Colorado (near Boulder) - Fort Ord, California circuit for analysis. MUFs are predicted by IONCAP for various sunspot numbers and the NRT update method, based on IONCAP and the foF2 data at Boulder, Colorado. The results are shown in Table 4. In this case, we do not use the four-hour sliding average on SSNAC (Boulder), but it is still better than the SSNE determined for ICED, and much better than the SSNs based on the 10.7 cm solar flux.

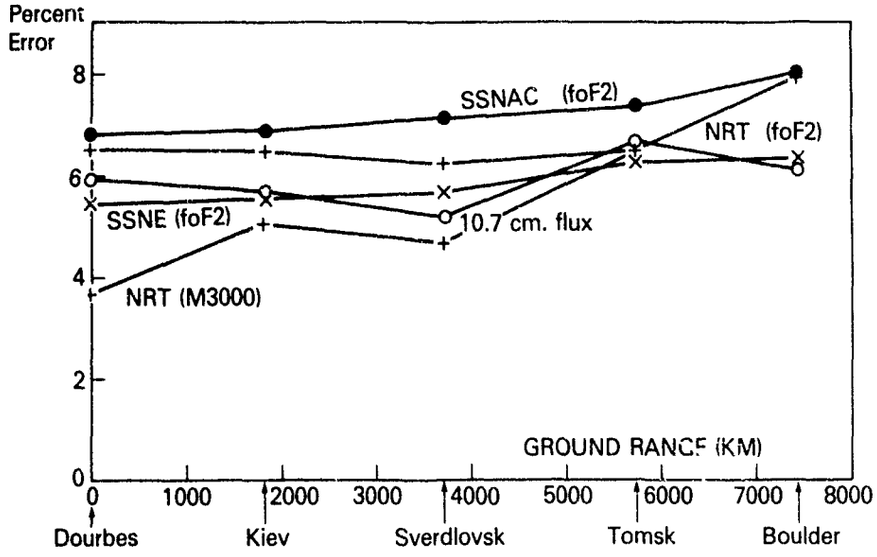


Fig. 4. ICED prediction error for M(3000) vs. sunspot number update method

TABLE 4. MUF IONCAP PREDICTION ERRORS FOR THE ERIE, COLORADO, TO FORT ORD, CALIFORNIA, CIRCUIT FOR 29 NOV-18 DEC 1982 FOR VARIOUS SUNSPOT NUMBERS

Type of SSN	Average (%)	% > 25%	% > 50%
Daily (10.7 cm)	32.0	48.4	17.0
Monthly Mean (10.7 cm)	29.6	43.1	15.7
SSNE	16.3	19.6	4.6
SSNAC (Boulder)	14.0	15.1	0.7

4. DISCUSSION AND CONCLUSIONS

In the absence of ionospheric data in a region of interest, ionospheric specification normally uses a statistical, climatological ionospheric model with appropriate inputs. One important input to many of these models is the sunspot number, and it has been shown how this can be specified from data at a remote update station, based on an ordering by corrected geomagnetic latitude (CGLAT). The most dramatic performance improvement over previous methods comes from NRT updates, with some averaging of local TID effects. There is a requirement that prediction points and update points are not too far apart in this method, but this separation apparently extends out to 4000 km or so. There is, however, some indication that this distance can be extended even further for foF2 predictions by a method which exploits a strong dependence on magnetic activity. With regard to the NRT update method, we find that for specification of the entire electron density profile, there is an advantage to determining two sunspot numbers in conjunction with an ionospheric model, one for foF2 and the other for M(3000), both of which can be routinely determined from vertical-incidence sounder data. These are the basis for a much-improved bottomside profile specification, in conjunction with the ionospheric model. The addition of approximately collocated TEC data, perhaps with assistance from other topside data sources, would be sufficient to specify one or more important parameters of the topside model, and the result should be greatly improved ionospheric specification out to large distances. Update stations could be deployed, each for ionospheric specification in a particular range of CGLAT values out to thousands of kilometers from the update point. Another caution, implied by the emphasis on CGLAT, is that update and prediction points should not be located in different ionospheric regions. For example, an update point in the subauroral trough may well be inappropriate for predictions at mid-latitudes or in the auroral region.

REFERENCES

Goodman, J.M. and M.H. Reilly, Shortwave propagation prediction methodologies, IEEE Trans. on Broadcasting, 34, 230-237, 1988.

Klobuchar, J.A., Ionospheric total electron content (TEC), Handbook of Geophysics and the Space Environment, Sec 10.9, AFGL, available from NTIS, ADA 167000, 1985.

Klobuchar, J.A. and J.M. Johanson, Correlation distance for mean daytime electron content, AFGL TR-77-0185, ADA048117, 1977.

Lloyd, J.L., G.W. Haydon, D.L. Lucas, and L.R. Teters, Estimating the performance of telecommunication systems using the ionospheric transmission channel, USACEEIA Tech. Rpt., vol. 1, EMEO-PED-79-7, Ft. Huachuca, AZ, 1978.

Reilly, M.H. and M. Daehler, Sounder updates for statistical model predictions of maximum usable frequencies on HF sky wave paths, Radio Sci., 21, 1001-1008, 1986.

Rush, C., M. Fox, D. Bilitza, K. Davies, L. McNamara, F. Stewart, and M. Pokempner, Ionospheric mapping: an update of foF2 coefficients, in ITU Telecomm. Journal, March, 1989.

Tascione, T.F., H.W. Kroehl, R. Creiger, J.W. Freeman, Jr., R.A. Wolf, R.W. Spiro, R.V. Hilmer, J.W. Shade, and B.A. Hausman, New ionospheric and magnetospheric specification models, Radio Sci., 23, 211-222, 1988.

Thomason, J., G. Skaggs, and J. Lloyd, A global ionospheric model, NRL Memo Rpt 8321, Washington, DC, AD-000-323, 1979.

Wrenn, G.L. and A.S. Rodger, Geomagnetic modification of the mid-latitude ionosphere: toward a strategy for the improved forecasting of foF2, Radio Sci., 24, 99-111, 1989.

AD-P006 296

UPGRADES FOR EFFICIENT 3D IONOSPHERIC RAY TRACING - INVESTIGATION OF HF NVIS EFFECTS

Michael H. Reilly
Naval Research Laboratory
Space Science Division
Ionospheric Effects Branch
Washington, DC 20375-5000

ABSTRACT

An earlier, three-dimensional (3D) ray-tracing program has been upgraded to include magnetic field effects, variable step sizes, improved propagation loss computation, and efficient homing and focusing computation algorithms. A combination of fourth-order Runge-Kutta and Adams methods, with error control, has been included. The program is applied to an investigation of high-frequency (HF) near-vertical-incidence-skywave (NVIS) effects. Model simulations of oblique ionograms are found to be consistent with experiment. Implications for NVIS geolocation are discussed.

1. INTRODUCTION

A three-dimensional ionospheric ray tracing method was introduced in a previous paper [Reilly and Strobel, 1988], hereafter denoted by RS. It used an analytic solution procedure for the ray path increments, which was based on a Taylor series expansion of the index of refraction about the beginning of the increment in local Cartesian coordinates. This propagation model was joined with a global, climatological ionospheric model [Thomason et al, 1979], which has been used in the RADARC program for over-the-horizon radar calculations. The propagation model properly included all ionospheric tilt effects. It also excluded artificial refraction effects which would arise from the ionospheric model's approximations for the height profile as a set of connected segments. Within each segment the electron density is a prescribed function of height, but one or more of the associated Taylor series coefficients is discontinuous at the boundaries between the segments. Care was taken to join the ray path increment solutions smoothly at these points.

The preceding ray tracing model did not go far enough to describe HF phenomena. Magnetic effects were neglected, which made the analytic solution procedure for the ray path increments possible. The program has now been rewritten to include magnetic effects, based on numerical solution procedures for these increments, although the procedure of expanding the electron density in a Taylor series about the beginning of the increment has been retained. This simplifies the function evaluations in the numerical solution, and is also the basis of another upgrade to the ray tracing model. Instead of fixed increment lengths, the Taylor series coefficients are used to determine the increment length, so as to conform to a prescribed rate of convergence of the Taylor series expansion. The size of the increment may also be modified by error control algorithms in the numerical solution. A third upgrade in the present program is the incorporation of homing for point-to-point calculations. A typical approach is to sequentially trace the primary ray, and then two displaced rays, associated with separate, slight deviations in the azimuth and elevation launch angles. The intercepts of these rays at the receiver height are used to approximate partial derivatives of latitude and longitude with respect to the launch angles, and these are used to correct the launch angles for the next homing iteration. In the present program, the partial derivatives are computed during the calculation of the primary ray, a more efficient process. A similar procedure for simultaneous computation of the partial derivatives was previously invoked for ionospheric focusing calculations [Nickisch, 1988]. The partial derivatives are also used in the present program in a revised calculation for spatial propagation loss. The above upgrades to the ray tracing model, and other improvements, will be discussed further in Section 2.

The program is applied to an investigation of HF near-vertical-incidence skywave (NVIS) effects in the Section 3. This is a fairly severe test of the model, since refraction effects are pronounced over this frequency range, and particularly so for NVIS propagation. The focus

91-09708

is of oblique ionogram results for the Driver, Virginia to NRL path with range 227 km. Oblique ionograms computed for the RADARC ionospheric model [Thomason et al, 1979] and the ICED ionospheric model [Tascione et al, 1988] are compared with experiment. Comparison of various numerical techniques with each other and with the rather dissimilar analytic solution technique [Reilly and Strobel, 1988] for the case of no magnetic field helps to establish their validity and relative advantages. Results are calculated for single-site location in various ionospheric approximations. The importance of ionospheric tilts and magnetic effects in the NVIS geolocation of transmitters is indicated. Section 4 closes with a discussion and conclusions.

2. RAY TRACING UPGRADES

In the region of validity for ray tracing, where the wavelength and inverse frequency are much less than length and time scales of variation in the ionosphere, the phase at a space-time point $\underline{X}$ of a wave emitted by a transmitter at the origin is

$$U = \int_0^{\underline{X}} \underline{K} \cdot d\underline{X} \tag{1}$$

where $\underline{K}$ and $\underline{X}$ are four-vectors in Cartesian coordinates given by

$$\underline{K} = (k_1, k_2, k_3, -\omega) \quad , \quad \underline{X} = (x_1, x_2, x_3, t) \tag{2}$$

One recognizes the three spatial coordinates of the wave vector k_i and position vector x_i, as well as the angular frequency ω and time t. The eight variables, K_λ and X_λ (λ = 1-4), are related by a local dispersion relation of the form

$$F(K_\lambda, X_\lambda) = 0 \tag{3}$$

Lighthill [1965] uses only the relations

$$K_\lambda = K_\lambda(X_\mu) = \partial U/\partial X_\lambda \quad , \quad \partial F/\partial X_\mu = 0 \qquad (\lambda,\mu = 1\text{-}4) \tag{4}$$

in a simple derivation of the eight ray trace equations

$$dX_\lambda/d\tau = \partial F/\partial K_\lambda \quad , \quad dK_\lambda/d\tau = -\partial F/\partial X_\lambda \tag{5}$$

for some parameter τ. In the present version of the ray trace program, the medium is considered stationary, so that frequency is constant in the medium, and the preceding eight equations are reduced to six:

$$dx_i/dt = -(\partial F/\partial k_i)/(\partial F/\partial \omega) \quad , \quad dk_i/dt = (\partial F/\partial x_i)/(\partial F/\partial \omega) \quad (i = 1\text{-}3) \tag{6}$$

To this is added an equation for the phase path delay

$$dP/dt = k_i(dx_i/dt)/\omega \tag{7}$$

The phase path length is cP. The expressions used for F are:

$$F = k^2 - \mu^2\omega^2/c^2 = 0 \quad , \quad \mu^2 = 1 - X/D \tag{8}$$

$$X = \omega_N^2/\omega^2 \quad , \quad \omega_N^2 = Ne^2/m\epsilon_0 \quad , \quad D = 1 - T \pm (T^2 + Y_L^2)^{1/2} \tag{9}$$

$$T = (Y^2 - Y_L^2)/(2(1 - X)) \quad , \quad Y_L = Y_j k_j/(k_i k_i)^{1/2} \quad , \quad Y_j = e\beta_j/m\omega \tag{10}$$

The index of refraction μ of (8) and (9) can take two values; the + sign is defined to be the ordinary (O) mode, whereas the - sign is defined to be the extraordinary mode. In these equations c is the speed of light, β_j is a component of the earth's magnetic field (in webers/m^2), and other parameters in these equations are as defined in RS. The summation convention for repeated subscripts is assumed throughout the paper, unless otherwise specified. To this point the formalism is similar to Cummack's [1980].

The equations (6)-(7) for a ray path increment are solved in local SEZ (south, east, and upward) Cartesian coordinates, where X is expanded in a Taylor series about the beginning of the ray path increment as

$$X = X_0 + A_i x_i - B_i x_i^2 \tag{11}$$

The $\underline{A}$ coefficient vector has the components (N_x, N_y, A) and the $\underline{B}$ coefficient vector has components (-A/2r, -A/2r, B), where the entities in parentheses are defined in RS. The expansion in (11) is complete to second order only in the case of spherical symmetry, or no tilts ($N_x = N_y = 0$). It is not difficult to obtain the analytic no-field solutions of RS from these equations. As indicated in the earlier paper, the A_i, B_i coefficients in (11) are obtained from specification of ionospheric parameters on a lat.-lon. grid of points, triangular interpolation to the beginning of the ray path increment, and the functional representation of the height profile in the ionospheric model. The specification of the magnetic field is handled similarly. The International Geomagnetic Reference Field (IGRF) spherical harmonic expansion determines the values of β_{iL} in

$$Y_i = (2\pi/\omega) \sum_{L=1}^{8} (r/r_E)^{-(2+L)} \beta_{iL} \quad , \tag{12}$$

on the same lat.-lon. grid of points. Here r and r_E are the radius to the field point and the radius of the earth. A Taylor series expansion of the magnetic field for the ray path increment

$$Y_i = Y_{i0} + M_{ij} x_j \tag{13}$$

is obtained analogously from (12) and the β_{iL} parameter specification.

It is inefficient to use a fixed increment length, which instead should vary with the length scale of variation in the electron density. Hence, while small increments are required to account for rapid variations in the D and E layers, larger increments can be used for the F layer and above. The criterion used to estimate increment size beforehand is a fixed rate of convergence of the Taylor series expansion in (11). If the ratio of first order to zeroth order terms is to be a prescribed value ϵ, and similarly for second order terms, then it makes sense to estimate Δt as the largest value which satisfies the four equations

$$|A_i (dx_i/dt)_0| \Delta t < \epsilon X_0 \quad , \quad |B_3| (dx_3/dt)_0^2 \Delta t^2 < \epsilon^2 X_0 \tag{14}$$

where the summation convention is not assumed here. The two equations associated with the height variable x_3 almost always determine Δt. Lower limits on Δt and X_0 in (14) are available user options. There is little point in spending computer time on tiny raypath increments in the D and lower E layer regions, where the ionospheric model is inaccurate.

In the case of ray tracing between a transmitter and receiver, it is necessary to estimate the launch angles of the ray at the transmitter, trace to the receiver height, determine the errors in latitude and longitude, and then correct the launch angles for the next homing iteration. A kind of Newton-Raphson procedure is used to relate the errors in latitude (ΔL) and longitude (Δl) from one iteration to the corrections in launch elevation ($-\Delta e$) and azimuth ($-\Delta a$) angles for the next iteration:

$$\Delta L = (\partial L/\partial e)\Delta e + (\partial L/\partial a)\Delta a \quad , \quad \Delta l = (\partial l/\partial e)\Delta e + (\partial l/\partial a)\Delta a \tag{15}$$

The partial derivatives are evaluated at the landing point for the primary ray at the receiver height, which is consistent with the validity of (15) to first order. The launch angles corrections are obtained by simply inverting (15). The partial derivatives are calculated by a procedure which is similar to that used by Nickisch [1988] for focusing calculations. For an infinitesimal displacement of a launch angle, δe in elevation or δa in azimuth, a ray infinitesimally displaced from the primary ray is generated. The equations for the associated displacements are obtained by applying the differential or displacement operator δ to (6) and (7) and keeping only the lowest order terms in δk_i and δx_i. Because, for example, $\delta(dx_i/dt) = d(\delta x_i)/dt$, the original set of seven differential equations for x_i, k_i, and P is augmented in this way by an additional set of seven differential equations for δx_i, δk_i, and δP for each launch angle displacement. A further step is the division of these displacement variables by the associated launch angle displacement, so that the extra set of equations is actually for the partial derivatives of x_i, k_i, and P with respect to the launch angle. Another set of seven equations is associated with the other launch angle. The four solutions for the partial derivatives of x_1 and x_2 at the receiver height are inserted into (15). This procedure turns out to be more accurate and more efficient than one of sequentially tracing a primary ray and two slightly displaced rays to the receiver height, and then calculating numerical differences for the partial derivatives. The latter procedure was developed first and then used to check the development of the former procedure, which has been found to cut the overall computer CPU time approximately in half.

The evaluation of the right-hand side of each of the equations in (6)-(7), hereafter referred to as the rate function for a rate equation, and the rate functions for the fourteen additional partial derivatives of x_i, k_i, and P with respect to the elevation and azimuth launch angles, is straightforward. When propagation loss results and homing to a receiver are not needed, the extra fourteen rate equations are dropped.

A partial set of initial conditions for the first ray path increment is

$$x_i,\ P,\ \delta x_i,\ \delta P = 0 \qquad (i=1\text{-}3) \tag{16}$$

$$(k_1,k_2,k_3) = (\omega/c)\mu(-\cos(e)\cos(a),\ \cos(e)\sin(a),\ \sin(e)) \tag{17}$$

The symbol δ now represents, for convenience, the two partial derivative operators $\partial/\partial e$ and $\partial/\partial a$. The initial conditions for the six variables δk_i are found by applying $\partial/\partial e$ to (17) and solving the three coupled linear equations for $\partial k_i/\partial e$ which result, and then similarly by applying $\partial/\partial a$ to (17).

The basic idea is unchanged from RS; one solves for a ray path increment, transforms the results to local Cartesian axes at the beginning of the next ray path increment, and then repeats the process. There are several numerical techniques for solving rate equations [Burden et al, 1981]. Three options are considered here. One option is the standard fourth order Runge-Kutta (RK4) method with four rate function evaluations (per rate equation) per ray path increment, which contribute with weights 1/6, 1/3, 1/3, and 1/6. The rate function evaluations are complicated, but are simplified by the Taylor series expansions and the use of Cartesian coordinates. The RK4 technique does not provide for control on the local numerical truncation error. The second option is the fourth order Runge-Kutta-Fehlberg (RKF54) method, with error control, which uses an RK method of order five to estimate the local truncation error of an RK method of order four. This technique requires six rate function evaluations per increment. The third option is a fourth order Adams (ABM4) type of method, which uses the Adams-Bashforth four-step explicit method for the "predictor" and one iteration of the Adams-Moulton three-step implicit method for the "corrector", with the starting values obtained from the RKF54 method. This technique has error control and requires only two rate function evaluations per increment. The conventional application of this uses equal step sizes, but it has been generalized to the present case, which uses unequal step sizes (group path delay increments). Although the details are different, error control has been used in the manner suggested by Burden et al [1981] to adjust the step size. It is only required to recalculate one rate function for the new predictor solution, when necessary to keep the local truncation error within a specified tolerance. Adjustment or recalculation of other rate functions is not required, unlike the case for equal step sizes. The predictor and corrector are used to estimate the local truncation error, and a weighted average of them determines the ray path increment solution. Rate functions from previous increments are used in the method, and they are transformed to the local Cartesian coordinates of the present increment.

It remains to indicate, without getting into complicated details, the way in which the primary and displaced rays are led up to and beyond a boundary between height profile function segments, up to and beyond an earth reflection, and up to the receiver height. In general, the rate functions do not change smoothly at a boundary between height profile segments. The size of the last primary ray increment before the boundary is reduced so that it ends approximately at the boundary. The end point solutions are calculated, and then are corrected for the presence of boundary curvature and tilts. Following this the primary ray ends accurately on the boundary, but the displaced rays do not. A corresponding correction procedure is applied for each displaced ray, and this introduces differences δG for the group path delays of the displaced rays, over and above that of the primary ray. The next increment for the primary and displaced rays thus begins on the boundary and uses new rate functions. Solutions with the ABM4 method reapply the RKF54 starting procedure at the boundary, as the ray path solutions are extended beyond it. The $\underline{k}$ and $\delta\underline{k}$ vectors are required to have the same directions at the beginning of a new increment as at the end of the preceding increment, although they are renormalized in accordance with (8). Hence, for example, the artificial, discontinuous change in rate functions at a boundary, introduced by the ionospheric height profile model, is not allowed to scatter the ray. The primary and displaced ray increments are forced to end at the receiver height in a similar manner. This is the end of the ray path if it corresponds to a prescribed number of "bounces". The bounce point number is an input parameter which is used to specify a particular mode of propagation (one-hop, two-hop, etc.). A bounce point is counted, and the solutions for it are stored in a file, whenever the ray either reflects from the earth or intersects the receiver height. Similar corrections are applied to the primary and two displaced rays at the earth's boundary, where specular reflection is assumed to take place. The new k_{in} and δk_{in} components for the start of the reflected increment are related to the corresponding numerical solutions k_i and δk_i of the preceding increment by

$$k_{1n} = k_1 \quad , \quad k_{2n} = k_2 \quad , \quad k_{3n} = -k_3 \tag{18}$$

$$\delta k_{1n} = \delta k_1 - 2k_3\delta x/r_E \ , \ \delta k_{2n} = \delta k_2 - 2k_3\delta y/r_E \ , \ \delta k_{3n} = -\delta k_3 - 2k_1\delta x/r_E - 2k_2\delta y/r_E \tag{19}$$

A bounce point solution specifies x_i, k_i, P, and group path delay G. It also specifies δx_i, δk_i, δP, and δG at the bounce point height when either the homing or the propagation loss option is selected. This information is sufficient for a variety of simulations, including those for ionosonde, time-of-arrival, and frequency-of-arrival systems. The program is

flexible; it can handle transmitter and receiver locations below, in, or above the model ionosphere.

The propagation loss calculation has also been upgraded. Loss is defined as

$$L = 10\log_{10}(P_t/P_r) \ , \tag{20}$$

where P_t is the power radiated by the transmitter, and P_r is the amount of this power received by a receiver. The program includes contributions from non-deviative absorption and earth-reflection loss, according to the ITS-78 model [Barghausen et al, 1969]. The non-deviative absorption contribution of the earlier program has been improved to include magnetic effects and better location of the D layer intersection. Another contribution included is the basic transmission loss L_b, which is the basic free space transmission loss of [Barghausen et al, 1969], generalized to include ionospheric focusing effects. It is found from

$$L_b = 32.448 + 20\log_{10}f + 10\log_{10}(A_r r_t^2/A_t), \tag{21}$$

where f is the wave frequency in MHz, A_r is the cross-sectional area of the flux tube associated with the primary ray and the two infinitesimally displaced rays at the receiver, and A_t is the cross-sectional area of this flux tube at a distance r_t km from the transmitter, which is small enough to assume that the three rays from the transmitter to this point are straight. It can be shown that

$$A_r r_t^2/A_t = |R/T| \quad , \quad R = \{\hat{s} \cdot [(\partial \underline{x}/\partial e) \times (\partial \underline{x}/\partial a)]\}_r \tag{22}$$

$$T = \{\hat{s} \cdot [d(\partial \underline{x}/\partial e)/dt \times d(\partial \underline{x}/\partial a)/dt]\}_0/(ds/dt)_0^2 \tag{23}$$

$$ds/dv = [(dx_i/dt)(dx_i/dt)]^{1/2} \ , \quad \hat{s} = (d\underline{x}/dt)/(ds/dt) \tag{24}$$

The expressions in parentheses or brackets with a 0 subscript are evaluated at the transmitter launch point from the initial conditions. Those with the subscript r are evaluated at the receiver from the ray path end point solutions. The contribution to L from deviative absorption is not included in the program. Further, the gains (in dB) of the transmitter and receiver antennas, in the launch and receiving directions, relative to an isotrope, should be subtracted in the calculation of L.

This completes the discussion of upgrades, one of which is the inclusion of magnetic effects. Similar upgrades have been incorporated into the program without magnetic effects, which uses the analytic solution procedure for ray path increments.

3. HF NVIS EFFECTS

As a test of the ray trace program, it has been adapted to a simulation of oblique ionograms. Ionograms have been measured at NRL (38.817N, 282.983E) with the chirp sounder transmitter at Driver, Virginia (36.817N, 283 5E), for a sounder range of 227 km. The measured ionograms for three successive morning hours for October 3, 1989, are shown in Figure 1. The simulation is for the middle ionogram, not only because it looks interesting, but also because there are large E-W tilts associated with the sunrise transition at this time. Values for the sunspot number and planetary magnetic index values, SSNE=143 and K_p=2.7, are used. The sunspot number is provided by Air Weather Service (AWS) for use with the ICED ionospheric model [Tascione et al, 1988]. It is a five-day running average which makes ICED fit foF2 data from a global network of vertical-incidence sounders.

The ionogram simulator steps through a prescribed set of frequencies. For each frequency the program uses the great circle azimuth angle and starts with the elevation angle e=15°. The ray trace program evaluates the range. If the difference between the calculated range and sounder range is more than 100 km, or between 50 and 100 km, or less than 50 km, the increment to the next elevation angle is 5, 2, or 1 degree, respectively. The search goes on until the range difference changes sign, at which point there is a linear interpolation, followed by homing iterations to the receiver at NRL. When homing is achieved, results from the solution for x_i, k_i, P, G, δx_i, δk_i, δP, δG, and L at the receiver are stored in a file. Then the search in elevation continues similarly from the point where the range difference changed sign until there is another sign change, at which point homing is repeated. The search goes up to e=89°. The search is interrupted if and when the ray escapes through the F2 layer, and the scan of frequencies stops when there are no oblique ionogram solutions.

The simulated oblique ionograms for 1400 (UT), one each for the RADARC and ICED ionospheric models, are shown in Figure 2. The RADARC ionogram lacks an F1 layer effect and looks more like the 1300 (UT) ionogram in Figure 1. The ICED ionogram exhibits more of an F1 layer effect than is seen in Figure 1, where it shows up as a ledge at 1400, which tends to be subsumed by the F2 layer at 1500. The simulated ionograms in Figure 2 show F2 layer features, including the O- and X- mode magnetic field splittings, which essentially agree with experiment. This tends to confirm the ray tracing procedures for inclusion of magnetic effects. It is interesting that

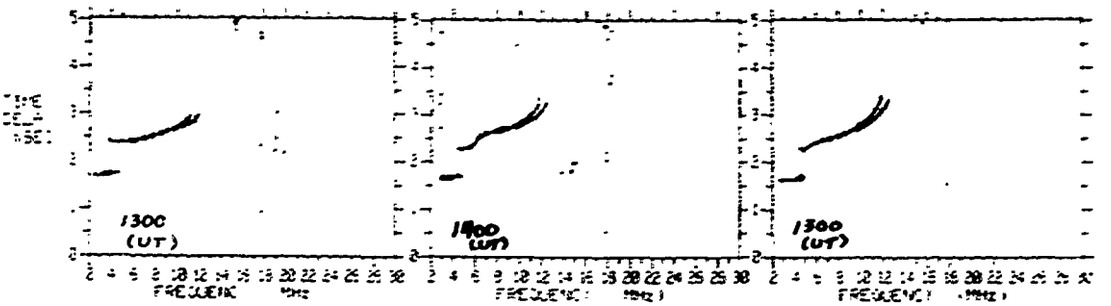


Fig. 1. Experimental oblique ionograms on the Driver, VA to MHL path - range 227 km.

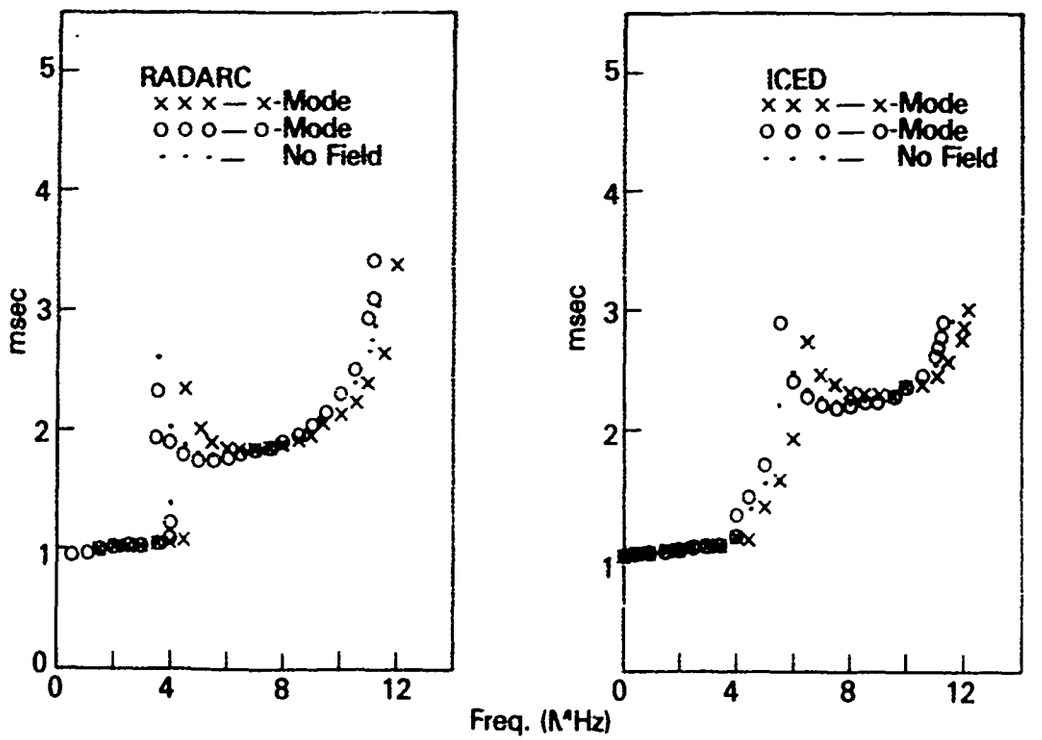


Fig. 2. Ray trace simulations of the 1400 (UT) ionogram in Fig. 1 for the RADARC and ICED ionospheric models.

the ionograms in Figure 1 show no effect from X-modes in the F layer trace near the E-F transition. Calculated RADARC propagation loss is higher for the X-mode than for the O-mode, by about 9 dB at 5 MHz and 6 dB at 6 MHz. It turns out that when the experimental threshold value is lowered, the X-mode solution in the 1300(UT) ionogram in Figure 1 shows up near the E-F transition, similarly to the RADARC calculations of Figure 2.

Also shown by dots in Figure 2 are results for the no-field case. There is an option in the ray tracing program MRAYTRACE for choosing the number of L values in (12) to be kept in the calculations for the magnetic field. The associated parameter is called NOLS. The no-field case is NOLS = 0, and the O- and X-mode (ordinary and extraordinary) results in Figure 2 are for NOLS = 8. The no-field results are closer to the O-mode results than the X-mode results. Such a simulation can be used as a guide for finding the true height profile directly from an oblique ionogram. A recent method for this [Reilly and Kolesar, 1989] assumes spherical symmetry of the ionosphere and no magnetic field.

It is of interest to compare the performances of the various numerical solution techniques with each other and with the analytic solution technique for the no-field case. In Figure 3 the ICED group path delay in Figure 2 for 10 MHz is plotted for the various techniques as a function of the Taylor series convergence ratio (TSCR) ϵ in (14), which determines the size of a raypath increment. The exact analytic solution technique is labelled as the RAYTRACE curve in Figure 3. Convergence toward the $\epsilon = 0$ result is clearly indicated. Unlike the other curves, this particular curve has had the provision of a lower limit on the raypath increment removed.

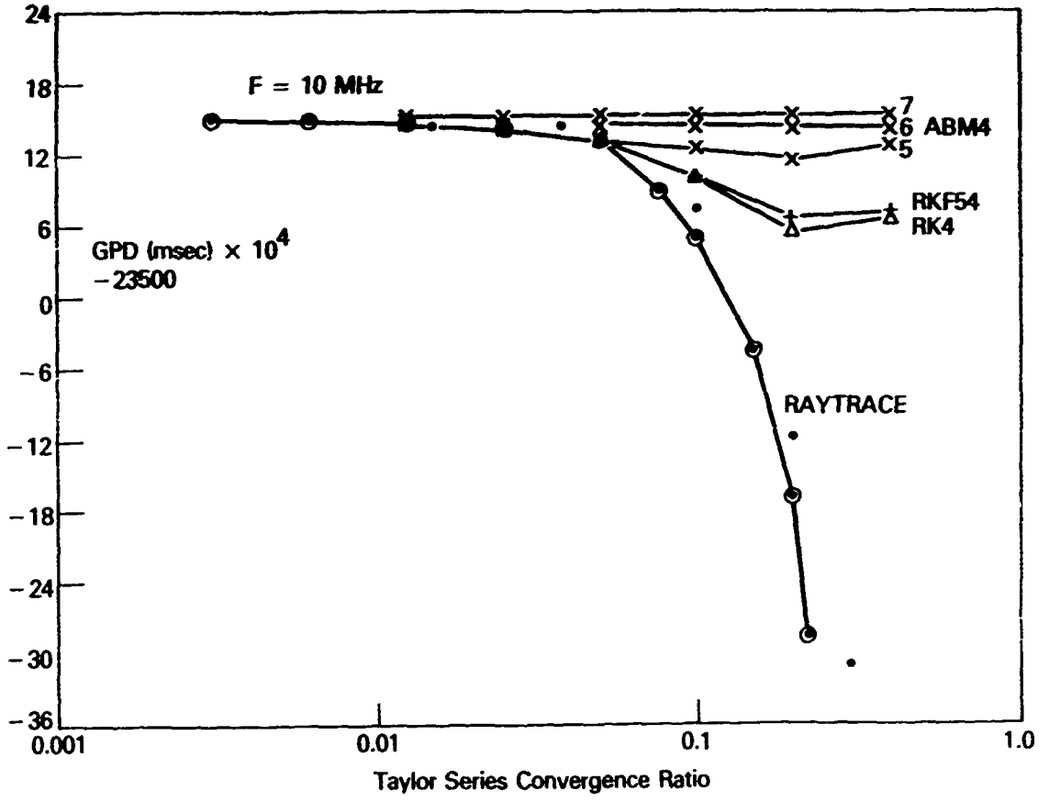


Fig. 3. Dependence of the ICED no-field 10 MHz group path delay result on Taylor series convergence ratio for various solution techniques.

The results associated with the imposition of the lower limit are indicated by the nearby isolated points. The points tend to coalesce with the curve for smaller ϵ values. The curves for RKF54 and RK4 are close to each other, although the RKF54 method provides step size shortening if the local numerical truncation error is not within tolerance. This tolerance was set at the level of 7 significant figures (NSF = 7) for ray path increment solution accuracy, and at this level, had very little effect. The RK techniques do not closely track the analytic soution technique until ϵ diminishes to about 0.05. The ABM4 curves in Figure 3, with numerical error control set at the levels of NSF = 5, 6, and 7, insert a significant amount of step size shortening. This occurs for NSF = 5 when $\epsilon > 0.05$. It is especially true when NSF = 6 and 7. This step size shortening is more pronounced in regions where strong ray refraction occurs. As a consequence results from the ABM4 method are not very sensitive to the choice of ϵ. The results in Figure 2 are very accurate, since they are associated with the ABM4 technique, NSF = 7, and $\epsilon = 0.015$. Suppose it is desired to obtain GPD in Figure 3 to five significant figures. Then one can choose RAYTRACE, RKF54, RK4, or ABM4 with NSF = 5 and $\epsilon = 0.027$; or one can choose ABM4 with NSF = 6 and $\epsilon = 0.05$ or 0.10; or one can choose ABM4 with NSF = 7 and $\epsilon = 0.10$. The CPU time on a VAX 6320 is indicated in Table 1. Evidently the ABM4 method is a good choice. It is even much faster than the analytic solution technique for the given level of accuracy. For all the methods, there were four homing iterations, with convergence obtained to within a millimeter. Considering that a primary and two displaced rays are associated with each iteration, then all the no-field results are accurately obtained in a time of about 0.4 seconds

per ray. The time for inclusion of magnetic effects (NOLS=8) for the O-mode, as indicated by the last three entries in Table 1, is a little more than twice this.

TABLE 1. CPU TIME OF VARIOUS METHODS (VAX 6320)

METHOD	NSF/NOLS	TSCR	TIME (Sec.)
RAYTRACE	NA	0.027	11.7
RKF54	7/0	0.027	15.9
RK4	NA/0	0.027	12.1
ABM4	5/0	0.027	12.5
ABM4	6/0	0.050	8.5
ABM4	6/0	0.100	5.3
ABM4	7/0	0.100	8.0
ABM4	6/8	0.050	18.3
ABM4	6/8	0.100	11.9
ABM4	7/8	0.100	23.7

The requirement for accuracy is greatly diminished by inherent error in ionospheric specification by a climatological ionospheric model. Typical errors in density specification are in excess of 10%. A change in $foF2^2$ at midpath by 10% is found to affect the converged value of GPD in the third significant figure, i.e., by several hundred units on the ordinate scale in Figure 3. Hence, it would be permissible to choose the ABM4 method, ϵ=0.2, and NSF=4. The no-field result is 4.6 on the ordinate scale in Figure 3, correct to four significant figures, and it is obtained in a CPU time of 2.3 seconds. Therefore, this is an appropriate choice of solution method parameters.

Finally, it is of interest to consider single-site geolocation of the Driver transmitter from angles of arrival at NRL, or the other way around if the roles were switched. The results for ICED and the 10 MHz ray are again considered. Angles of arrival at the receiver for the various modes are indicated in Table 2. Geolocation results depend on the ionospheric approximation. It is often assumed in practice that the ionosphere is spherically symmetric, characterized by vertical-incidence sounder results at the receiver, and the magnetic field effects are usually neglected. This is one of various ionospheric approximations to be investigated in a ray tracing simulation. In the case of geolocation of NRL from Driver, the simulation includes: either the O-mode or the X-mode angles of arrival, either the full (F) ionospheric specification with tilts or the spherically symmetric ionospheric specification at Driver (D), and either the inclusion of magnetic effects (NOLS = 8) or the lack thereof (NOLS = 0). Hence, the ionospheric approximation is specified by three indices. For example, XF0 refers to the X-mode angles, the full ionospheric specification, and neglect of the magnetic field.

Results of the NRL geolocation simulation are shown in the top portion of Figure 4. The position of the NRL site is indicated by the open circle, which is associated with the full ionospheric treatment (0F8 and XF8). Neglect of ionospheric tilts accounts for something like a 0.1° miss, or

TABLE 2. ANGLES AT THE RCVR IN THE GEOLOCATION SIMULATION

XMTR	RCVR	MODE	El.(deg.)	Az.(deg.)
NRL	Driver	O	71.6917	349.183
		X	69.5286	349.009
Driver	NRL	O	70.3429	167.834
		X	68.5034	167.907

about 10 km. miss distance, and neglect of magnetic effects accounts for a similar error. Evidently the effects can be additive when both are neglected. The errors are slightly smaller for the corresponding Driver geolocation from NRL, but the same sort of effects occur. The displacements of the reflection points from the midpath (MP) point for the 0F8 and XF8 rays are also shown in Figure 4. It appears from Figure 2 that these errors should be more pronounced for 11 MHz. A larger effect would also be expected if the receiver to transmitter path were precisely in the direction of the ionospheric tilts [Reilly and Strobel, 1988]. An improved geolocation technique would use sounder measurements at one or more sites to update an ionospheric model like ICED, identify the mode (O or X) through polarization measurements, along with the angles of arrival, and use the ray trace program for proper inclusion of magnetic effects and ionospheric tilts.

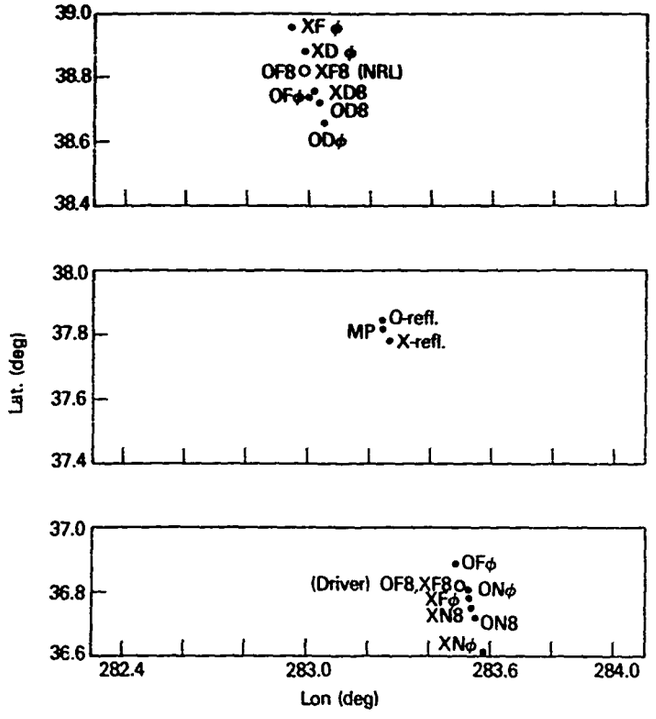


Fig. 4. Single site geolocation results for the ICED 10 MHz ray in various ionospheric and magnetic field approximations. Ionospheric reflection points near midpath (MP) for the OF8 and XF8 rays are shown.

4. DISCUSSION AND CONCLUSIONS

Several improvements have been made to the ray tracing program since the earlier paper. The inclusion of magnetic effects, detailed here, has necessitated a substantial revision. The earlier program has been useful in checking the validity of other upgrades. These include the Taylor series expansion for determining the size of the ray path increments, improved homing and focusing calculations, associated treatments for crossing boundaries between model profile segments and for earth reflections, and other efficiencies. The program retains the use of Taylor series expansions, including one for the magnetic field, and the use of local Cartesian coordinates and transformations in passing solutions from one ray path increment to the next. A numerical multistep Adams-Moulton solution technique with numerical error control has been adapted to this approach. The treatment of equal step sizes in this method has been generalized for unequal step sizes, and the technique turns out to be well suited to this ray trace program, even superior to the analytic solution method for the case of no magnetic field. The result is an accurate and efficient program, within the constraints of a climatological ionospheric model description, which seems capable of handling any radio system application for frequencies in the HF band and above. Further upgrades could include a tropospheric refraction model, graphical outputs, a global collision frequency model, ionospheric model improvements, a model for statistical assessment of TID and irregularity effects, propagation loss improvements, and adaptations to particular radio systems, among others.

The model has been tested against and found to be consistent with oblique ionogram data for HF NVIS propagation, which exhibit strong magnetic effects. The sunrise transition data are also associated with strong ionospheric tilts, and the effects of these and the magnetic effects were found to contribute importantly in HF NVIS single site geolocation of a transmitter.

ACKNOWLEDGMENTS

The cooperation of M. Daehler and L. Quinn in supplying experimental ionograms is appreciated, as are the assistance of P. Kehres in the preparation of the manuscript and comments from J. Goodman and F. Kelly.

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, Tech. Rep. ERL 110-ITS78, Environ. Sci. Serv. Admin., U.S. Dept. of Commerce, Boulder, CO, 1969.

Burden, R.L., J.D. Faires, and A.C. Reynolds, Numerical Analysis, Prindle, Weber, and Schmidt, Boston, MA, 1981.

Cummack, C.H., An ionospheric ray-tracing technique, Radio Sci., 15, 1121-1125, 1980.

Lighthill, M.J., Group velocity, J.Inst.of Math and its Appl., 1, 1, 1965.

Nickisch, L.J., Focusing in the stationary phase approximation, Radio Sci., 23 (2), 171-182, 1988.

Reilly, M.H. and E.L. Strobel, Efficient ray tracing through a realistic ionosphere, Radio Sci., 23 (3), 247-256, 1988.

Reilly, M.H. and J.D. Kolesar, A method for real height analysis of oblique ionograms, Radio Sci., 24 (4), 575-583, 1989.

Tascione, T.F., H.W. Kroehl, R. Creiger, J.W. Freeman, Jr., R.A. Wolf, R.W. Spiro, R.V. Hilmer, J.W. Shade, and B.A. Hausman, New ionospheric and magnetospheric specification models, Radio Sci., 23, 211-222, 1988.

Thomason, J., G. Skaggs, and J. Lloyd, A global ionospheric model, Rep. 8321 (AD-000-323), Nav. Res. Lab., Washington, DC, August 1979.

AD-P006 297

ANGLE OF ARRIVAL OBSERVATIONS OF HF RADIO WAVES[1]

Zebulon R Jeffrey
Andrew Australia, Technology Park, Australia 5095

Leo F McNamara
Andrew Government Systems,
2916 National Drive, Garland, TX 75041

ABSTRACT

This paper describes selected total angle of arrival observations of signals from HF transmitters at ranges from 41 to 1150 km, over a wide range of ionospheric conditions. Particular attention is paid to the physical interpretation of the observations, especially in terms of the effects of travelling ionospheric disturbances (TIDs) and multi-moded propagation. Observations made during a mid-latitude particle-E event are also briefly described.

1 INTRODUCTION

The observations described here are a small subset of observations made at two sites near Dallas, Texas, namely Rowlett (32.9N, 263.5E) and Floyd (33.2N,263.7E), and in Adelaide, Australia. The Rowlett observations were mainly of medium-range (400 - 800km) and long-range transmitters (WWV Fort Collins, at a range of 1150km), and were made during a very disturbed week in March 1989 (March 10 - 18). This week was punctuated by shortwave fadeouts, negative ionospheric storms reaching the "G" condition, and even a major particle E event (SESC, 1989; IPS, 1989). The Floyd observations, on the other hand, were made during an ionospherically quiet week from 9 to 14 October, 1989, just before the onset of an exceedingly eventful week which saw several major flares and their accompanying terrestrial effects. The Adelaide observations were made during December 1988, an uneventful week (IPS, 1988).

All of the angle of arrival (AOA) observations were made with either the Andrew AOA system, SKYLOC, or the Andrew ionosonde, TILTSONDE. SKYLOC is a 7-element interferometric angle of arrival system, which uses phase and amplitude measurements at each of the antennas to deduce the total AOA of the incoming wavefront. The procedures used in SKYLOC, and the accuracy of the angle measurements, are described by Jeffrey et al (1988), while McNamara (1988a) has described ionospheric aspects of a set of AOA observations made on long-range transmitters in Australia. As deployed in the field in Texas, SKYLOC made a measurement of the AOA and the quality of that measurement at the rate of 2 per second. The quality of each AOA measurement is characterized by its WSP (Winkler Spectral Processing) index, which takes into account several factors which include the linearity of the wavefront, and is small for a linear wavefront. Those measurements failing specified quality criteria are rejected, since the assumptions inherent in the data processing would not then be valid. In general, all data not corresponding closely to single-moded propagation are rejected.

AOA observations could be made on vertical monopoles (for the low elevation, long range transmitters) or on circularly polarized loops (for high angle, short range transmitters).

[1]Prepared for the Ionospheric Effects Symposium, Washington DC, May 1990.

91-09705

TILTSONDE is a derivative of the Andrew ionosonde SKYSONDE (used at Rowlett), extended to include the capability of measuring the AOA of its own returning signals. It takes in many of the sophisticated data processing procedures used in SKYLOC, but is deployed with only a 3-element antenna array and consequently no wavefront testing. AOA measurements are made on specified spot frequencies, at the rate of 2 seconds per "tilt" measurement. TILTSONDE can select either O or X mode polarization, and generates the corresponding ionogram in a background mode. As deployed at Floyd, TILTSONDE revisited the same frequency every 20 seconds.

The TILTSONDE observations, and the SKYLOC observations for short-range transmitters, are discussed in terms of the deviations of the downcoming ray from the zenith, towards the north and towards the east. In short-range situations, the high elevation angles mean that the azimuth of the ray has very little significance, since a very small actual movement in the reflection point can result in a large change in azimuth. The north (N) and east (E) zenithal angles are related to the azimuth (A) and elevation (E) by N = atan(cos A / tan E) and E = atan (sin A / tan E).

2 OBSERVATIONS OF WWV (FORT COLLINS)

WWV was the default transmitter during the March observing period, to be observed when other transmitters (all non-cooperative) could not be heard. However, because of the disturbed nature of the period, WWV was observed more often than had been scheduled, since its high power meant that it could still be heard during absorption events. In fact, many of the other intended transmitters simply went off the air when their communications failed. WWV was observed on 2.5, 5.0, 10.0 and 15.0 MHz at different times of day. During some periods, two frequencies were monitored simultaneously.

The data presented here are chosen to illustrate several points:-

1. The minimum spread in azimuth and elevation
2. The effects of multi-moded propagation
3. The effects of TIDs on even this long circuit
4. The effects of particle E

2.1 Minimum Spread in AOA

Figure 1 shows a numerical histogram display of elevation angle vs azimuth for Rowlett cases 166 and 167 combined, which were both for 15.0 MHz at around 1800 local time. The resolution of the plot is 1 degree, and the bin counts range from 1 to 15, with 11 to 15 being represented by A, B etc. Counts greater than 15 are denoted by an asterisk. Note that in this particular case asterisks hide the fact that over 300 counts occur in some bins The first column of the plot is the elevation angle, while the second is the total count for that elevation angle. There is no line printed for any elevation angle with zero counts.

The data included were for a WSP of 9 or less, which included 75% of all the data. The average elevation angle in this case is 25.3 degrees, with 88% of elevations lying in the 24, 25 and 26 degree bins. Similarly, about 95% of azimuths lie in the 321 and 322 degree bins. The bulk of the observations fell in the 322 degree bin for minutes 1 to 7, and minutes 49 to 60, and in the 321 degree bin between minutes 8 to 48. The average azimuth of 321.4 degrees is within 0.1 degrees of the true value. As an aside, even this tight distribution was excelled for case 48, for which 99.5% of azimuth observations fell into the one bin.

Propagation simulation using both the Classical method (reflection from a plane mirror at a height h'(fv)) and analytic raytracing through a multi-segmented quasi-parabolic N(h) profile (McNamara, 1988b) showed that the propagation mode was a 1F2 mode. Since such modes often cause systematic deviations of 1 degree or more in azimuth, the very close agreement of the observed and correct azimuths must be regarded as fortuitous. In general, SKYLOC can determine the azimuth of an incoming wave with an error much less than any systematic shift imposed by the ionosphere.

2.2 Multi-moded Propagation

Figure 2 shows the elevation vs azimuth plot for Rowlett Case 161 (10 MHz, 1300LT). The main interest in this case is the very large spread in elevation angle, accompanied by only a small spread in azimuth. The observing period was 53 minutes. This spread appears to be due to the presence on the circuit of 4 propagation modes, 1E (5 degrees), 2E (15 degrees), 1F2 (28 degrees) and 2F2 (44 degrees). All 4 modes are on basically the same azimuth, which makes it a little harder to substantiate the identifications of the modes present. It is quite common for the azimuths for E and F modes to differ by more than one degree. The interpretation is, however, consistent with the results of propagation simulation.

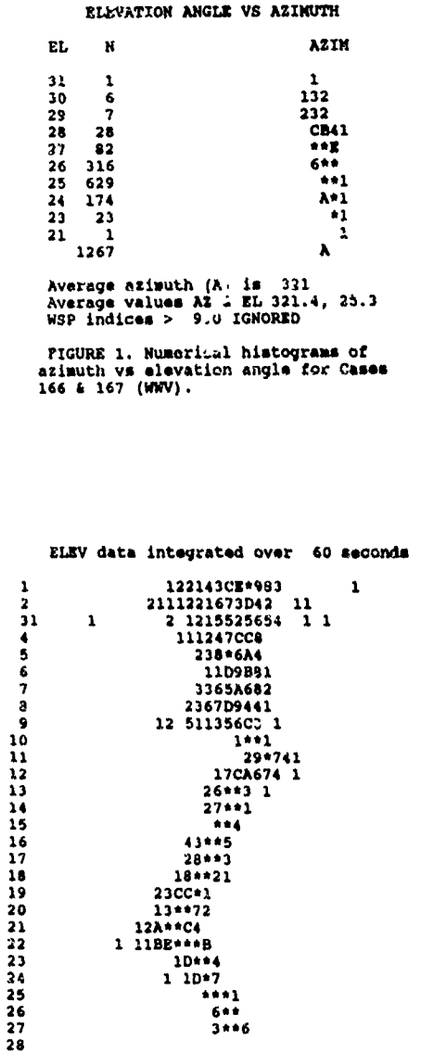


FIGURE 1. Numerical histograms of azimuth vs elevation angle for Cases 166 & 167 (WWV).

FIGURE 3. Numerical histograms of the elevation angle, as a function of time. (Case 110, WWV).

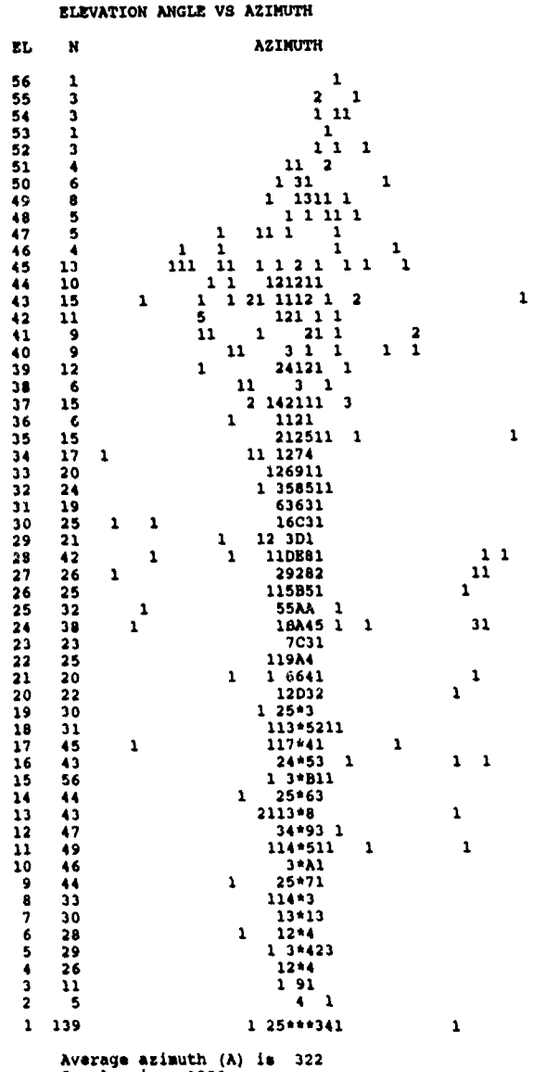


FIGURE 2. Numerical histograms of azimuth vs elevation angle for Case 161 (WWV).

The figure shows a change in data quality at 38 degrees, as the mode changes from 1F2 to 2F2, the latter mode having a much larger spread in the observed azimuths. Azimuth vs time plots show that the azimuth (as does the WSP also) becomes more spread for both periods for which the 2F2 mode was the dominant mode (minutes 1 to 20, 28 to 36). The high count in the 1 degree bin (where mathematically impossible solutions are placed) is commensurate with having so many modes present, with a much higher chance of multi-mode interference. This case is typical of those in which it is difficult to obtain quasi uni-modal propagation (QUMP) because of the requirement that (N-1) of the N possible propagation modes must effectively fade away.

2.3 Effects Of TIDs

Figure 3 shows a numerical histogram plot of elevation vs time, with the data being grouped into one minute intervals, for Rowlett case 110 (10.0 MHz, 2030LT).

The figure shows the presence of a strong 20-minute TID in the elevation angle. The azimuth remains essentially unchanged at 321 degrees, but the elevation swings between 13 and 23 degrees. The fact that there is no obvious effect on the measured azimuths indicates that the TID was travelling almost directly along the circuit.

The WSP indices also exhibit a periodic behaviour for the last 15 minutes of the observing period, and show a strong correlation with elevation angle. Both the WSP and elevation data show a sudden decrease in spread at minute 10.

2.4 Effects Of Particle E

Rowlett cases 53 to 59 were all observed during a large particle E event, which lasted from about 1730 LT to 2130 LT (UT Days 13/14 March). Observations were made on different WWV frequencies, sometimes on two frequencies simultaneously (since most HF communicators appeared to have given up for the night). All were characterized by more angular spread than normal for the high-power WWV transmissions, and showed the presence of both Es and normal propagation modes affected by the passage of the waves through the particle E layer. Figures 4 to 6 show the data for case 54 (5.0 MHz, 2100 LT). Figures 4 (WSP), 5 (azimuth) and 6 (elevation) show that the data undergoes a qualitative change after the first 10 minutes or so. The early data show very scattered values of the three parameters. The elevation data for the first ten minutes yield an average of about 18 degrees, but this value is rather uncertain. The average azimuth is close to the "correct" value, but this average is again uncertain.

Minutes 11 to 16 show an azimuth increase of about 4 degrees, and an elevation decrease of about 12 degrees, indicating the passage of a large tilt through the reflection point. The azimuth then remains fairly steady, while the elevation swings back up towards the higher values, after passing through a minimum of around 9 degrees. Propagation simulation indicates that any normal propagation below about 17 degrees would be via the E layer, while propagation above that angle would be via the F2 layer. The interpretation of the observations in terms of propagation modes hinges on the interpretation of the time series of elevation angles. Figure 4 (all WSP values) suggests the presence of a large TID between minutes 11 and 28. However the low-WSP data suggest that there were two propagation modes present during this time, with elevation peaks of around 10 degrees between minutes 16 and 26, and around 20 to 25 degrees between minutes 11 and 15, and during minute 27.

If we accept this latter interpretation, we would have a low-angle Es (particle E) mode between minutes 16 and 26, and a higher-angle 1F2 mode at other times. The mode for the first 10 or so minutes would be an F2 mode, with a large amount of scatter and attenuation imposed by the passage of the rays through a lower non-reflecting particle E layer, either on the way up or the way down (or both). The F2 mode continues, but with less scatter, until minute 16, suffering a large azimuth swing not untypical of the F2 layer.

```
WSP indices > 99.0 IGNORED
Sample size   898
WSP data integrated over  60 seconds

 1      11   2    1  12  212122 222   21 1  2112 1 11   2                1
 2           1          111   11          1   1  2   1111 1 1  1 1         1
 3         1  1    1    1 2 1  1       1  1    1     2      1   1      1   11
 4          1   11    1  1          1   3    21 1   11
 5       1 1     11  1   1    1   1          1   1  21   2      1   1    1   1
 6          1   1 1   1  1  11    21 1 21   1   1 1 1     1              11
 7          1 1 1     1   1   1       11   1 1  11 1           1
 8     22  123331 11 1 23   1 2        1           11       1           1
 9    1 1 21   1131 112      11 11 1   1              1           1  1
10                  2 11 2    1   1  112            1       1
11    2 12141732341112  1   1    1              1
12     1  365373321  3 1    1
13   1116442111351 3   1   11          1   1
14     2135513421  1 2  2 2 1         1  1       1
15      1  1 31 13 11 1  1131 1 1  1         11
16   2  1 32112313 4111  11          1    1 1
17   11234736243 21 11     1
18  2113 33224222231        1                1
19   1244456511  12111             1           1
20    24426822111   11
21   126285844412
22  1136532 5361311   1
23  1 2125573544211    1
24    22244497622      1
25     221252113 11 1111 11 1  11        1             1
26   215464 392 311 1 1   1
27     4E8135 12 1 1 2 1                  1
28
```

FIGURE 4. Numerical histograms of the WSP index, accumulated into one minute groups. Case 54, WWV.

```
WSP indices > 99.0 IGNORED
Sample size   898

AZIM data integrated over  60 seconds

 1          1 1  12 3229434511      11              1
 2   1       11  1221  1 1  2      2          1 1 1  2    2  11
 3                13222 1    1  2  1               1    1 1
 4                11212  32        1  11     1
 5       1      2 1 113 11 131 2   1  1 1
 6               2211152 121   21                          1
 7            1  1 2   3311 1 1
 8               1112134464113   11
 9    1             1 31331431 1 11
10 1  1           2     111111  1 21                 1
11     1      1  1 5AA64 2
12               1 2989621  1
13                13545A44  1    1
14               36179A2
15              111 12313222131 1
16                     4476711  1
17                    8*74
18                   15*A1  1
19                 1  6*6
20                   44*3
21                   16*92
22                    6*A
23                    7**
24                    9*6
25                    4C771 1
26                   1**3
27                  11C*73
28
```

FIGURE 5. Numerical histograms of azimuth angle, accumulated into one minute groups. Case 54, WWV.

3 OBSERVATIONS OF KLC

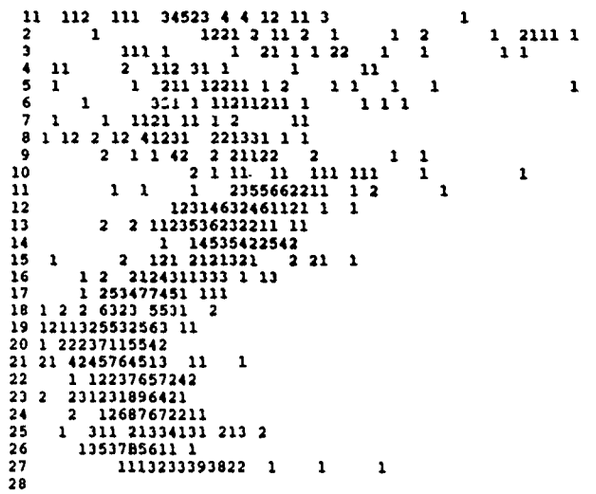
Radio station KLC is at Arcadia, Texas, 412 km from Rowlett. The data presented illustrate :-

1. Large swings in AOA
2. The effects of Particle E
3. Multi-moded propagation
4. Frequency dependence of angular spread

3.1 Large Swings In AOA

Figures 7 and 8 show the time variation of the azimuth and elevation for case 18 (8.666 MHz, 2000 LT). This case shows a dramatic swing in azimuth of about 15 degrees within the first 5 minutes of the observing period (Figure 7), with the elevation dropping about 5 degrees at the same time. In Figure 9, the high elevation - low azimuth clump of data corresponds to the beginning of the observing period. Figure 10 shows a dearth of low WSP indices for the first few minutes, indicating a weak signal and/or a distorted wave front at this time. It is the later data which correspond more closely to the correct azimuth.

Propagation simulation shows that the propagation could not have been by a normal 1E (penetration) or 1F2 mode (KLC is well inside the skip distance). The propagation mode was probably a 1Es mode, more or less in keeping with the occurrence of Es at Rowlett, which disappeared during the observing period. (The Es, not Rowlett). The off-azimuth propagation must therefore have been due to reflection from the leading edge of an Es cloud, which moved across the circuit. It cannot, of course, be attributed to a TID.



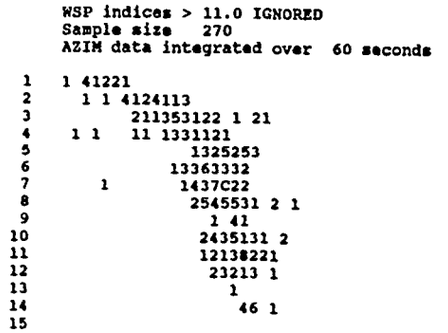


FIGURE 7. Numerical histograms of the azimuth, with the data accumulated into one minute groups. Case 18, KLC.

FIGURE 6. Numerical histograms of elevation angle, accumulated into one minute groups. Case 54, WWV.

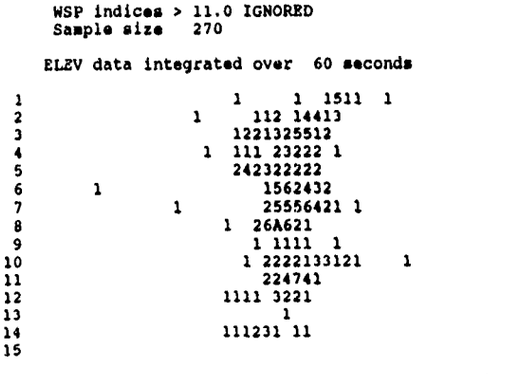


FIGURE 8. Numerical histograms of the elevation angle, with the data accumulated into one minute groups. Case 18, KLC.

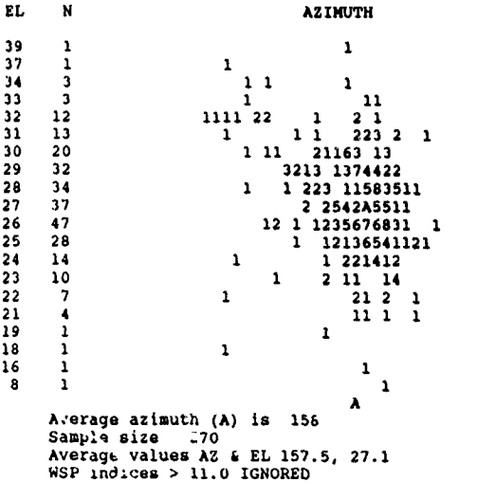


FIGURE 9. Numerical histograms of the azimuth vs elevation angle. Case 18, KLC.

3.2 Effects Of Particle E

Case 52 was observed just after the start of the particle E event, at around 1700 LT. As shown by Figure 11, the AOA data are exceedingly spread. Two-thirds of the WSP indices in this case exceeded 28. In contrast, most of the WWV observations under normal conditions had WSPs less than 10 or so. Extensive analysis has failed to reveal any structure within the AOA distributions, indicating that the particle-E layer was very irregular. This is not unexpected when the irregular nature of particle precipitation is considered.

It is interesting to note that observations during the particle E event of the somewhat more distant radio station WLO (Mobil, Alabama) at a range of 834 km were far less spread than the KLC observations. This is probably attributable to the fact that the KLC signals at a higher frequency and elevation angle had to penetrate further into the layer before being reflected.

3.3 Multi-moded Propagation

Figures 12 and 13 show the AOA observations for case 87 (8.666 MHz, 1020 LT). Figure 12 shows that there were two propagation modes in this case, with elevation peaks around 56 degrees and 33 degrees, and on essentially the same azimuth. The higher angle mode is the normal 1F2 mode. The high equivalent vertical frequency for the lower mode (4.9 MHz) shows that this mode cannot be a normal 1E mode, while there was no F1 layer to support a second mode at this elevation angle. It is therefore deduced that this mode was an Es mode. HF absorption was particularly high at the time, but there is some indication of Es on the local ionogram.

Figure 13 shows that the Es mode lasted for about 6 or 7 minutes at the start of the observing period. The Es propagation was associated with relatively high WSP values, with very few low WSP values being obtained for the first 9 minutes of the observing period. The high WSP values and very spread elevations for the Es mode indicate that the Es was a very patchy layer in this case.

The azimuth plot with time shows a definite 20-minute TID for the last part of the observing period, in contrast with the relatively constant elevation angle, suggesting that the TID was travelling more or less east-west across the circuit.

3.4 Frequency Dependence Of Angular Spread

Figures 14 and 15 show the azimuth-elevation plots for cases 90 and 91, which were recorded at around 1230 LT on frequencies of 8.666 MHz (Figure 14) and 13.038 MHz (Figure 15). Apart from showing that the elevation angle is significantly higher for the higher frequency case, the plots also show that the spread is much less for the higher frequency. This feature can be attributed to a much stronger dependence of ground range on elevation angle for the higher frequencies.

This effect is true in general, and was seen often in the WWV data. However its effects tends to be swamped by the presence at the lower frequencies of more than one propagation modes.

4 OBSERVATIONS WITH TILTSONDE

In its current implementation, TILTSONDE can make AOA ("tilt") measurements on up to 5 selected frequencies, while generating a normal ionogram as a background task. The data presented here are chosen to illustrate the variations of the tilts with time, and with operating frequency. Both of these variations are manifestations of the passage of one or more TIDs through the F2 layer during the observing period.

```
WSP data integrated over  60 seconds

 1     13223 24111225  32 11 1 1  1  1   1 1   2
 2    131152562245l1 4 3  1      1 11  1111  1
 3   213443523353311111112 2     2    1    1  1
 4   222 12345321321212   1 11     1      1
 5   243234124A3221 222123    11       1    1         1
 6   312414362252232l111    112  1    1   1 2 1     1        1
 7 1 1312568547 422  13211  22 1    1        1 1
 81  12154662374232S1 11 22 112 1 2 11     1     1 1
 9     21 1113 21312  121121          1      1
10   2145413121133242l 1211 21        1                 1
11   11 5223667464A341  11  1    1
12  1 1 3 1244A666661213  11 1111 1     1  1
13          1  3  1                 1 1
14   3131 2 12 2   111  1         11
15

    WSP indices > 99.0 IGNORED
    Sample size   755
```

FIGURE 10. Numerical histograms of the WSP index, with the data being accumulated into one minute groups.

```
                NORTH DEVIATION VS EAST DEVIATION

40        1       1        1
39          1                                1     1
38             1
37        1                                                    1
36                                          1
35                                    1
34       1 1 11
33  1             1
32                                          1
31                      1                   1
29          1     1  1                1 1  111                      1
28         1                            1              1
27          11         1    1            1
26   1                      2         1
25     1     1                          1 11 1                 1
24  1                1    1              11
23              1        1             1   1     1    1     1
2211 2        1  1              1       11  2                    1
21     1     1 1  J        1          1   11              11
20          1     1     1   1  2          1   1     111          1
19   1                      1  11     1 1 1    11            1
18   1     1       1  1              1     1       1     1
17    11              1               1                   1
16   2 1                        1                  1  1        1
15        2          1       1 1          1    111 1          1
14   1      1    2    1              2 211    1   12   1      1
13      1 12  1                     1   2   1  1       1     1 1
121 1      2   1 1             11     2  1  1              1
11   1 1      21    1 1     1  1 11     1111    1                1
10  1    11     1  1   1    1           112  1                  1
 9      13    1    1               1 1 1 21             1 1
 8  1 11 2                  1     1 1  1 1           1
711  11    1 11 2  1 1   1 12 11 21   11   11  11          1
63  211       1         11        3 1   1    1 21     1
52 1      1    1 1     1             2    1  1       1 1  1
4  1 11  11         1 1    1  2  11  111 2      1     1   1   1
3  3 11   1 21     11    1 1    1       2     2     1      21
21 11 11 1 1     1     1                1     1 2 3       11     11 1
12    1     11 11   11   1     11 1 21        11 11 31    2     1 1
                              A
Average EAST deviation (A) is   27
Average NORTH deviation is at row   13
    Sample size   386
    Average values ZN & ZE              -27.4       -3.0
    Limits on N & E angles -40   0
    Limits on N & E angles -30  30
    WSP indices > 99.0 IGNORED
```

Figure 11. North zenith vs east zenitn angle plots for all data. Case 52, KLC.

ELEVATION ANGLE VS AZIMUTH

```
EL    N                      AZIMUTH

66    3                        11  1
65    1                      1
63    3                       1 1 1
62    4                           211
61    8                       11 1 311
60   17                        3121521 2
59   29                       42617441
58   37                     1 1123B683 1
57   39                     1 1 15CC331
56   50                       18**41  1
55   31                      11 398521 1
54   19                        14A31
53    9                       12122          1
52   17                       113561
51   12                      1 25211
50    6                       1121  1
49    7                    1   2 4
48    1                         1
47    3                        21
46    1                          1
45    1                                1
43    4                        2  11
42    2                         11
41    1                         1
40    2                        1 1
39    3                        111
38    1                          1
37    1                          1
36    5                        22 1
35    4                        1111
34    3                        1 1 1
33    9                       1 1411   1
32    4                       1  111
31    2                         11
30    4                       1 2  1
29    2                     1     1
28    1                       1
27    2                        2
26    1                       1
25    2                      11
24    1                                1
23    3                       1 1 1
22    2                       11
                               A
Average azimuth (A) is  162
Sample size   357
Average values AZ & EL 162.0, 52.2
Limits on azimuth  140 180
WSP indices > 13.0 IGNORED
```

FIGURE 12. Numerical histograms of azimuth vs elevation angle. Case 87, KLC.

4.1 Observations At Adelaide

Figure 16 shows N-S and E-W zenithal angles as a function of time for frequencies of 5.5, 6.5 and 7.5 MHz over a period of 120 minutes starting at 08 UT (1700 LT) on 19 December 1988. The highest frequency data is at the top of each plot.

```
   ELEV data integrated over  60 seconds

1 1   1  13 2322211412 23  3
2       212 2256212 413 1111 1       1              1
3      113 212113313133321   11
4         1 542123 22 44113 1
5  1 12 12 1 1 11 5221112 2     1     1
6           211  1142  1      2  121 1    11    1  12
7    1  1   1 22                  11 1 122 3211 1 3336
8            1                        21312  5268322  1
9       1   1     1              1    1 1 21 3 3252 2
10                                   1 2123 3221153
11                                      11172336432
12                                       1255D555 1
13                                      1  54C0A941
14                                         3A*94232
15                                         157631461
16                                     1 11 125357422  1
17                                     11122311 5623 2
18                                        221315311
19                                     14371145456853 1
20                                    1  236383B6BE7422
21                                     1 141674788BA652
22                                       4335845D8BA64
23                                   11   112  13215171  2    1
24             2      1           2112111121 21 1 1 142111 1     1
25

    WSP indices > 99.0 IGNORED
    Sample size  1003
    AZ & EL limits  140 180,  10  90
```

FIGURE 13. Numerical histograms of elevation angle, accumulated into one minute groups. Case 87, KLC.

```
   ELEVATION ANGLE VS AZIMUTH

EL    N                    AZIM

63    1                      1
61    1                      1
59    3                    1   2
58    2                     1 1
57    2                      2
56   16                    138211
55   31                   1296742
54   49                   1B*A45
53   76                    8***42
52   65                   277**9
51   81                    ?***71
50   52                    3**911
49   24                   137714 1
48    8                     3311
47    4                     3 1
46    1                   1
                            A
    Average azimuth (A) is  162
    Sample size   416
    Average values AZ & EL 161.8, 52.2
    Limits on azimuth  140 180
    WSP indices > 10.0 IGNORED
```

FIGURE 14. Numerical histograms of elevation angle vs azimuth. Case 90, KLC.

```
   ELEVATION ANGLE VS AZIMUTH

EL    N                    AZIMUTH

66    7                      1141
65  104                      7**
64  317                    1**D9**
63  203                  1  *7***4
62   68                     A**
                              A
  Average azimuth (A) is  163
  Sample size   699
  Average values AZ & EL 162.7, 63.6
  Limits on azimuth  140 180
  WSP indices > 10.0 IGNORED
```

FIGURE 15. Numerical histograms of elevation angle vs azimuth. Case 91, KLC.

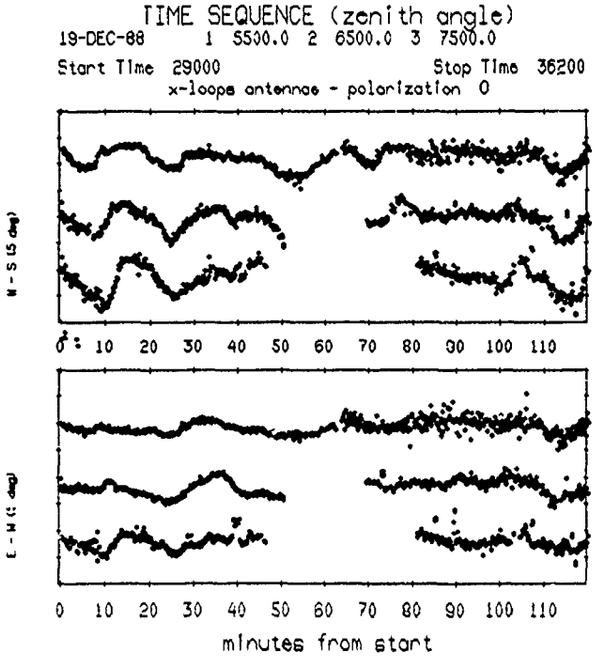


FIGURE 16. Time series of zenithal angles recorded at 5.5, 6.5 and 7.5 MHz. Adelaide.

The first 30 minutes of the N-S plot show the effects of a TID with a period of about 20 minutes. The curves at different frequencies are well correlated, with a delay from the highest frequency to the lowest of about 2 minutes per MHz. In the E-W plot, only the 5.5 and 6.5 MHz curves are well correlated during this period, with a similar propagation delay. The E-W component of the TID appears to have been dissipated before it reached the altitude corresponding to a plasma frequency of 7.5 MHz. In general, the tilts at different frequencies do not appear to be very strongly correlated.

This particular set of data also illustrates quite nicely the effects of an Es layer, which was blanketing at 5.5 and 6.5 MHz, but non-blanketing at 7.5 MHz. The layer became blanketing at 5.5 MHz several minutes before it became so at 6.5 MHz, with the reverse situation when it disappeared. It is a moot point whether or not the Es was caused by the TIDs (this effect can be seen in movies of 20-second ionograms). The 7.5 MHz data show a dramatic change in quality at minute 62, especially in the E-W component of the tilt, which becomes very spread. This change can be interpreted in terms of diffraction from discrete Es clouds, as the Es changes from blanketing to non-blanketing. The spread nature of the tilts persists for the rest of the observing period, on all three frequencies.

Figure 17 shows 4.5 MHz tilt observations for 20 December, starting at about 1700 LT. Interpretation of these data is facilitated by the virtual height plot of Figure 18, which shows that the largely chaotic tilts observed in the first hour corresponded to a period when the virtual height was double valued. In fact, the E-W plot does show some semblance of order, with the tilts increasing systematically with time - for example, between minutes 40 and 50. This type of behaviour is characteristic of larger disturbances which provide more than one reflection surface. The h'(t) traces for these show splits and Zs corresponding to large off-vertical angles of return.

4.2 Observations At Floyd

Over 100 hours of TILTSONDE observations were made at Floyd, on several frequencies, and at all times of day. The data presented here are for case 68, which was one of the most interesting of all the cases. This case was recorded at around 0900 LT on 10 October. Tilt measurements were made on a frequency of 7.9 MHz, in conjunction with AOA measurements by SKYLOC of signals on the same frequency from cooperative transmitters at Plano (range from SKYLOC of 41 km, on an azimuth of 250.8 degrees) and Bonham (45 km, 14.8 degrees). The SKYLOC observations will be described in the next section. The interesting features of the TILTSONDE tilts are the TIDs which they display, and the correlation, with appropriate time delays, of these tilts with effective tilts derived from the Plano and Bonham observations.

The time series of the tilts will be discussed in the next section, in conjunction with the Plano and Bonham observations. Figure 19 shows a plot of the north zenithal angle vs the east zenithal angle. TILTSONDE is at the intersection of the grid lines ("X"). The 20-second data have been accumulated into 1 minute sets, and a character plotted at the corresponding zenith angles. The characters start with "1" at time zero, and increase with time through to "9", then "A" to "Z", and finally through the shift characters on a VT340 keyboard. The lines connecting the points therefore trace out the trail of the apparent reflection point in the sky. Such plots are jokingly called "skytrails".

It can be seen that the skytrail is by no means a random walk across the sky. In fact, it can be interpreted in terms of a TID travelling N-S (or S-N, since there is an ambiguity of direction without Doppler information). The N-S swings of the apparent position result from reflections from E-W aligned corrugations in the iso-electronic contours.

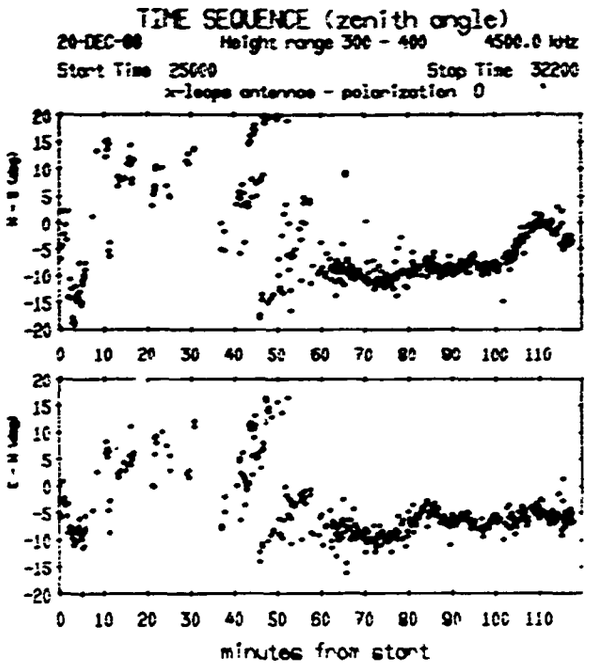


FIGURE 17. Time sequence of zenithal angles at a frequency of 4.5 MHz. Adelaide.

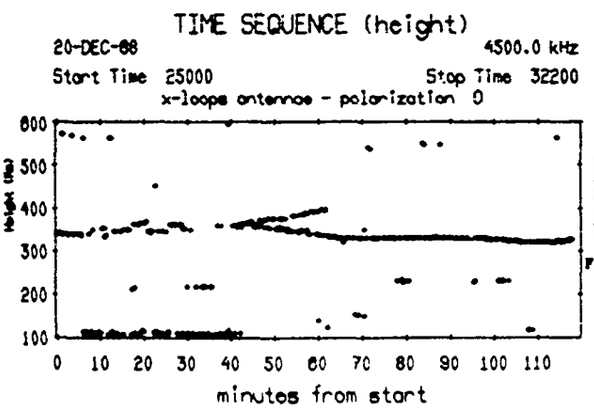
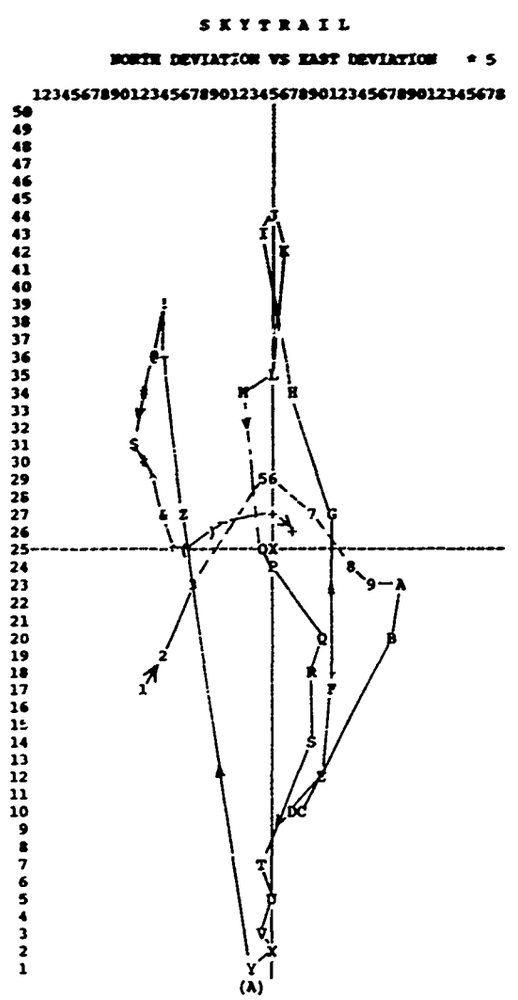


FIGURE 18. Time sequence of virtual heights at 4.5 MHz. Adelaide.



Average EAST deviation (A) is 23
Average NORTH deviation is at row 24

Sample size 49
Average values ZN & ZE * 5 -1.3, -2.2
Limits on N & E angles * 5 -25 25
Limits on N & E angles * 5 -25 25
WSP indices > 99.0 IGNORED

FIGURE 19. Skytrail for TILTSONDE tilts at 7.9 MHz. Floyd Case 68.

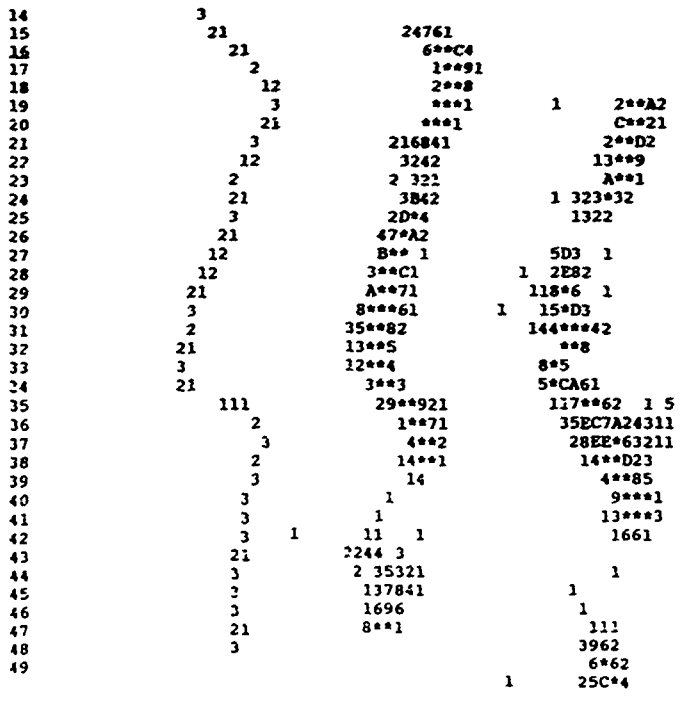


FIGURE 20. Numerical histograms of north zenithal angle for TILTSONDE, PLANO and BONHAM, with lags of 1 and 4 minutes introduced.

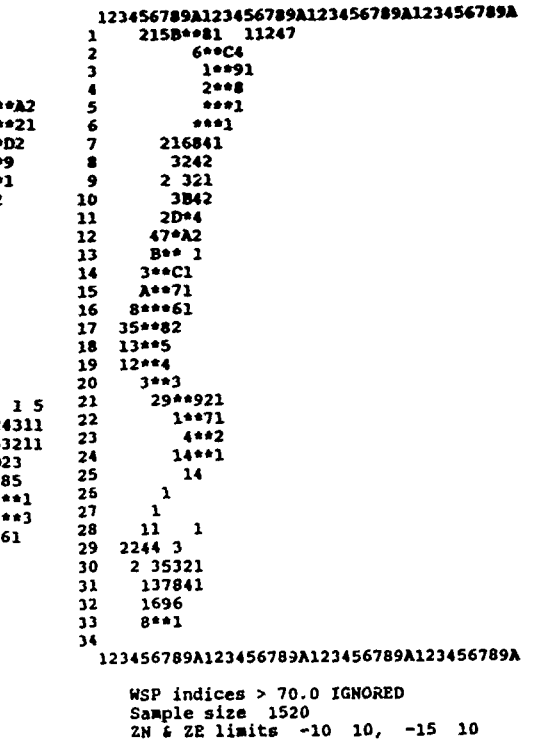


FIGURE 21. Numerical histograms of north zenith angle, with the data accumulated into one minute groups. Case 68, Plano.

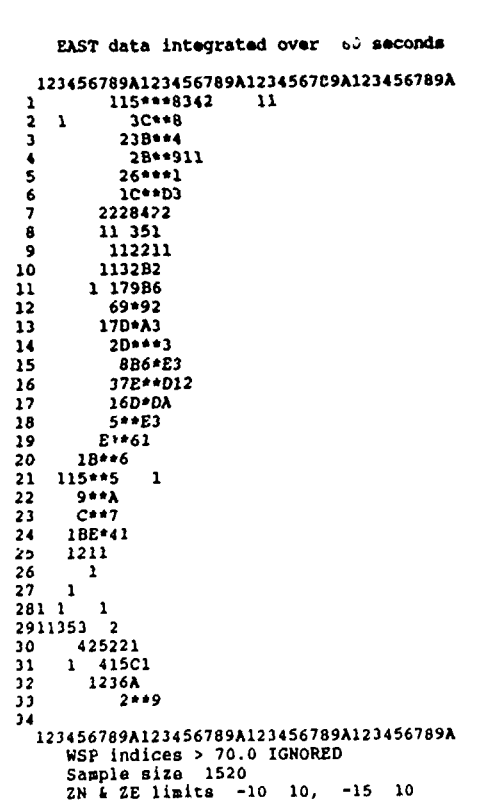


FIGURE 22. Numerical histograms of the east zenith angle, with the data accumulated into one minute groups. Case 68, Plano.

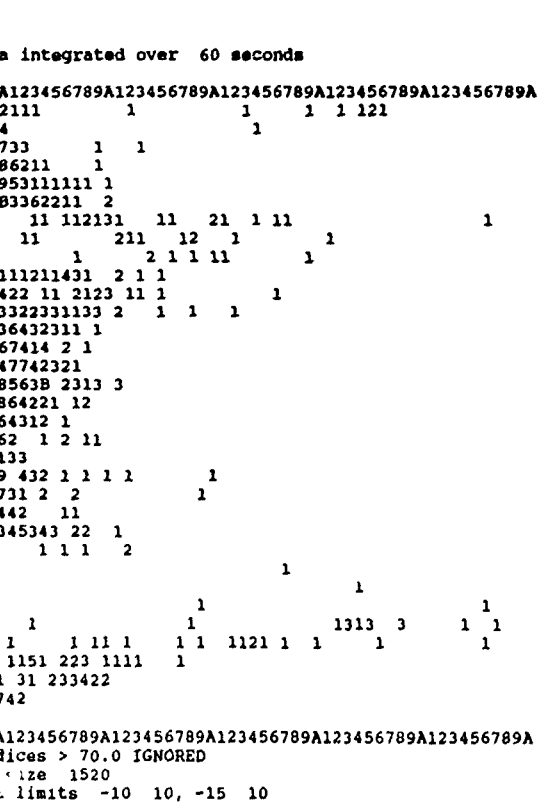


FIGURE 23. Numerical histograms of the WSP histograms, with the data being accumulated into one minute groups. Case 68, Plano.

5 SKYLOC OBSERVATIONS OF PLANO AND BONHAM

Figure 20 shows the time variations of the N-S zenithal angles for case 68 for Plano and Bonham, as well as the corresponding TILTSONDE observations. The Plano and Bonham data are effective tilts at the circuit midpoints, deduced from the AOA observations, the observed h'(fv), and knowledge of the real location of the transmitter. The three time series have been overlaid with lags of 1 minute between the Floyd (TILTSONDE) data and the Plano data, with Floyd leading Plano, and with a lag of 4 minutes between the Bonham and Floyd data. These lags correspond to a TID travelling from the north to the south at about 5 to 10 km/min, which is a typical velocity for the medium scale (period of 15 to 20 minutes) events. The E-W zenithal plots do not show a clear correlation of the E-W tilts at the three locations, suggesting that any variations are just variations along the long axis of the TID.

Figures 21 and 22 show the time variation of the zenithal angles for the Plano data. Apart from the well defined TID, the plots show an interesting very low productivity at two times during the 33-minute observing period. These occur at minutes 7 to 9, and at minutes 25 to 28. Similar decreases in the productivity occurred for the Bonham data. The explanation for this low productivity appears to lie with the WSP. Figure 23 shows the time variation of the numerical histograms of the WSP. It can be seen that not only does the total number of counts in any one minute drop at the times noted, but also that the WSP climbs to much higher values during the few minutes. This suggests that the low productivity arose from the fact that the wavefront being sampled became very distorted, leading to rejection of the most unreliable data points by the various filters, and the determination of high WSP values for the remainder. Gross distortion of the wavefront from the required plane wavefront is not unexpected in the presence of such large TIDs.

6 CONCLUSION

The extensive set of AOA observations collected at Rowlett and Floyd have not yet been analyzed in detail, with most attention to date being focussed on the interesting cases. The analyses to date have shown firstly, that SKYLOC and TILTSONDE can provide high quality AOA observations on a routine basis, and secondly, that there is nearly always one or more TIDs passing through the ionosphere at any one time.

7 ACKNOWLEDGEMENTS

SKYLOC and TILTSONDE were developed by Andrew Australia, and fielded in the USA by Andrew Government Systems, with the support of Codem Systems Inc. We would like to acknowledge the efforts of all those concerned, with a special word of appreciation to the "volunteers" who operated the systems around the clock during the field trials providing the present observations.

8 REFERENCES

1. IPS Radio and Space Services. Solar Geophysical Summary, December 1988. Dept of Administrative Services, Sydney, Australia, 1988.
2. IPS Radio and Space Services. Solar Geophysical Summary, March 1989. Dept of Administrative Services, Sydney, Australia, 1989.
3. Jeffrey, Z R, P T Middleton and C Winkler. Accurate Measurements of the Total Angle of Arrival of HF Skywaves. AGARD Conference Proceedings No 441 "Ionospheric Structure and Variability on a Global Scale and Interaction with Atmosphere, Magnetosphere". Munich May 1989.
4. McNamara, L F. Ionospheric Limitations to the Accuracy of SSL Estimates of HF Transmitter Locations. AGARD Conference Proceedings No 441 "Ionospheric Structure and Variability on a Global Scale and Interaction with Atmosphere, Magnetosphere". Munich May 1988a.
5. McNamara, L F. Ionospheric Modelling in Support of Single Station Location of Long Range Transmitters. J Atmos Terr Phys, 50(9), 781-795, 1988b.
6. SESC. Preliminary Report and Forecast of Solar Geophysical Data, 14 and 21 March. Space Environment Services Center, Boulder, Colorado, 1989.

AD-P006 298

HF/VHF PROPAGATION RESOURCE MANAGEMENT USING EXPERT SYSTEMS

D. P. Roesler
Collins Defense Communications
Rockwell International Corporation
Cedar Rapids, IA 52498

ABSTRACT

Propagation management in the HF/VHF spectrum is very complex for adaptive skywave communications systems. Understanding and readily coping with these complexities is part of an ongoing independent research and development project within Collins Defense Communications. This paper will describe the development of a propagation management system, the various resources used, and how expert systems are being applied. The latter will include the kind of rules developed and the types of expert systems explored. Because products developed are to support mobile/portable radio systems, PC-based expert systems will be used for implementation.

For benign long-term propagation conditions, tools such as IONCAP and AMBCOM represent a significant source of information to satisfy most skywave communications engineering requirements. However, in day-to-day operations, timely sources of knowledge can help HF/VHF operators adapt to changing conditions. Monitoring of government ionospheric research, conducting transmission experiments, collecting extensive data and analyzing actual HF-communications operations enable the development of a valuable propagation knowledge base.

Providing resources that can be useful to adaptive HF/VHF skywave operations requires data collection from several accessible sensors. Data from these sensors must then be processed and made available to the operator for evaluation. This is done by networking remote-sensor data to a central computer (PC or workstation). Some of the sensor data available is from the Space Environmental Service Center (SESC) (computer bulletin board and satellite broadcasts) and from HF-sounding networks. The HF operations data consist of company Communications-Central operators' observations and automatic HF-station channel-quality data. The system is being designed so that other sensors that become available/useful can be added.

Propagation rules are being developed based on observations made during ongoing research. One example to be presented is the impact of flare activity on HF point-to-point and network communications. Problems associated with "massive" automatic data monitoring, merging, processing and interpretation will also be discussed.

Expert systems used in this project have three major applications. One application is to facilitate/alleviate operator training on how to access useful tools (e.g., propagation models) or sensor data. The second is to automatically monitor the sensor data and to dynamically advise the communications operator on propagation stability and frequency selection. A third, future, application will be to automatically control the communications station.

91-09704

SYSTEM INTRODUCTION

The HF/VHF (1.6- to 88-MHz) spectrum enables communications beyond line-of-sight that are feasible for mobile platforms and independent of satellites. This type of communications relies on frequencies dependent on the state of the ionosphere. These frequencies are used for near vertical incidence skywave (NVIS), meteor burst, and long-haul skywave propagation. The development of propagation resources for real-time frequency management is part of an ongoing independent research and development project within Collins Defense Communications (CDC). The basic approach for this research is identifying useful data resources (sensor inputs), making them available, and applying them to useful HF/VHF propagation rules. This requires accessing the immense amount of propagation and communications knowledge available within the DoD and

associated government research community. Making these vast interdiscipline sources available and presenting guidance to radio operators in a real-time system can be accomplished with a type of expert system described herein.

Because the ionosphere changes with diurnal cycles, seasons, solar inputs, and magnetospheric instabilities, the optimum propagation frequency also changes. Many of these frequency variations can be consistently accounted for and are generally predicted by models, such as IONCAP, AMBCOM, BLINK, and MiniMuf. By using these models, most communications systems can be designed to accommodate the necessary frequency changes that can be consistently expected. The model predictions are used to look at the gross frequency variations a proposed system must accommodate. These predictions will dictate the required frequency ranges for the necessary antennas and RF equipments.

In real-time operations, the optimum frequency variation can be very dynamic. These dynamics appear as severe variances in signal strength or optimum frequency changes that can occur within seconds to tens of minutes. The main contribution of this project will be to eventually reveal and utilize the observables that can be used to predict propagation dynamics.

Many different ionospheric research groups have been studying the mechanisms that generate these dynamics (directly and/or indirectly) for many years. HF communicators have been dealing with these dynamics from the beginning of radio usage. This project merges these two sources of knowledge into a useful tool that can be utilized at a communications terminal.

The types of dynamics that disrupt a communications link usually occur over a small fraction of total attempted contact time. Modern equipment has improved to the point of minimizing the impact of all but the most severe dynamic conditions. Automatic scanning types of radio control, when available, help immensely in minimizing operator involvement in maintaining communications when such variations occur [1]. Such an approach is most effective when the proper channels are being scanned. The determination of the optimum scan frequency list for an automatic controller or a designated list for manual operation is a critical task. The goal for most HF/VHF communicators is to utilize the optimum frequency at all times. However, that is not always possible (or necessary) with available resources under real-time conditions. Given certain inputs and knowledge, however, real-time communications can be enhanced.

Objective:

This paper will describe the preliminary effort put forth toward achieving the following objective. The objective of this project is to amass the information and knowledge that can support real-time operations to maximize ionospheric-communications connectivity, throughput and reliability. This knowledge can be stored in an expert system shell to be made available to operators or, eventually, directly applied to control equipment. Although the CDC expert system described here, called PropMan (propagation manager), is still in the developing stage, the basic idea can be conveyed.

The first level of concentration is on propagation phenomena that have been observed to have a significant (not necessarily frequent) impact on communications. Although many of these phenomena are known and the effects documented in various sources, the implications are not readily available to the untrained operator responsible for communications. This expert system is being developed with the idea that it will be portable and readily (or automatically) accessible at the communications terminal utilizing available data resources.

System Design:

At this stage of system design, it has been assumed that all useful resources are obtainable and available. Obviously (not ideally) in certain field operations this will not be the case. Certain sites will have various types of data resources updated with varying timeliness. However, this system will be designed to automatically make use of what is available. The desired inputs, or data resources, are sensors that indicate the solar-geophysical, communications and equipment conditions. Useful inputs are certainly not limited to what is described here. There is no known finite bin of collectable observables that will predict ionospheric propagation dynamics satisfactorily for all users. With the advent of fast, small, economical computer resources, there exists a vast potential for remote sensing applications dealing with propagation. The computer resource perspective is also applied in the propagation modelling arena. Almost any model can be, or will be, available in near-real time or on-line to support critical communications links. The most useful models will be the ones that can utilize available real-time sensor data. The availability/collectability of the sensor data is a very important issue. The

more passive and robust sensors are the best choice for many reasons. This leads to a lot of creativity in indirectly utilizing the local environment for various useful propagation/ionospheric indicators.

The first emphasis on this system design will be to isolate inputs/events that have been revealed to coincide with disturbance of benign propagation. (Obviously additional and future revelations will need to be accommodated.) These propagation disturbances will be described by two temporal categories: short term (seconds to hours) and long term (hours to days). Many modern HF systems are susceptible to solar-geophysical events that will lead to blackouts without warning. This need not be the case. A system goal is to establish rules that generate a warning for the radio operator or network of impending communications disturbances. In a multimedia communications system (LF, MF, HF, VHF, UHF ...), this will enable the operator or system to reroute traffic (to other frequencies) or broadcast a warning to the operators for the impending disturbance.

More detailed knowledge about the underlying form of the disturbance will be reflected in rules for recommended actions in reacting to the warning. This leads to a second emphasis to establish rules characterizing the disturbance. Two categories in this area are indirect and actual characteristics. A polar cap absorption (PCA) event will not directly impact a midlatitude path. However, a PCA event will have an actual impact on the polar paths. These are examples of an indirect and an actual disturbance characteristic rule respectively.

A third emphasis in system design is to automatically/remotely capture the observable data that is useful. The problem is two-fold in that there is limited knowledge of the data available and its use. There exists a tendency to make rules out of only the data available (narrow approach). There is also a contrasting tendency to focus only on the ideal comprehensive solution that is not available and useable at the present time. There does exist a compromise that will enable an iterative development scheme starting with the sensors available and incorporating future ones.

Each sensor input deemed useful will require a unique interface into the expert system. The most generic approach is to format unique sensor data into a file for the expert system to access. These files can then be accessed by diskette, hard disk or local area network (LAN). The sensor data can then be automatically incorporated into the expert system shell or put into it manually. Figure 1 illustrates the PropMan System.

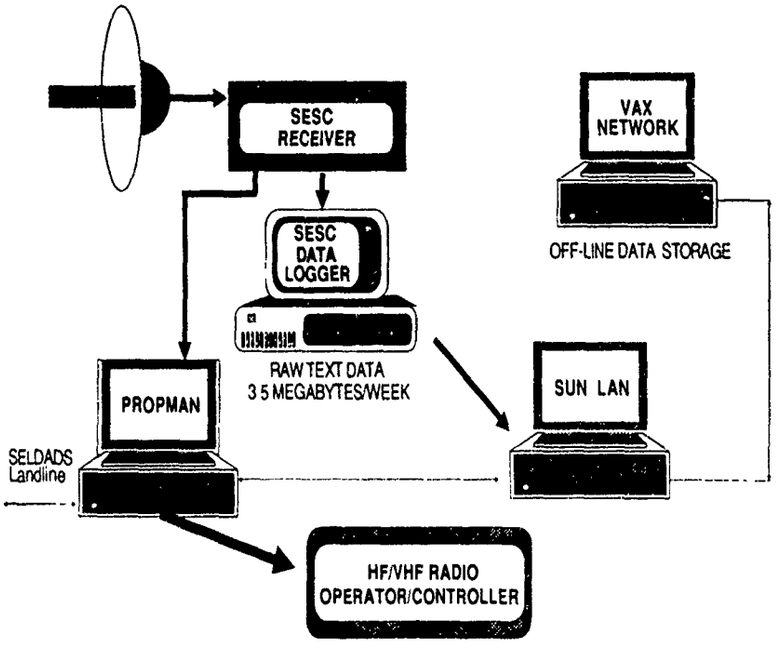


Figure 1. PropMan System.

Internal Propagation-Disturbance Research:

Communications experiments performed internally (within CDC) have observed dynamic (seconds to minutes) frequency and fading fluctuations. Some of these fluctuations are readily observed to correlate with certain data from the Space Environmental Services Center (SESC) satellite broadcasts. Other fluctuations do not result from readily observable solar-geophysical events. These internal fluctuations are more dependent on the global ionospheric plasma characteristics. Much remains to be done in characterizing (in real time) the plasma that influences, sustains and/or perturbs the ionosphere used for HF/VHF communications. Research groups that are involved in the diurnal and seasonal fluctuations of the ionosphere are not necessarily considering communications applications. Real-time adaptive HF/VHF systems need to be aware of new real-time models that the ionospheric researchers find. Incorporation of these new models and sensors will be necessary until the ionospheric medium is satisfactorily understood and characterized.

For the present, internal research has concentrated on the SESC solar-geophysical data which is readily available via the satellite broadcasts. The SESC satellite broadcasts are received every minute and contain roughly 300 bytes per transmission. The SESC data consists of by-the-minute, hourly, daily and alert data on solar-geophysical measurements.

For HF communications channel data, various test signals used in internal experiments, such as the Advanced Link Quality Analyzer (ALQA) [2] and the (Selective calling and scanning) SELSCAN automatic comms processor (ACP) LQA, are used. CDC experiments use periodic soundings on designated (licensed) frequencies and scan lists to indicate the hourly variations and qualities. An on-going CDC study to measure HF channel characteristics is the step-sounding experiment between Cedar Rapids, Iowa and Dallas, Texas. ALQA units are used as a signal generator and a detector. The sounding procedure transmits on one of 30 frequencies every minute. The frequencies used range from 2 to 88 MHz. The date, time-of-day, frequency, and the ALQA parameters are automatically collected and logged. The ALQA detector measures signal-to-noise ratio, multipath spread, Doppler spread, fading rate and other channel parameters. This testing has been performed for the past few years, enabling long-term and short-term HF characteristics to be observed between Cedar Rapids and Dallas.

Monitoring of a SELSCAN network has also been automated to continuously log the digital data. This data consists of date, time-of-day, channel and LQA values. This data can be used to determine how certain locations within the network are impacted by propagation disturbances.

The two major sources of solar-geophysical data accessible by the expert system are the SESC and the Space Environment Laboratory Data Acquisition and Display System (SELDADS). Even though they have similar origins, each can be accessed through a separate medium. Thus, if the satellite link goes down for SESC, the SELDADS can be used via landline. There also exists another landline source, the SESC bulletin board, that contains only the previous day's data.

SELDADS contains an immense amount of solar-geophysical data. SELDADS data has been found useful for determining propagation disturbances. Such data can be made available to the expert system through the use of automated PC retrieval. In SELDADS, there exists the USAF/NOAA bulletins and geophysical activity summaries. The USAF/NOAA HF bulletin is prepared at the Space Environmental Support Branch, Air Force Global Weather Central (AFGWC). This bulletin summarizes and forecasts HF propagation reports every six hours. The information supplied in this bulletin establishes gross conditions for the northern hemisphere. Although the information is very general, the indication of high-latitude propagation conditions has been very useful in supporting auroral/polar-path propagation analysis. Explanations describing certain significant events are also supplied. This is a very valuable service that real-time HF operations should use for determining current default conditions within the last six hours. Another useful bulletin is the joint USAF/NOAA solar and geophysical activity summary for the previous day. This enables a user to have knowledge about energetic events such as flares and noise bursts. Also included are the solar indices, magnetic indices, proton events and summaries. The solar event log is another useful source for monitoring optical and RF events. This data is updated every hour, enabling a user to get the previous hour's activity of the current day. The SELDADS data is kept available on the bulletin board for the previous 30 days.

In determining disturbed conditions, current values of certain types of solar-geophysical information should be compared to established quiet-time values. Descriptive statistics have been (and continue to be) computed to determine averages and deviations for values, such as the Boulder total magnetic field, proton fluxes, proton influences, geosynchronous magnetic field components, X-ray fluxes, X-ray influences, and others. This enables the expert system to know

immediately which values are abnormal. This type of information is then used to match up propagation rules that can be useful for the current communications scenario.

When radio operators initially look at all of the potential information, they quickly become saturated by the quantity. Knowing what is currently deviating significantly from the mean usually requires an unavailable amount of training and observation time. The time necessary to manually access, retrieve, or log the SESC or SELDADS data usually interferes with communication operations. A major convenience to the operator is for the expert system to do this work in parallel with the communications mission. The expert system will get the useful information and display to the operator only what is important to the immediate communications scenario. Ionograms are a good example of how useful information can be too complex or time consuming for a communications operator to analyze while handling communications traffic.

The continuous monitoring and displaying of changing SESC data has enabled some interesting observations. For an example, the SSN and solar flux for the previous day is broadcast every five minutes along with the 90-day mean solar flux. Due to the natural rotation of the sun, the solar flux and sunspot number vary with a 28-day cycle. SESC broadcasts the previous day's Boulder SSN, 10-cm solar flux and the 90-day mean solar flux. Figure 2 represents the 1989 daily solar flux and sunspot number as broadcast by the SESC. Long-term propagation models use the average for the month. This makes it somewhat ambiguous to use a real-time SSN for propagation predictions with these models. The day-to-day variation of the SSN has not been observed to immediately correlate with day-to-day MUF variations. SSN and solar flux have been observed to vary by a hundred units in a 24-hour cycle without a simultaneous proportionate change in MUF. There can be up to several days of delay for the LUF/MUF to respond to solar irradiance variations.

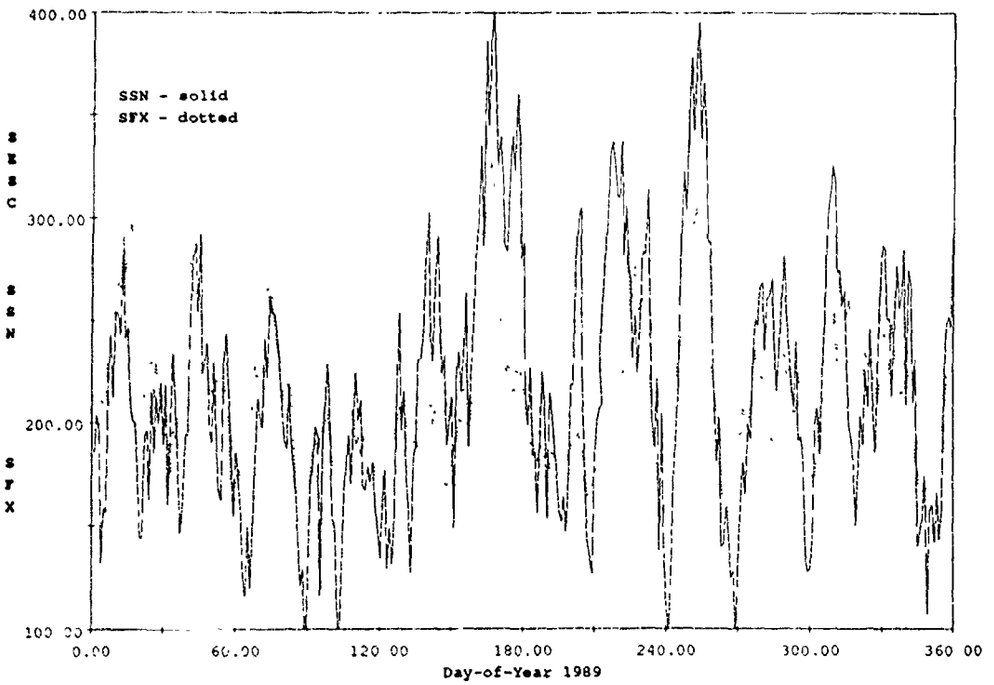


Figure 2. 1989 Daily Solar Flux and Sunspot Number.

Since IONCAP is currently being used as a general baseline to describe benign conditions, an approximate SSN will be satisfactory. In this system a default sunspot number will be used for the IONCAP SSN input. The SSN number is to correspond to with the amount of solar activity that has been radiating the ionosphere. This approach for determining a satisfactory SSN for this system application can easily be modified when observational evidence is apparent for improvement to FOT forecasting.

Rule Developments:

The most blatant short-term disturbances that impact HF communications that have been available for real-time observations are the X-ray flares. The X-level (1-8 angstrom flux > 10E-4 watts/square meter) X-ray events have been observed to severely impact HF communications data. X-ray events starting at M5 (flux = 5E-5 Watts/square meter) have been observed to cause a significant disturbance. The duration, peak value, optical size and timing of the flares are all factors that must be considered in creating rules in the expert system. The time of day that the flare occurs will determine if a short-term effect will be noticed. A night-time propagation path is usually not impacted by flares unless the path crosses the auroral oval. The immediate short-term and long-term effects of these flares have been observed several times during this solar-cycle peak. Figures 3 and 4 indicate HF step-sounding data and X-ray flux, respectively, representing a direct short-term disturbance. Figure 3 represents the HF channel data (ALQA) observed over three step-sounding cycles. Frequency (0-25 MHz) and SNoR (greater than 25) are both represented on the vertical axis. Day-of-year is on the horizontal axis. Figure 4 represents the X-ray flux (1-8 angstroms) on the vertical axis. These two figures represent data recorded simultaneously. Notice how the X-ray flux is increasing at the end of the first step-sounding cycle. The second cycle indicates the effect the X-ray flare is having. The third cycle indicates HF recovery at the higher frequencies. The time duration between cycle 1 and cycle 3 is 60 minutes. The flare had at least a 30-minute impact on the HF channel. This impact caused the low frequencies to become unusable and higher frequencies to propagate. Several examples of this type of event have been recorded during 1989.

Based on some basic knowledge of the ionosphere, general scenarios can be envisioned. If a large flare occurs causing an extraordinary high influence of high-energy protons (several hours later) and the magnetic K index at Boulder (or preferably College, Alaska's magnetometer) is greatly increasing, the auroral oval is going to become very large. Knowing the auroral oval's diurnal geographical movements, it is possible to predict the likelihood of it impacting the propagation path in question. Knowing correlations from the CDC transauroral HF experiment, TAHFE [3] [4], performed between Barrow, Alaska and Cedar Rapids enables the presentation of the communications implications to the operator.

The expert-system rules know the quiet-time averages (hourly, daily, monthly, yearly) for the X-ray flux and the proton flux (available through various sources). If data is available to the expert system to indicate that the current values are far beyond the norm, a warning and advisory can be made to the operator. The communications implications based on events (either through research or on-site data recorded experience) can be contained in the advisories. For an example, the midlatitude MUF's may be significantly impacted by the midlatitude trough (and other effects) from 24 to 36 hours after a flare occurred, depending on the size of the auroral oval. The sunrise/sunset transition-time effects may also be more dynamic and weaker 24 to 36 hours after the flare.

The rules set up in the expert system are based on the known cause and effects that have been observed. Some correlations result in observed events that coincide with certain communications impact. These impacts seem to correlate to HF propagation disturbances but are really only indirectly related to the actual source of the disturbance. If these events are observable, and the correlations are repeatable, then the resulting rules will still be used.

Many major flares induce polar cap absorption that has blacked out high-latitude propagation for several days at a time. The August 12-19, 1989 flare activity consisted of five X-level flares that blacked out most HF propagation at latitudes above 60 degrees north (based on AFGWC HF bulletins). Except for the first flare, the rest occurred at or after local (Cedar Rapids) sunset. The first flare occurred at around 1425Z on August 12th and caused a major disturbance (instantaneous complete blackout for at least five minutes) to daylight propagation paths within the continental US until about 1622Z. The PCA was observed within minutes by monitoring the Thule riometer and within an hour by monitoring the Eielson AFB sounder, both available in the SESC broadcast. Within SELDADS, at 1800Z, the next AFGWC bulletin indicated their observations. The subsequent flares that occurred at or after sunset had little effect on midlatitude HF propagation but continued to fuel the PCA event. Keeping track of and presenting (interpreting) these types of developments automatically to the radio operator will be a significant contribution of the expert system.

System Implementation and Operation:

The goal of this project is to reveal and apply useful relationships between observable solar-geophysical data and disturbed communications. The expert system goal is to display the outcomes of rules that have been developed (and are still being developed) to someone or something that can use them. This can be done using internally developed software or a

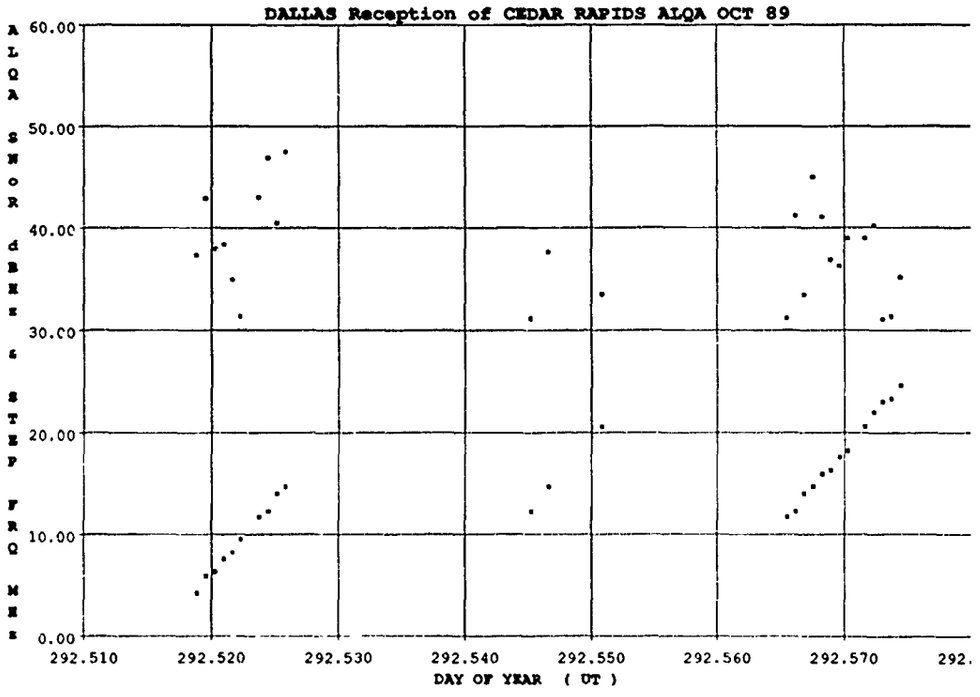


Figure 3. Cedar Rapids to Dallas ALQA Stepsounding Data.
(14 Minutes per Horizontal Tick Mark)

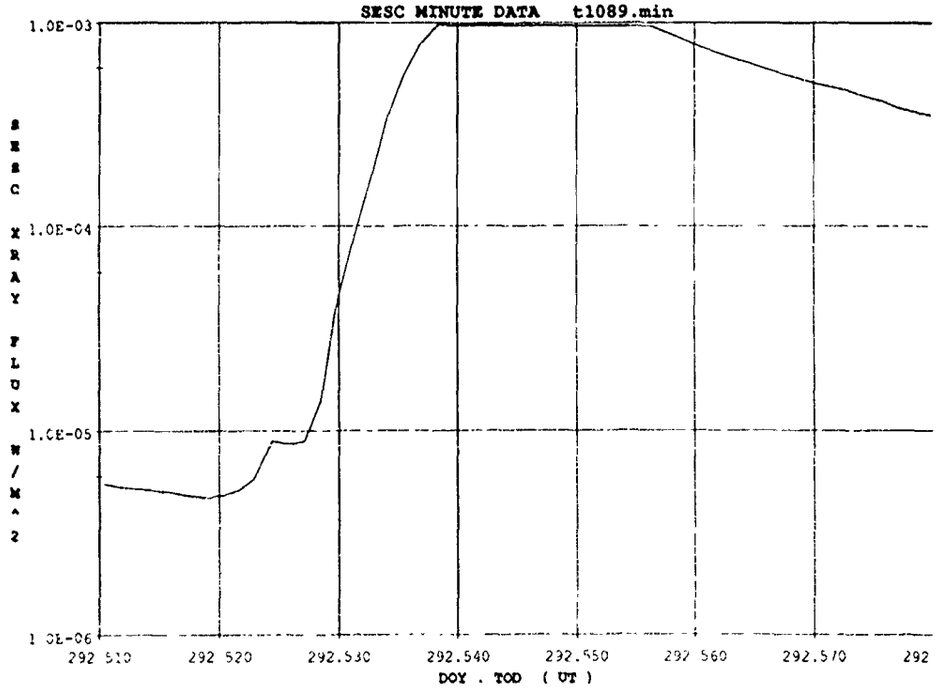


Figure 4. SESC X-ray Flare Data

commercial package. A commercial package can potentially constrain development of a project (of this nature) to being limited to the package's applications. This project has utilized a commercial expert system package (TI's Personal Consultant Plus) but has also developed software to streamline the data for the most efficient capture by the commercial software. Since LISP is used in the commercial software, the package runs very slowly when using LISP functions. It is much more efficient to create customized PASCAL or assembly routines that can execute much faster for data retrieval or sorting functions. Experience in using a particular commercial package is also a factor that is very important. An initial attempt at developing a customized expert system using an earlier choice of expert system shell was too time consuming and resource demanding. Operation and development on an expert system can become more resource demanding than the research of the actual problem. The package selected for this project enables execution of external programs for customizing PropMan for specific applications without excessive development time.

Since the primary goal of this expert system is to present useful propagation and communications information to a controller, simplicity is desired. This system (on an IBM AT clone) has been created with a default bootstrap. The only necessary input to start operations is current date and time-of-day on the PC clock. A configuration file is set up for the specified site location and the primary communications site. A menu can also be selected to change site locations for the automatic IONCAP execution. When the system is turned on, IONCAP (ITS PC.25) Method 16 will automatically be executed for the current and subsequent hour (which currently requires 15 seconds per hour to execute). All of the IONCAP information, including the predicted optimum frequency (FOT) between the primary site and another site, will be immediately available to the system. With this information, the system is immediately aware of the benign long-term propagation predictions based on the current part of the diurnal cycle, season, and sunspot activity. A 10.5-year sunspot-cycle curve has been generated and stored in a file based on the 1976 to 1986 sunspot cycle.

It is well known that long-term model predictions for frequency of optimum transmission (FOT) or MUF's will have various degrees of accuracy when compared to real-time data. The usefulness of the IONCAP program is to enable the prediction of a useable band of frequencies. This will give the operator/controller a list of frequencies under benign conditions. The IONCAP prediction will reflect the consistent frequency variation due to the diurnal and seasonal changes. Thus, when a disturbance occurs, the frequency variation can be compared to the norm.

Once the system is up and running (currently 2 minutes for bootstrap), the system will monitor any inputs available and display to the user current advisories, IONCAP-predicted LUF, FOT, MUF, predicted received SNoR's for other frequencies, site locations, distances and recommended actions. The SESC data controller is physically near enough the expert system to enable continuous monitoring (RS232 serial lines). The SESC data stream lasts for about 25 seconds in each minute. An assembly routine is executed by the expert system to capture specified data within the SESC data stream. This data is then evaluated by the rules that produce the necessary responses. If the SESC satellite link goes down, the expert system can sense this and will execute modules to access SELDADS or the SESC bulletin board periodically. The immediately useful advisories that can be presented to the operator are the warnings of disturbances, how these disturbances will impact the communications link, and what can be done in frequency management and/or network control.

Conclusions:

With all of the research performed over the last 30 years, a tremendous amount of experience in both ionospheric research and communications has been available. What is described here is an attempt to gather such useful knowledge into a black box immediately accessible to a radio operator or controller. Much of the past knowledge, although useful, is from a different era where data logging, data analysis and remote sensing was much more time consuming and expensive. Modern computer resources have opened a new arena to pursue many new perspectives in HF propagation research. The resources that are available within the US Department of Commerce, NOAA, SESC, the National Geophysical Data Center and others are immense. Utilizing those resources (both knowledge and data) for communications applications must be accomplished. The expert system is a tool for interpreting and presenting this information in an efficient and useful manner. Finding solutions to HF blackouts, maximizing throughput, reliability and connectivity should be the primary focus. Future progress needs to be harnessed in a manner similar to the system described in this paper. This systems's design will enable future upgrades to embody new developments.

Future Goals:

Ideally, in a future scenario, the PropMan expert system will be compatible with a software package (e.g., CDC's Command Control Management System) that controls all the RF equipments and requires minimum operator intervention. Wave-form decisions that can be adaptively applied to programmable modems will also be eventually included.

The developing Air Force sounder network, the Navy chirpsounder network and sounder networks around the world have a great potential for future global ionospheric modelling. This data may eventually be available to future real-time communications networks utilizing multimedia information systems. PropMan development will utilize such resources when available.

The Advanced Prophet system developed by NOSC [5] is an excellent example of a comprehensive approach in analyzing various communications scenarios. This system has an X-ray flare model that will use available real-time flux values [6]. Since this model is not yet available (other than MiniMuf 3.5 and MiniMuf 85) to this project, a comparison of predicted frequency recoveries with CDC's observed data has not been made. Advanced Prophet has many useful modules that could be merged with an expert system [7]. Future access to this system for analysis and comparisons will be sought.

A propagation modelling/prediction scheme needs to be developed for real-time communications applications. Knowing how seasonal, diurnal, and averaged solar activity impacts long-term ionospheric propagation is not enough. It is necessary to focus on hourly and by-the-minute inputs that can change/drive the ionosphere. Theoretical analysis of the ionosphere should be made available describing the perturbations and flow of the ionosphere in real time. There are many research efforts in progress for various special interests. It is hoped that, in the near future, a multiorganizational effort will be enacted to solicit cooperation with the many available research and experimental resources. Hopefully, each organization will benefit and will be able to utilize more reliable and higher throughput ionospheric communications for their own special applications.

REFERENCES

[1] Ames, J. W., "Development and Testing of Adaptive HF Radio Techniques," (SRI) DNA-TR-84379, October 1984.

[2] Bliss, D. H., "A New HF-Link Parameters/Quality Analysis Approach," MILCOM'83 Conference Proceeding, October 1983.

[3] Bliss, D. H., Roesler, D. P., Hunsucker, R. D., "Preliminary Results from a Trans-Auroral HF Experiment," MILCOM'87 Conference Proceeding V4., October 1987.

[4] Roesler, D. P., "Analysis of Data from A Transauroral HF Experiment," Rockwell International Working Paper WP88-2010, March 1988.

[5] LaBahn, R. W., "Operational Users Manual for Advanced Prophet on MS-DOS-Based Microcomputer Systems," NOSC TD-848, August 1985.

[6] Argo, P. E., "X-ray Flare and Short-Wave Fade Duration Model," NOSC TR/255, August 1978.

[7] Rose, R. B., "Prophet and Future Signal Warfare Decision Aids," DTIC, AD-A205 615, January 1989.

AD-P006 299

FREQUENCY SELECTION AND INTERFERENCE EFFECTS FOR METEOR BURST SYSTEMS

David J. Cohen
U.S. Department of Commerce
National Telecommunications & Information Administration
Annapolis, MD 21401

ABSTRACT

The factors affecting the choice of frequency for a meteor burst system are identified and reviewed. Operationally, various physical models of the meteor burst reflection process are commonly used to predict the following parameters: information throughput, burst length, D region absorption and amplitude of the signal return. These parameters have specific frequency dependencies. Additional operational parameters such as interference, frequency allocations, and the prevalence of other ionospheric propagation modes at VHF frequencies also are frequency dependent. These frequency dependent parameters are analyzed to aid the meteor burst system developer in choosing a frequency of operation.

INTRODUCTION

The geophysical characteristics of meteors and the ionosphere determine the bounds on the frequency region for meteor burst system operations. Below 30 MHz the signal returns from the meteor scatterers are masked by ionospheric reflection returns and above 100 MHz signal returns from the meteor scatterers are much too weak. The optimum frequency region for operation of meteor burst systems is considered to be the 40-50 MHz band. The particular choice of frequency for a particular application is made by a careful consideration of geophysical and operational factors along with assuring compatibility with the electromagnetic environment. These considerations are the subject of this paper.

GEOPHYSICS

The magnitude of the signal return from an underdense meteor trail reflection is [Sugar, 1964]

$$P_R(t) = \frac{P_T G_T G_R}{r^3} S(\xi) \exp[-T(\xi) f^2] \exp(-2t/t_o) \quad (1)$$

where

$P_R(t)$ = received power

P_T = transmitter power

G_T, G_R = antenna gain

$S(\xi), T(\xi)$ = geometric and physical dependent factors

t_o = $U(\xi)/f^2$ where $U(\xi)$ = geometric and physical dependent factors.

91-09703

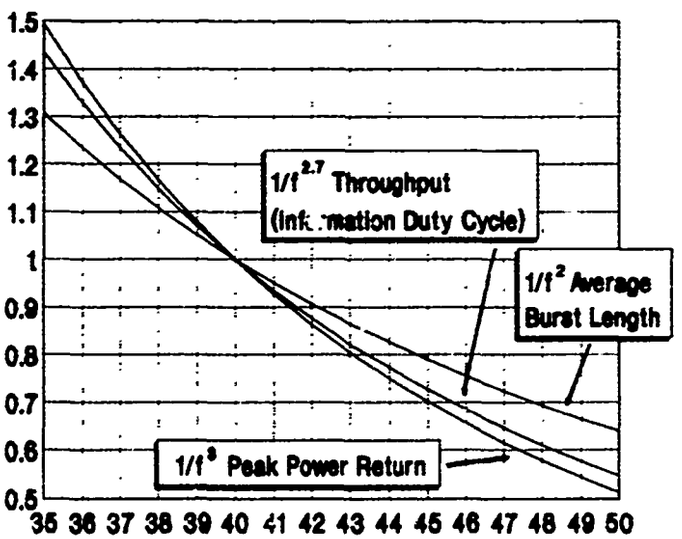
Note that the frequency dependence is f^{-3}. The peak signal return for overdense echoes also has a frequency dependence f^{-3}. Considering both underdense and overdense echoes, the average channel time duration of a meteor reflection varies as f^{-2}. Combining these two factors together and including galactic noise (which determines the required threshold signal) which varies as $f^{-2.3}$, the frequency dependence of the information duty cycle (I_C) for a meteor burst system is

$$I_C \sim f^{-2.7} \qquad (2)$$

where

I_C = information duty cycle (CCIR Report 251-4).

Figure 1 plots together each of these frequency dependent terms over the frequency range 35-50 MHz. In Figure 1, each of the terms is normalized to a value of 1 at f=40 MHz. It can be seen that the Information Duty Cycle, I_C, which is directly related to the throughput of the meteor channel is about three times larger at 35 MHz than at 50 MHz. The peak signal power level is also about three times larger at 40 MHz than at 50 MHz. Similarly, the average time duration of a meteor signal is twice as large at 35 MHz than at 50 MHz.



Frequency Dependent Terms in the Range 35-50 MHz.

Figure 1

Figure 1 suggests the use of a frequency below 40 MHz for an undisturbed ionosphere. However, when frequencies below 40 MHz are used, other ionospheric modes such as Sporadic E, ionoscatter, and F region reflection are more likely. These additional propagation modes occur more frequently as the frequency decreases below 40 MHz. When these modes are prevalent, their signal strengths will tend to exceed or mask the meteor signal return. These modes, although a bonus, may result in multipath effects that may degrade the communications capability.

The U.S. government allocated bands in the range 35-40 MHz are the fixed and mobile bands 36-37 MHz and 38.25-39 MHz. These bands, below 40 MHz, should be considered for government meteor burst use, particularly if there is frequency congestion and groundwave interference in the optimum frequency region 40-50 MHz. From a multipath and interference perspective, the best times are those when other ionospheric propagation modes are less likely to occur. The COMET system operated successfully in the early 60's using frequencies between 35-40 MHz [Bartholme and Vogt, 1968].

FREQUENCY SELECTION ABOVE 50 MHz

The attenuation effect on meteor burst systems of D region absorption is under 1 dB for normal ionospheric conditions and is not included in determining system parameters. It is generally the case that to ameliorate D region absorption effects, meteor burst systems use higher frequencies above 50 MHz. The usual assumed expression for the absorption loss has a $1/f^2$ frequency dependence. The choice of a frequency during an absorption event is not simple since the peak signal level is dependent upon two competing factors [Ostergaard et al, 1986]. Both the signal amplitude and the absorption decrease with frequency.

An additional factor during absorption events is that the received background cosmic noise level (cosmic, galactic noise) may be less because of the added absorption of cosmic noise when it comes through the D-region of the ionosphere. Another loss factor effect, Faraday rotation, shows no clear frequency dependence [East et al, 1987]

Regulatory frequency support for meteor burst systems is also an issue above 50 MHz. The only government allocated band in the frequency range 50-100 MHz which permits transmitter power greater than 1 watt is 74.8-75.3 MHz. This band is allocated to the Aeronautical Radionavigation Service and is not appropriate for meteor burst use. However, government operations may employ frequencies in nongovernment bands between 25-2400 MHz for tactical and training purposes subject to regulatory conditions (NTIA Manual, Section 7.15.3).

INTERFERENCE EFFECTS

An important factor when choosing a frequency of operation for a meteor burst system is to insure that the meteor burst system is compatible with existing equipments in the telecommunications environment. The choice of frequency should take into account possible interference to and from meteor burst systems. VHF propagation mechanisms [Weitzen et al, 1987], other than meteor burst trails, which may cause interference in meteor burst system environments are: ionospheric (sporadic-E, auroral scatter, ionospheric scatter) and tropospheric (diffraction, line of sight propagation, and troposcatter). The interference potential of these propagation mechanisms is discussed in the sections which follow. A complete description of the electromagnetic compatibility of meteor burst systems is found in Cohen et al [1989].

IONOSPHERIC INTERFERENCE EFFECTS

CCIR Report 259-6 [1986a] identifies those ionospheric reflected signals which may lead to interference at frequencies between 30 and 300 MHz. The most prevalent of these interference signals are reflections or scatter from: (1) meteor ionization, (2) Sporadic E reflection and scatter, and (3) regular F-layers. Because the maximum propagation range of these types of ionospheric signals can extend to ranges of 500-4000 km, there is the potential for interference over large geographical areas. The most likely occurrences of each of these interference sources is time (e.g., diurnal, seasonal, and yearly) dependent and also spatially dependent (i.e., low-latitude, temperate, equatorial). The actual times of occurrences of each of these interference signals can only be predicted statistically. Each of these different types of ionospheric propagated interference signals will be examined separately.

METEOR IONIZATION

The signals from meteor trails are short lived and may cause at most a short noise burst interruption to a wanted signal. To perform an interference analysis it is assumed that received signals from meteor trail

reflections have signal magnitudes of - 110 dBm with returns separated on an average of 4 to 20 seconds. The average duration of a signal is assumed to be .58 seconds [Oetting, 1980]. Miller and Licklider [1950] have determined the impact of short interruptions on percent word articulation score. Figure 2 shows the Miller and Licklider measured listening performance data for random interruption bursts of noise. Figure 2 plots noise-time fraction (the proportion of time the noise is on) versus frequency of interruption (number of interruptions per second) for equal articulation score contours. The typical meteor burst signal described above has a noise time factor = (.58/4) = .14 and the number of noise interruptions/second = (1/4) = .25. This point is plotted in Figure 2.

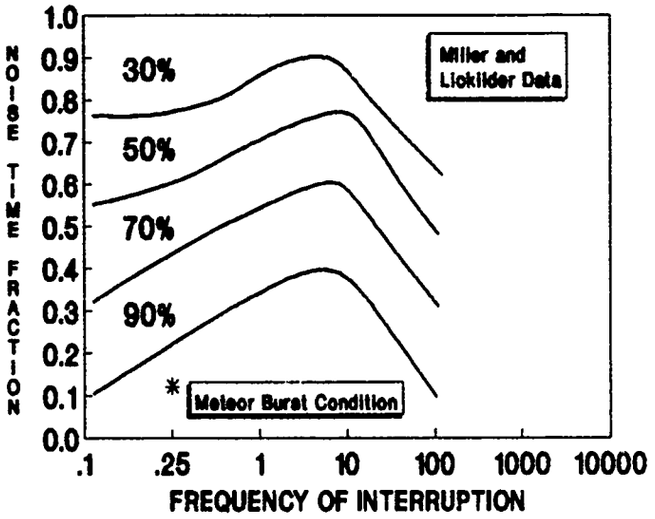


Figure 2

The location of the meteor burst operation point in Figure 2 shows that the percent word articulation score in ionospheric meteor burst interference is above 90% which is more than acceptable for most, if not all, communication purposes. Moreover, the typical signal return of -110 dBm is a relatively weak signal on a par with galactic noise levels and is unlikely to cause interference to other VHF systems such as land mobile. The possibility of interference is further reduced since each meteor acts as a directional antenna and focuses the energy it radiates to the ground in a small "footprint" with dimensions of 5 x 40 km. It is concluded that meteor scatter signals are not a likely source of interference to other users of the VHF spectrum.

SPORADIC E REFLECTIONS

The potential for meteor burst station signals that are reflected from Sporadic E ionization to cause interference can be assessed by use of the CCIR recommended Sporadic E field strength calculation method [CCIR Rec. 534-2, 1986]. This calculation method was used to compute expected signal returns from Sporadic E ionization for the equipment parameters: P_T = 10 kw, $G_T = G_R$ = 10 dBi. Figure 3 plots the calculated expected received power as a function of transmitter frequency and separation distance between the transmitter and receivers. The critical frequency of Sporadic E was 13 MHz for the calculations. Figure 3 shows that the signal powers from Sporadic E returns are of the same order of magnitude (i.e., -110 dBm) as meteor trail returns.

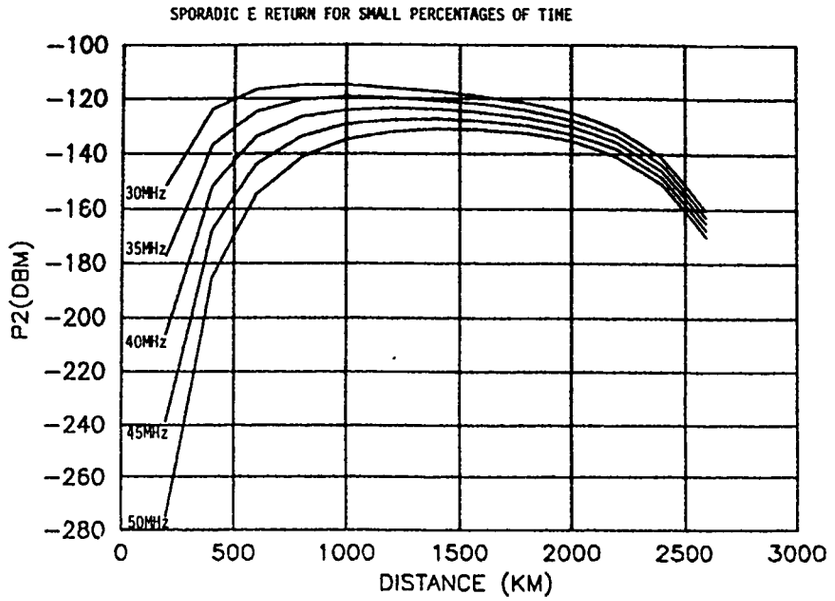


Figure 3

The maximum percentage of time for Sporadic E reflections at VHF frequencies is in the range of 1 to 8% with the probability of occurrence decreasing with frequency.

REGULAR F REGION REFLECTIONS

When the solar cycle is at a maximum, VHF signals up to 60 MHz can propagate to long distance via reflection from F_1 and F_2 layers in both temperate and low latitudes (CCIR Report 259-6]. These ionospheric signals can originate from a number of interference sources including VHF-TV, land mobile, and meteor burst. The duration of interference from F reflections can range up to a couple of hours. However, the occurrence times are random and the time percentage of occurrence is small, (about 1%), making it difficult to develop spectrum management procedures for this kind of interference. Some outage time due to ionospheric interference must be expected.

SUMMARY OF VHF IONOSPHERIC INTERFERENCE

The common quality of ionospheric interference for frequencies above 30 MHz is that the times of occurrence are random with a relatively low probability of causing a communication outage. The probability of interference will be greatest during the times of high sunspot activity. It is concluded that it is not necessary to develop specific compatibility criteria for ionospheric propagated VHF interference to or from meteor burst equipment. Real time equipment adjustments such as power, frequency or antenna changes are required to solve these interference problems.

TROPOSPHERIC INTERFERENCE

Typical equipments in the frequency range 35-50 MHz band are land mobile systems and consumer devices such as cordless telephones. Also, the TV video IF frequency is at 45.75 MHz with a bandwidth of 6 MHz (43.75-48.75 MHz) and the audio IF is at 41.25 MHz with a 200 kHz bandwidth.

The possible tropospheric interference paths between meteor burst systems and other telecommunication systems operating in the 30-50 MHz frequency region are illustrated in Figure 4.

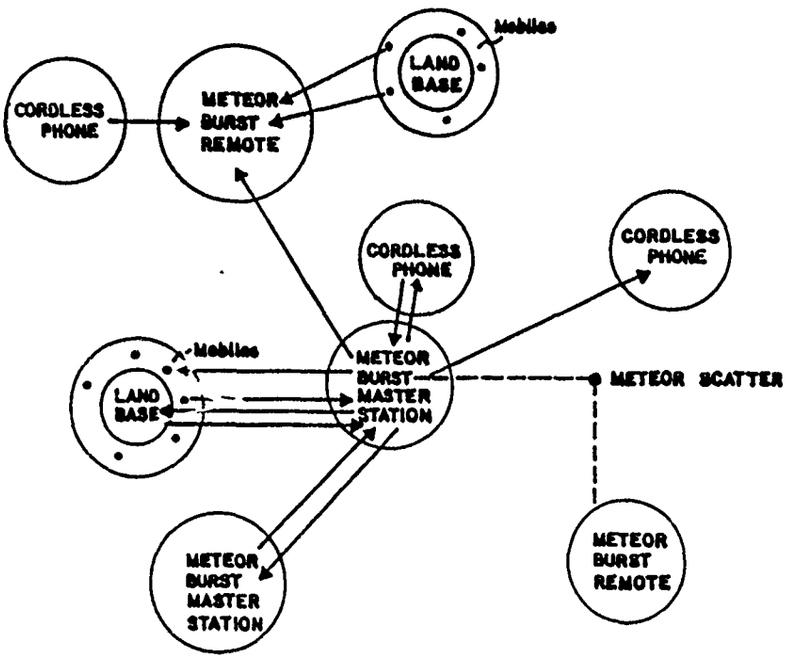


Figure 4

Meteor burst master stations may be both a source (probe signal) and receptor to VHF tropospheric (e.g., groundwave) interference. The separation distances between equipments for compatible cochannel sharing (joint simultaneous operation) with land mobile operations are in the range 90-320 km. When the equipments are on adjacent channels, the land mobile separation distances are in the range 7-60 km.

The signals transmitted from remote meteor burst stations, assuming they transmit only when polled, are intermittent and unlikely to affect the intelligibility of other VHF systems (See Figure 2). Thus, they are not judged to be an interference problem. Meteor burst remote stations, however, are susceptible to groundwave interference from other VHF equipments. The land mobile separation distances for cochannel use are 90-250 km and for adjacent channel use 10-38 km.

Non-licensed consumer devices, such as cordless telephones, operate (unprotected and on a noninterference basis) along with U. S. government systems in the upper 46 and 49 MHz. Meteor burst systems operating cochannel may receive interference when located as far as 10 km away from these devices.

Meteor burst systems themselves, if operated on the same channel, also require a separation distance to avoid mutual groundwave interference. The required separation distances are in the range 300-1200 kilometers.

REFERENCES

Sugar, G. R., Radio Propagation by Reflection from Meteor Trails, Proceedings IRE, Vol. 52, pp. 116-136, 1964.

Bartholme, Pierre J. and Irmfried M. Vogt, "Comet - A New Meteor Burst System Incorporating ARQ and Diversity Reception," IEEE Trans. on Comm., Vol. COM-16, pp. 268-278, April 1968.

Ostergaard, J. C., J. E. Rasmussen, M. J. Sowa, J. M. Quinn and P. A. Kassey, The RADC High Latitude Meteor Scatter Test Bed, RADC-TR-86-74, July 1986.

East, J. H., S. A. Maier, and I. Jamil, Polarization Rotation Effects on the Link Budget and Throughput of the Meteor Burst Communications Channel, Ionospheric Effects Symposium, 1987.

NTIA, Manual of Regulations and Procedures for Federal Radio Frequency Management, U.S. Department of Commerce, National Telecommunications and Information Administration, Washington, D.C., revised September 1989.

Weitzen, J. A., M. J. Sona and R. A. Scofidio, "Characterizing the Multipath and Doppler Spreads of the High-Latitude Meteor Burst Communication Channel," IEEE Transactions on Communication, Vol. COM-35, No. 10, October 1987.

Cohen, David, William Grant and Francis Steele, "Meteor Burst System Communications Compatibility, NTIA Report 89-241, U.S. Department of Commerce, National Telecommunications and Information Administration, March 1989.

CCIR Report 259-6, VHF Propagation by Regular Layers, Sporadic E or other Anomalous Ionization, Doc. XVI Plenary, 1986a.

Oetting, J. D., An Analysis of Meteor Burst Communications for Military Application, IEEE Transactions on Comm COM-28, pp. 1591-1601, September 1980.

Miller, G. A. and J. C. Licklider, Intelligibility of Interrupted Speech, Journal Acoustical Society of America, Vol. 22, No. 2, March 1950.

CCIR Recommendation 534-2, Methods for Calculating Sporadic-E Field Strength, Doc. XVI Plenary, 1986b.

AD-P006 300

EFFECTS OF ABSORPTION ON HIGH LATITUDE METEOR SCATTER COMMUNICATION SYSTEMS.

J.C. Ostergaard[1], J.A. Weitzen[1], P.A. Kossey, A.D.Bailey,
P.M.Bench, S.W Li[1], J.R.Katan[2], A.J. Coriaty, J.E. Rasmussen

Geophysics Laboratory (AFSC), Hanscom AFB Ma 01731-5000.

1. University of Lowell Center for Atmospheric Research, Lowell MA.
2. Naval Underwater Systems Center, New London CT.

ABSTRACT

Data acquired with the Geophysics Laboratory's High Latitude Meteor Scatter Test-Bed between Sondrestrom Air Base and Thule Air Base, Greenland during the solar disturbances of March and August, 1989 are presented. These events provided a unique opportunity to observe a number of naturally occurring disturbance effects on meteor scatter links operating in the frequency range (35 to 147 MHz) covered by the test-bed. The effects range from signal absorption to system noise variations. The properties of ionospheric absorption in general is discussed and illustrated with computations using electron density profiles from the September 1978 Solar Proton Event (SPE). It has been found that accurate measurements of high levels of ionospheric absorption with riometers pose special problems. These problems are identified and discussed. The data acquired during the March and August solar disturbances are then related to the zenith absorption measured at Thule, and the influence of absorption, as well as system noise variations, are discussed. The two events presented are very different. The August event was dominated by ionospheric absorption which affected meteor arrival rates and duty cycles primarily at the lower frequencies (35 and 45 MHz) although some effects could also be seen at the higher frequencies (65 to 147 MHz). The March event combined weak ionospheric absorption with large solar noise bursts. The effects of this event on the test-bed were dominated by increased solar noise at all frequencies. The relative influence of solar noise and ionospheric absorption during SPE events is discussed along with speculation as to the validity of frequency dependence conclusions based on testing of the JANET system.

1.Introduction

Polar Cap Absorption (PCA) events are caused by high energy (> 10 MeV) protons of solar origin, emanating from the sun during Solar Proton Events. The particles are funnelled into the polar cap ionosphere along the earth's magnetic field lines creating increased ionization at altitudes between 45 and 90 km. The collision frequency in this part of the ionosphere is high, and the increased ionization causes absorption of radio waves passing through this region. The absorption is a function of frequency and the path length within the absorbing medium. Lower frequencies are absorbed more than higher frequencies and the absorption increases with path length.

The increased ionization caused by SPE's influences all radiowaves passing through the D-region of the ionosphere. Medium- and Shortwave communication systems can be severely influenced or even interrupted due to absorption. Meteor scatter systems operating in the VHF band are expected to suffer less due to the lower level of absorption at these frequencies. Information confirming the existence of PCA influence on meteor scatter links exist (1,2), but no extensive, quantitative evaluation of these effects over the frequency range of 30 to 150 MHz used for meteor scatter communication has been undertaken.

The Air Force's Geophysics Laboratory (GL) operates a 1200 km diagnostic meteor scatter link between Sondrestrom AB and Thule AB, Greenland. The link is situated entirely within the Polar Cap and operates at 35, 45, 65, 85, 104 and 147 MHz, Figure 1. A second diagnostic link between Sondrestrom AB and Narssarssuaq in southern Greenland is 700 km long. This link passes through the auroral oval and operates at 45, 65, 85, and 104 MHz. The links, constituting the GL High Latitude Meteor Scatter Test-Bed, are used to measure meteor scatter propagation and communication properties during quiet and disturbed ionospheric conditions. The propagation properties include: meteoric signal arrival rates, duty cycles, signal durations and fading rates. From these, communication properties such as throughput and waiting time for messages of given length can be derived for a range of postulated modulations and signaling protocols. Descriptions of the links, and the results obtained have been presented in a range of papers and reports,(3,4,5,6,7,8,9,10). No SPE's occurred during the first years of operation of the test-bed. However solar disturbances in March, April, August, September and October 1989 provided both absorption events and large variations in solar noise. The data acquired during the March and August 1989 disturbances have been

91-09702

selected for presentation here, as they show a very illustrative combination of absorption and noise effects. An analysis incorporating all the disturbances of 1989 will be published later.

The absorption on a meteor scatter path during a PCA occurs in the lower D-region of the ionosphere in the height range of 45 to 85 km, well below the height of the meteor trails, Figure 2. In general, the absorption A_{f1} for one pass through an absorbing layer at the frequency f1 and at an elevation angle of z can be approximated as:

$$A_{f1} = A_{f0}(f0/f1)^n \text{seca}(z) \text{ dB}$$

where f0 is a reference frequency lower than f1, A_{f0} is the zenith absorption at f0, and n is an exponent describing the frequency dependence of the absorption. The exponent n has a value of approximately two for low values of absorption. Seca(z) is a factor describing the ratio between the path at zenith and the path elevated 90-z degrees. The factor would have been equal to the secant of the zenith angle if the earth was flat and the absorbing layer began at the surface (see Figure 3). However, due to the earths curvature, the seca factor does not increase above 6 for an absorbing layer at 45 to 85 MHz. The absorption decreases with an increase in frequency, and increases with the length of the path through the absorbing region. The influence of the seca factor should be noted as the absorption values expressed in the above formula are in dB. Thus a zenith absorption of 6 dB translates into a slant path absorption of 25 dB for a 10 deg elevation angle. A signal propagating between the stations of a meteor scatter link is attenuated twice; once propagating through the absorbing region of the ionosphere from the transmitter to the meteor trail and again from the meteor trail to the receiver. For a quiet site, the receiver noise power is dominated by galactic noise at low VHF frequencies, decreasing with frequency to the power - 2.3. The noise is also absorbed by the ionosphere, but only through one passage of the absorbing region.

Thus, the higher frequencies are less affected by the absorption than the lower frequencies, but at the same time the higher frequencies offer less communication capacity during undisturbed ionospheric conditions. The frequency dependence of the capacity can be expected to decrease with frequency to the power -2.7 (11). This exponent includes the effects of smaller scattering efficiency, shorter trail durations as frequency increases and a galactic noise limited receiving system. An optimum frequency of operation for a meteor scatter link must then exist for a given absorption level. This paper presents and examines propagation data from the August 1989 PCA as a function of frequency in an effort to determine this relationship. The above simple relations are insufficient for an accurate quantification of the frequency dependence of absorption effects during PCA's and a more general review of ionospheric absorption is presented below prior to the discussion of the test-bed data.

2. Ionospheric absorption at VHF frequencies.

The properties of ionospheric absorption at VHF is presented in this section with a discussion of the dependence of absorption on collision frequencies, electron density profiles, wave frequency and the height of the absorbing region.

Following the presentation in (12), the ionospheric absorption (A) in the D-region can be computed after the Appleton-Hartree formula as:

$$A = \text{Const} \int Nv/(v^2 + (w \pm w_L)^2) \, dS$$

where N is the electron density, v is the collision frequency, w is the wave angular frequency, w_L is the gyro angular frequency and dS is the incremental path length. This formula implies a constant collision frequency. The collision frequency is not constant, so the formula cannot be used for accurate computation of the absorption. The equation is useful, however, to illustrate that the absorption is dependent on the electron density profile, the wave frequency and the collision frequency, and that the absorption follows a frequency-square relationship if the wave angular frequency is much larger than the collision frequency and the gyro frequency. At VHF the wave frequency is much larger than the gyro frequency, but the collision frequency at altitudes between 45 and 60 km is of the same magnitude as the wave angular frequency. The frequency dependence of the absorption in this height range will therefore differ from the frequency-square relationship, and the variation from this relationship for the overall path is a function of the fraction of the total absorption which occurs in this height range. Sen and Wyller have formulated a generalized magneto-ionic theory that does not imply a constant collision frequency (13). This theory provides more accuracy in the computation of the ionospheric absorption A at VHF frequencies:

$$A = \text{Const} \int N \, a/w \, C_{5/2}(a) \, ds$$

where N is the electron density, w the wave angular frequency, a = w/v' and dS is the incremental path. $C_{5/2}(a)$ is one of a special class of integrals found in semiconductor theory. The values of these integrals can either be found in tables (14) or computed. The value of the collision frequency v' used in the expression has been discussed by Thrane and Piggott (15).

The magnitude of the expected advantage of operating a meteor scatter communication systems at higher frequencies to overcome absorption is a function of the frequency dependence of the ionospheric absorption. A frequency dependence different from the simple frequency-square relationship depicted by the Appleton Hartree formula for cases where the collision frequency is smaller than the wave angular frequency will reduce the advantage of the higher frequencies. A dependence of the absorption proportional to frequency or even independent of frequency, such as can be envisioned if the absorption exclusively occurs at a height range where the collision frequency is much higher than the wave angular frequency, would seriously limit or eliminate any advantage of operation at higher frequencies.

The ITSA raytracer (16) modified to incorporate the Sen-Wyller formulation for the complex index of refraction in the ionosphere has been used to compute the zenith absorption at noon on 23 and 26 September, 1978. The computations, presented below, are aimed at examining the frequency dependence and height of occurrence of the ionospheric absorption during a PCA event. The computations are based on electron density profiles obtained from the GL oblique VLF ionosounder at Thule (17) and collision frequency profiles presented by Thrane , Piggott and Friedrich (15,18). Figure 4 shows the electron density profiles and the collision frequency profile used for the computations. The increased ionization below 90 km, as compared to the quiet day profile, is clearly seen. The zenith absorption at 30 MHz was approximately 10 dB at noon on September 23 and approximately 3 dB at noon on September 26. The electron density increases two to four orders of magnitude below 55 km on the 23rd where the electron density exceeds $10^6/m^3$ at 48 km. The profile for the 26th shows little increase at 45 to 52 km but one to two orders of magnitude increase of the electron density in the range 55 to 65 km.

The collision frequency profile is presumed independent of the increased ionization. The temperature of the D-region is not increased more than a few degrees Kelvin during even large particle events (E.V.Thrane pers. comm.) and such changes will introduce negligible changes in the collision frequency. Figures 5 and 6 show the computed cumulative and specific absorption at noon on September 23, 1978 as a function of height and frequency. The cumulative absorption increases as expected with height and more than 80% of the absorption occur below 80 km. Significant absorption takes place between 48 and 55 km where the collision frequency is high relative to the wave angular frequency. As the meteor trails occur between approximately 85 and 110 km, the absorption takes place below the meteor region. The specific absorption is largest at 60 km for a wave frequency of 30 MHz and it decreases both in magnitude and altitude with an increase in frequency. This is expected as the maximum absorption should occur at altitudes where the wave angular frequency is comparable to, or smaller than, the collision frequency. Figures 7 and 8 show the computed cumulative and specific absorption at noon on September 26, 1978 as a function of height and frequency. The absorption is less than at noon on the 23rd and occurred below 90 km. The maximum specific absorption at 30 MHz occurred at 65 km and little change is seen between 55 and 75 km. The height of the maximum specific absorption still decreases with frequency but very little absorption takes place below 55 km. Thus, the absorption mainly occurred between 60 and 90 km on the 26th of September when the total zenith absorption was approximately 3 dB, and between 55 and 85 km on the 23rd when the total zenith absorption was approximately 10 dB. This means that the absorption is not totally dominated by the increased ionization below 60 km. The increased ionization above 60 km contributes a large part of the total absorption, even at an absorption level of 10 dB.

The frequency dependence of the cumulative absorption at noon on 23 and 26 September, 1978 are presented as functions of frequency and height in Figures 9 and 10. The absorption decreases with frequency at all heights as expected. Also, the rate of decrease is least at the lower heights, and is not linear with frequency when presented in double logarithmic coordinates. The rate of decrease of absorption with frequency gets larger with height, and also more linear. If the slopes of the curves were linear then an exponent describing the frequency dependence of the absorption could be determined for each height. Analysis of such exponents have been presented by

Lerfald, Little (19) in the frequency range 5 to 50 MHz. The slope is not linear however, and adoption of fixed exponents to describe the frequency dependence of absorption must be accompanied by a description of the frequency range and method of fitting used to determine the slope. The exponent for the same absorption event can vary considerably for different frequency ranges. Although not strictly valid, the concept of an exponential fit to the frequency dependence of absorption can be useful for analysis of absorption data. The following exponents have been derived as a linear fit to the absorption in the frequency range 30 - 100 MHz: At noon on the 23rd the computed absorption at 50 km is almost independent of frequency, as the collision frequency is comparable to or greater than the wave angular frequency. However, the cumulative absorption at this height is only a small fraction if the total absorption. At 60 km, the absorption is almost inversely proportional to frequency, and at 90 km the absorption varies as frequency to the power -1.5. At noon on the 26th, the absorption is almost frequency independent at 55 km, and close to a frequency-square dependency at 90 km. The variations of the exponents with height for the two days are presented in Figures 11 and 12. The differential exponent increases from 0 to 2 in the height range from 50 to 70 km. The cumulative exponent reaches a value of 1.5 and 1.8 respectively for the two cases. The overall exponent is lower on the 23rd because more of the total absorption occurred below 60 km on this date compared to the 26th. In both cases, absorption occurring above 60 km constitutes a large part of the total absorption and the overall exponent cannot be totally dominated by the low values found between 45 and 60 km.

The two examples shown here do not cover all possible electron density profiles that can be expected during SPE events, but they are believed to phenomenologically represent a wide range of absorption conditions. These can be summarized as follows: The ionospheric absorption during a PCA event occurs in the height range from 45 to 80 km. The larger absorption levels contain substantial contributions from the lower part of this range, whereas the lower absorption levels are dominated by absorption in the range 60 - 75 km. Although an exponential fit to the frequency dependence of the absorption is strictly not valid, such exponents derived for the frequency band 30 to 100 MHz range from -2 at very low levels of absorption to -1.5 or less when the absorption exceeds 10 dB at 30 MHz. However, the exponent of -1.5 observed during the 10 dB absorption of September 23rd, indicates that the absorption level is unlikely to become frequency independent even during very high naturally occurring absorption levels.

3. Measurement of ionospheric absorption at VHF.

Ionospheric absorption is usually measured with a riometer operating at 30 MHz. The principles of riometers and their use is described in a range of excellent references (12,20) and only a few aspects of the use of riometers, especially important for measurement of large absorption values such as encountered during PCA's, will be mentioned here. A riometer is basically a power meter connected to an antenna pointed at zenith, and the absorption $A_{riometer}$ is measured as:

$$A_{riometer} = 10\log(P_{abs}/P_{quiet}) \text{ dB}$$

where P_{quiet} is the power measured at the antenna feed point during quiet ionospheric conditions, and P_{abs} is the power measured during an absorption event. Power can be measured either in Watts (P) or in degrees Kelvin (T) where P = kTBw, k is Bolzmann's constant (1.38 E^{-23} W/Hzdeg) and Bw is the bandwidth of the power meter. Temperature as a power measure is conveniently independent of bandwidth and will be used in the following, implying this relation between power and temperature exists. The diurnal variation of the noise during quiet ionospheric conditions is termed the quiet day curve, and is used as a reference for the absorption measurements. At 30 MHz the noise of an antenna pointed at zenith is of galactic origin at a quiet receiving site, so the decrease in the measured noise is due to absorption in the ionosphere. Friedrich (21) has determined the ionospheric absorption at 30 MHz during quiet ionospheric condition to be approximately 0.01 dB, which is insignificant for the absorption measurement.

The riometer installation, including antenna, is not a completely linear instrument with respect to the magnitude of the galactic noise power at zenith. The riometer itself has a dynamic range determined by the dynamic range of the variable noise source in the riometer. This has a lower limit set by the physical temperature of the termination resistor in the noise source, approximately 290 - 300 K. The upper limit is set by the excess noise ratio of the noise source, but this is usually much higher than the highest level of galactic noise. The output voltage from the riometer is usually the DC voltage across the termination resistor of the noise source, or other measure of the DC current flowing through it. It is usually assumed that the output

voltage of the riometer is proportional to the noise temperature at the antenna feed point less 290° K. This assumption is valid for riometers using temperature limited vacuum tube diodes as noise sources, but may not be valid for riometers using variable solid state diodes. Experience with the GL riometers at Thule shows that it is advisable to check the calibration of solid state riometers with a primary noise standard before use.

A moderate gain antenna such as a five element Yagi or an array of two element Yagis is often used with riometers. The computed radiation pattern for the riometers at Thule is shown in Figure 13. The main beam of the antenna is pointed towards zenith, but the -3 dB beam width is approximately 50 degree. Thus, galactic noise arriving at zenith angles in the range from 0 degree to beyond 25 degree will contribute to the noise temperature of the antenna. Due to the extra length of the path through the ionosphere, noise arriving at angles different from zenith will suffer greater absorption than noise arriving at zenith. The absorption measured by the riometer is then larger than the absorption at zenith. Also, the noise temperature of the ionosphere itself contributes to the antenna temperature. The electron temperature of the ionosphere in the height range where the absorption takes place is 200 - 270 K (22), and the antenna temperature, T_{ant}, is determined as:

$$T_{ant} = T_{gal}A + (1-A)T_{ionosphere}$$

where A is the value of the absorption and $T_{ionosphere}$ is the temperature of the ionosphere. For small values of A the ionospheric temperature introduces insignificant errors in the measurement of absorption as the galactic noise temperature (T_{gal}) is in the range of 7000 K to 10,000 K. The ionospheric contribution gains significance as the absorption increases. Figure 14 shows the ratio between the measured absorption and the zenith absorption in dB for the Thule riometer with the ionospheric temperature as a parameter. The ratio, which is dependent on the radiation pattern of the antenna used, is approximately 1.3 at low values of absorption, consistent with the value found in the literature (12). It decreases with an increase in absorption. The slope of the curves change at approximately 5 dB zenith absorption from being dominated by the antenna pattern beam width to being dominated by the influence of the ionospheric temperature. At approximately 13 - 15 dB zenith absorption the measured absorption is equal to the zenith absorption and at higher values the zenith absorption is larger than the measured absorption. If the ionosphere is envisioned as an opaque screen, and the galactic noise as a uniform source of light, then the light pattern seen by the riometer antenna during an absorption event will grow gradually darker at the circumference than at the center of the field of view of the antenna until it obtains a uniform darkness corresponding to the radiation caused by the ionospheric temperature.

The conclusion of the above discussion is that the term 'riometer absorption' as used in the literature may be very different from 'zenith absorption' during large PCA events, and suitable corrections must be included before riometer data can be used for accurate determination of absorption levels during such events. A detailed analysis of these corrections are presented in (23). The absorption measurements presented in this paper have been corrected for the influence of antenna patterns, feeder cable losses and noise source non-linearities. An ionospheric temperature of 200 K has been assumed and the absorption values represent the zenith absorption to the accuracy of the corrections.

4. The influence of PCA's on system noise.

The noise in a meteor scatter receiving system is composed of galactic noise, receiver front end noise and man made noise and interference. For a properly designed system, man made noise and interference can be completely suppressed leaving only the receiver noise and the galactic noise. The receiver noise temperature for a properly constructed receiver should not exceed 500 K throughout the frequency range 30 - 150 MHz. The average galactic noise temperature is a function of frequency ranging from 10,000 K at 30 MHz to 250 K at 150 MHz. Also, the galactic noise contribution to a receiving system, ie. the receiving antenna temperature, is the sky noise temperature integrated over the antenna aperture. This is a function of the antenna gain and the noise temperature vs azimuth and elevation and the properties of the terrain below the antenna.

The computed elevation pattern at boresight for the horizontally polarized, five element, 45 MHz Yagi antenna used with the test-bed receiver at Thule AB is shown in Figure 15. The antenna is mounted 1.5 wavelengths above the ground to obtain optimum illumination of the common scattering volume between Sondrestrom and Thule (4). The

radiation properties of the antenna itself were computed from the dimensions and spacings of the antenna elements using the method of moments. A computer program originating from the Technical University of Denmark, specifically tailored for analysis of Yagi antennas, was used (24), but other programs such as NEC (25) could have been used as well. Next the radiation pattern of the antenna mounted above the partially reflecting ground in Thule was computed using the complex Fresnel reflection coefficients for the ground. The resulting antenna gain shows an increase of 5 dB at an elevation angle of 8 degrees relative to the free space gain of the antenna. This added gain, caused by the ground reflection, is highly advantageous for the operation of the link, as it increases the received signal power from meteor trails near the hot spots close to the midpoint of the link. The antenna gain at elevation angles exceeding approximately 25 degrees is of no consequence for the reception of meteor scatter signals from Sondrestrom, as these all arrive at lower elevation angles. However, the antenna noise temperature being the galactic noise temperature integrated over the antenna aperture, contains contributions originating from all elevation angles. During undisturbed ionospheric conditions this temperature is dominated by the galactic noise arriving at low elevation angles within the main beam of the antenna where the gain is largest. The galactic noise arriving at low elevation angles will be absorbed more than the noise arriving at high elevation angles during a PCA event, as the low elevation angle paths must traverse a much longer path through the absorbing region of the ionosphere. Consequently, the galactic noise seen by the high elevation sidelobes will contribute a larger and larger part of the antenna noise temperature as the absorption level increases. This means that the antenna temperature can not be expected to decrease much more than the zenith absorption at a given frequency.

This mechanism favors operation at the higher frequencies during absorption events as both signal and noise are absorbed less than at the lower frequencies. A large signal absorption at the lower frequencies would not be detrimental to the operation of a meteor scatter link if the noise was absorbed by the same amount as the signal, but the channel capacity at low frequencies will not only suffer from the much larger signal absorption as the signal traverses the absorbing region twice and the noise only once, but the noise originating from the high angle sidelobes of the receiving antenna will further limit the capacity. Also, the ionospheric temperature eventually limits the antenna temperature to values greater than 200 - 300 K at all frequencies. The obtainable lower limit is finally influenced by the receiver noise, but this can easily be reduced to approximately 100 K, a value which will not impair the operation of the receiving system. Thus, higher gain antennas with low sidelobes may be particularly advantageous for long meteor scatter links if resistance to absorption is important, not only due to the gain but also due to the reduced sidelobe noise contribution.

Although a receiver noise temperature of 100 K is technically feasible, it may not be possible to exploit it at the lower frequencies where absorption levels so high that the antenna temperature decreases to 200 - 400 K most certainly will be accompanied by large signal path losses, so large that the meteor signals will have all but disappeared before the antenna temperature has decreased to this point. A noise figure of 3 dB (290 K) is easily obtained with standard transistors. This noise figure will be sufficiently low for systems operating at 30 - 50 MHz and with effective radiated transmitter powers of 10 - 20 kW. At higher frequencies such as 80 - 100 MHz, a lower noise figure may be of use to make sure the receiver noise does not at any instant degrade the sensitivity of the receiving system.

5. The August 1989 PCA

A major SPE event occurred in the middle of August 1989. The associated polar cap absorption began in the afternoon of the 12th and reached a maximum 30 MHz zenith absorption of 13 dB around noon on the 13th as shown in Figure 16. The absorption then decreased throughout the 13th and 14th to reach a low of 2 dB at midnight on the 14th and then increased again to reach another maximum of 7 dB at noon on the 16th. The absorption reached a third maximum of 2.5 dB at midnight on the 19th and then decreased until noon on the 20th when the event ended.

The receiver noise absorption at the six frequencies of the Thule meteor scatter receiver during the period 10 to 23 August is presented in Figure 17. Absorption (ie. a decrease in receiver noise) is represented with positive numbers, and an increase in noise by negative numbers. The days leading up to the event show noise increases ranging from 4 to 8 dB at all frequencies. The largest increases are found at 104 MHz. Similar noise increases are found at Narssarssuaq and must be associated with the early stages of the event. The 35 MHz record shows large noise increases throughout the period of the PCA, as well as noise absorptions. These noise increases appear it Thule in the early spring and disappear in the late fall. Their first occurrence in the

spring of 1989 coincided with the sharp increase in the sunspot number. Their origin is unknown, but they are believed to be interference propagated into Thule by F2-layer reflections south of the area affected by the absorption. Such signals are known from North East Greenland where HF paths at 18 MHz or higher are regularly open to Europe during PCA's.

Solar X-ray emissions are known to create strong radio noise bursts at the beginning of a PCA and such emissions are most probably the cause of the enhanced receiver noise occurring on all frequencies between noon and midnight on the 12th. The noise absorption increases to a sharp maximum of 12 dB at 35 MHz and 8 dB, 4 dB and 2 dB at 45, 65 and 85 MHz respectively in the afternoon on the 13th. Little noise absorption is seen at 104 and 147 MHz. The maximum 30 MHz zenith absorption was 12 - 13 dB but the noise levels at the receiver frequencies did not decrease much more than the expected zenith absorption at the respective frequencies. This confirms the result of the discussion presented above that the receiver noise can not be expected to decrease much more than the zenith absorption value at any frequency even at very high absorption levels.

The meteoric arrival rate and the duty cycle for August 1989 for signals exceeding -116 dBm at Thule are shown in Figures 18 and 19. The arrival rate at 35 MHz was very intermittent due to dominating sporadic E-layer propagation at this frequency. The arrival rates for the other frequencies ranged from an average of approximately 6 meteors/min at 45 MHz to an average of 1, 0.3, 0.2 and 0.08 meteors/min at 65, 85, 104 and 147 MHz respectively. Large diurnal and day to day variations were observed during the undisturbed periods. The effect of the absorption was clearly seen. The link was interrupted at frequencies below 104 MHz during most of the 13th and decreased arrival rates were seen throughout the duration of the PCA event. The decrease was frequency dependent with the largest decreases occurring at 35, 45 and 65 MHz. Much smaller decreases were seen at the higher frequencies (85, 104 and 147 MHz). 104 MHz was not interrupted, but the arrival rate decreased slightly during the 13th and 14th. The arrival rate at 147 MHz is generally very small, and it often happens that no trails are detected within a 20 min measurement period. Tho arrival rate is therefore undefined regularly even during quiet ionospheric conditions and it was not possible to differentiate between the interruptions due to a small arrival rate and those due to absorption during the PCA. During the period from 11 to 13 August a slight increase in the arrival rate was seen twice a day on all frequencies except 35 MHz. This was due to the Perseid meteor shower that peaks in this period.

The duty cycle, which is a measure of the average communication capacity, shows much larger diurnal and day to day variations than the arrival rates, often exceeding an order of magnitude during the quiet periods. This must be kept in mind when assessing the influences of absorption on the communication capacity. In calculating the duty cycle statistics shown in Figure 19, the sporadic E-layer contribution is included and is the reason for the 100% values at 35 MHz. The contribution of the Perseid shower was more clearly observed in the duty cycle statistics than in the arrival rates, especially at 85 and 104 MHz. The duty cycle was also very frequency dependent during the PCA event. The largest decreases are seen at 35, 45 and 65 MHz, whereas smaller decreases are seen at 85 and 104 MHz. The changes observed at 147 MHz during the PCA were predominantly increases rather than decreases. They were most likely due to the Perseid shower meteors and it appears that the absorption had very little effect at 147 MHz.

The duty cycle obtained for optimum selection of the operating frequency for a SNR of 17 dB on the Sondrestrom-Thule link during the PCA event is shown in Figure 20. The frequencies are 45, 65, 85, 104 and 147 MHz. 35 MHz has been eliminated, as this frequency was dominated by sporadic E-layer propagation, and the contribution of the meteor trails was very difficult to assess. The chosen frequencies and the zenith absorption at 30 MHz are also shown. 45 MHz was the best frequency of operation whenever the zenith absorption was less than 3-5 dB and 104 MHz was chosen when the zenith absorption exceeded approximately 10 dB. 65 and 85 MHz were selected to cover the intermediate range. The highest frequency, 147 MHz, was rarely selected except during the very short times when the zenith absorption reached its maximum on the 23rd and when strong noise bursts were present at the lower frequencies. The absorption was very frequency dependent and the link was never totally interrupted, but the duty cycle dropped almost two orders of magnitude from 15% to 0.15% during the most intense parts of the absorption when frequencies lower than 104 MHz were interrupted. This change seems very large, but it mostly reflects the difference between the average duty cycle at 45 and 104 MHz during quiet ionospheric conditions. It should be noted that interruptions at the lower frequencies occurred at higher levels of absorption than the levels of absorption where the frequency changes took place, and 104 MHz was not interrupted at all. If only 45 and 104 MHz were available, the link could still be

sustained, although with a somewhat decreased capacity for zenith absorptions in the range of 5 to 8 dB.

6. The March 1989 PCA.

Two solar disturbances occurred in March 1989. The zenith absorption at 30 MHz for the period 9 to 22 March is shown in Figure 21. A polar cap absorption was observed beginning March 9 and ending March 13. The zenith absorption at 30 MHz reached values between 1 and 3 dB except for a few hours at noon on the 13th when the absorption reached 6 dB. A very large solar flare occurred on the 17th and the associated zenith absorption at 30 MHz reached 7 dB in the early afternoon of the 18th and disappeared around midnight on the 20th.

The noise absorption for the six frequencies of the Sondrestrom-Thule link are shown in Figure 22. Positive numbers denote absorption and negative numbers an increase in the receiver noise relative to the undisturbed noise level. The solar flare on the 18th produced a marked noise absorption of 6 dB at 35 MHz, 4 dB at 45 MHz and 2 dB at 65 MHz. Very little absorption is observed at the higher frequencies. The PCA that occurred between March 9 and 13 should not have caused appreciable noise absorption due to the low level of zenith absorption. This indeed is true. Only the absorption peak at noon on the 13th resulted in discernable absorption at 35, 45, 65 and 85 MHz. Very large increases of the receiver noise were observed, however, on all frequencies around noon in the period 10 through 16 March. The increases were smallest at 35 MHz, being less than 10 dB, but exceeded 15 dB at all other frequencies. The occurrence and disappearance of the noise was coincident with the local sunrise and sunset at Thule. The receiving antennas at Thule were pointed to the south, and it was very apparent that the noise was present only when the sun was in view of the antennas. A 104 MHz antenna pointing north did not receive the noise except for a short period around sunrise and sunset when the sun was within view of the antenna sidelobes.

The meteoric arrival rates and duty cycles for March 1989 are presented in Figures 23 and 24. Both the absorption and the increase in receiver noise caused the arrival rates and duty cycles to decrease. The 7 dB zenith absorption associated with the solar flare caused interruption at the 35 and 45 MHz, whereas the higher frequencies experienced smaller decreases as the frequency increased. At 65 MHz the arrival rate dropped by a factor of 15 from 1.5 meteors/min to 0.1 meteor/min and the duty cycle dropped from 0.8% to 0.03%. At 85 MHz the arrival rate and duty cycle dropped from 0.8 to 0.1 meteor/min and 0.3% to 0.03% respectively. Even less changes were seen at 104 MHz where the arrival rate dropped from 0.8 to 0.3 meteors/min and the duty cycle from 0.2 to 0.08%. The influence of the absorption was clearly frequency dependent. The influence of the solar noise emissions on the other hand decreased the arrival rates and duty cycles approximately two orders of magnitude at all frequencies. The effect was largest at 65, 85 and 104 MHz, consistent with the large solar noise power increases at these frequencies. The 30 MHz zenith absorption at the time of the solar noise bursts did not exceed 2 - 3 dB and therefore the effects of the disturbance was dominated by the solar noise emissions. This would not have been the case if the link was orientated from Thule to Sondrestrom as the receiving antennas would have been pointed north and the receiver noise would not have increased significantly due to the solar noise emissions. The disturbance would still have caused absorption on the path, but the channel capacity would only have been reduced by the absorption. The channel capacity of the two links if operated in a duplex mode would have been strongly influenced by the large reduction suffered by the south-to-north link as link management would be very difficult or impossible with one of the links having a channel capacity approximately two orders of magnitude lower than the other.

The existence of strong solar noise emissions in the VHF band has been known to geophysicists for many years, but it has received little attention in conjunction with meteor scatter propagation evaluations. The solar temperature is very high compared with the ambient galactic noise temperature even during undisturbed ionospheric conditions but the spatial extent of the sun is too small to influence the antenna temperature for the five to ten element Yagi antennas often used with a meteor scatter links. The fact that the noise emitted by the sun being a point source during the March event exceeded the undisturbed level by 10 - 25 dB indicates that the noise temperature of the sun must have increased by several orders of magnitude. The noise emissions were apparently not accompanied by a large flux of high energy protons as the ionospheric absorption did not exceed 3 dB during most of the event. The August PCA on the other hand had high levels of absorption but few solar noise emissions. Also, communication systems using FM modulations such as frequency shift keying have limiting IF amplifiers in the receivers. The demodulated noise level, from such receivers does not change with a change of the antenna temperature, so link degradations are easily ascribed to signal

loss rather than to noise increase unless separate noise measurements are performed.

The influence of the November 1960 SPE on the performance of the JANET meteor scatter system operating in Northern Canada (2) indicate that both the 45 and the 104 MHz links were interrupted extensively by the event. Riometer absorption in excess of 18 dB was reported for the early parts of the event, and this level of absorption would certainly have strongly influenced both frequencies. However, strong solar noise emissions in the VHF spectrum were also reported for this event. The SPE effects on both of the JANET links has stimulated speculation that absorption on meteor scatter systems would be frequency independent during PCA's, eliminating the advantage of working at higher frequencies. It may be, however, that the interruption of the 104 MHz link was not caused by absorption as much as the increased solar noise.

Summary

Data acquired with the 1200 km GL High Latitude Meteor Scatter Test-Bed between Sondrestrom AB and Thule AB in Greenland during solar disturbances in March and August of 1989 have been analyzed and presented. The solar disturbance of August 1989 resulted in a Polar Cap Absorption event reaching 12 dB zenith absorption at 30 MHz. The test-bed, operated at 35, 45, 65, 85, 104 and 147 MHz, was significantly influenced by the absorption. All frequencies lower than 104 MHz were interrupted for periods ranging from a few hours to one day during the part of the event where the zenith absorption exceeded 10 dB. No interruptions occurred at 104 MHz although a slight decrease of duty cycle was noted. The absorption was highly frequency dependent. Operation at frequencies in the range 65 to 104 MHz out performed 35 and 45 MHz during the most severe absorption. The solar disturbances in March 1989 produced a weak PCA with a maximum zenith absorption of 3 dB. Little effect was seen from the absorption, but very strong solar noise bursts decreased the signal-to-noise ratio on all frequencies and caused interruptions. It is speculated that solar noise bursts present during previous Solar Proton Events could have been the basis for expectations that the absorption at VHF frequencies during these events was frequency independent. That is not supported by the results presented in this paper.

References

(1) Crysdale, J.H., Analysis of the performance of the Edmonton-Yellowknife JANET circuit, IRE Trans. Com., March 1960.
(2) Maynard, L.A., Propagation of meteor burst signals during the polar disturbance of November 12-16 1960. Can. J. Phys. V39, p628, 1961.
(3) Ostergaard, J.C., J.E. Rasmussen, M.S. Sowa, J.M. Quinn and P.A. Kossey, Characteristics of high latitude meteor scatter propagation over the 45 to 104 MHz band, AGARD Conf. Proc., AGARD-CP-382, paper 9.2, 1985.
(4) Ostergaard, J.C., J.E. Rasmussen, M.S. Sowa, J.M. Quinn and P.A. Kossey, The RADC High Latitude Meteor Scatter Test-Bed. Tech. Rep. RADC-TR-86-74, Rome Air Development Center, July 1986.
(5) Weitzen, J.A. and S. Tolman, A technique for automatic classification of meteor trails and other propagation mechanisms for the Air Force high latitude meteor burst test bed, Tech. Rep. RADC-TR-86-117, Rome Air Development Center, AD173133, 1986.
(6) Weitzen, J.A., A data base approach to analysis of meteor burst data, Radio Science, Vol. 22, No. 1, pp 133-140, January-February 1987.
(7) Weitzen, J.A., USAF/GL Meteor scatter data analysis program. Users guide. GL-TR-89-0154, june 1989.
(8) Weitzen, J.A., M.J. Sowa, R.A. Scofidio, J. Quinn, Characterizing the multipath and doppler spreads of the high latitude meteor burst channel, IEEE Trans. Com. V.COM-35, no. 10, Oct. 1987.
(9) Ostergaard J.C., Meteor Burst propagation and system design. Special course on interaction of propagation and digital transmission techniques. AGARD R-744, 1987.
(10) Sowa, M.J. Quinn, J.M., Rasmussen, J.E., Kossey, P.A., Ostergaard, J.C., A statistical analysis of polar meteor scatter propagation in the 45 - 104 MHz Band. AGARD Conf. Proc., AGARD-CP-419, paper 44, 1987.
(11) Eshlemann, V.R., On the wavelength dependence of the information capacity of meteor burst propagation. Proc. IRE, Dec 1957.
(12) Manual of ionospheric absorption measurements. K. Rawer ed. World Data Centre A Report UAG-57. 1976.
(13) Sen, H.K., Wyller, A.A., On the generalization of the Appleton-Hartree Magnetoionic formulas. J. Geophys. Res. Vol 65, p 3931, 1960.
(14) Dingle, R.B., D. Arndt, S.K. Roy, The integral Cp(x) and Dp(x) and their tabulation. Appl. Sciences Res. 6B p155, 1957.
(15) Thrane, E.V., W.R. Piggott, The collision frequency in the E- and D-regions of the ionosphere. J.A.T.P. V30, p135, 1966.

(16) Jones, R.M. A versatile three dimensional raytracing computer program for radio waves in the ionosphere. U.S. Dept. Commerce, 1975.
(17) Kossey, P.A., J.P. Turtle, R.P. Pagliarulo, W.I. Klemetti, and J.E. Rasmussen, VLF reflection properties of the normal and disturbed polar ionosphere in northern Greenland, Radio Sci, Vol. 18, No. 6, Nov-Dec 1983, p. 907-916.
(18) Friedrich, M., K.M. Torkar, Collision frequencies in the high latitude D-region. J.A.T.P. V45, p267, 1983.
(19) Lerfald, C.M., C.G. Little, D-region electron density profiles during Auroras. J. Geophys. Res. V69 no.13, July 1964.
(20) Cormier, R.J., Riometry as an Aid to Ionospheric Forecasting, AFCRL TR-70-0689. Dec. 1970.
(21) Friedrich, M., K.M. Torkar, High latitude plasma densities and their relation to riometer absorption. J.A.T.P. V45, p217, 1983.
(22) Cospar International Reference Atmosphere 1972, Akademie-Verlag, Berlin 1972.
(23) Ostergaard, J.C., Derivation of zenith absorption from riometer measurements with special emphasis on the GL riometers at Thule AB., in publication.
(24) Hald, J., J.E. Hansen, Dipole arrays with loading networks. Electromagnetics Institute Rep. R257, R258, Technical University of Denmark, 1982.
(25) NEEDS, The Numerical Electromagnetic Engineering Design System. Version 1.0 February 1988. Dist. by the Applied Computational Electromagnetics Society.

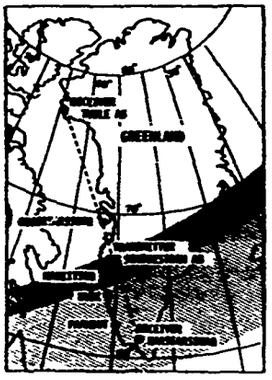


Fig. 1. Geographical locations of the GL High Latitude Meteor Scatter Test-Bed.

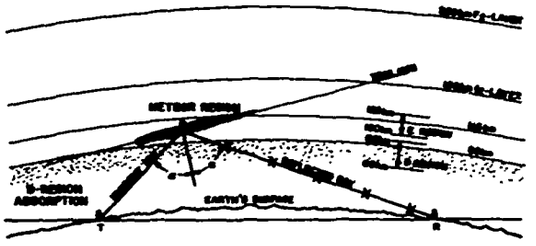


Fig. 2. Cross section of a meteor scatter propagation path.

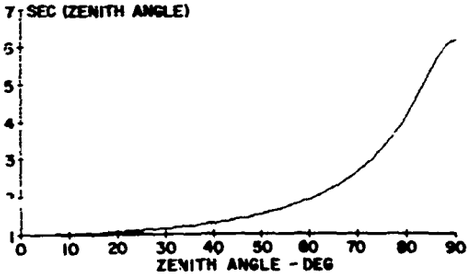


Fig. 3. Ratio of path length at zenith vs path lengths at zenith angles in the range from 0 to 90 deg through the ionosphere's D-region.

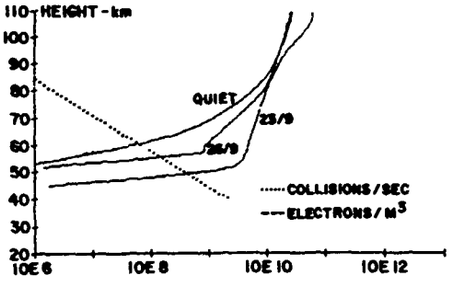


Fig. 4. Electron density profiles at noon on 23 and 26 Sept. 1978 at Thule.

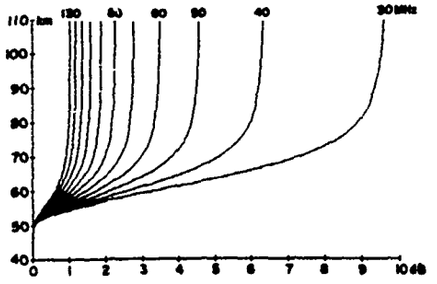


Fig. 5. Computed cumulative zenith absorption at noon 23 Sept. 1978.

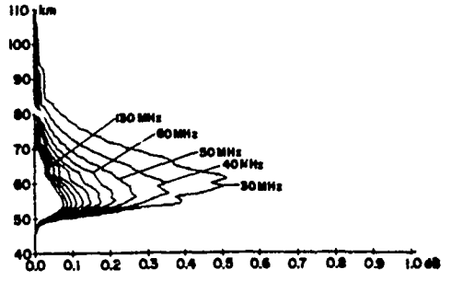


Fig. 6. Computed specific zenith absorption at noon 23 Sept. 1978.

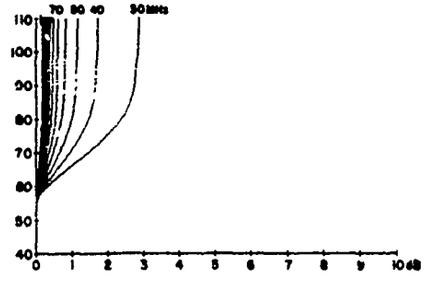


Fig. 7. Computed cumulative zenith absorption at noon 26 Sept. 1978.

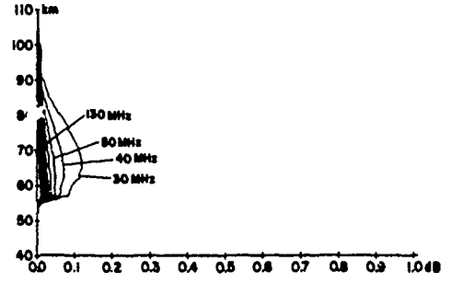


Fig. 8. Computed specific zenith absorption at noon 26 Sept. 1978.

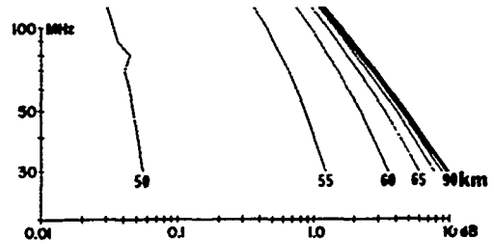


Fig. 9. Computed frequency dependence of the cumulative absorption at noon 23 Sept. 1978.

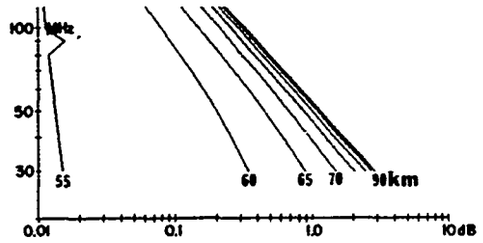


Fig. 10. Computed frequency dependence of the cumulative absorption at noon 26 Sept. 1978.

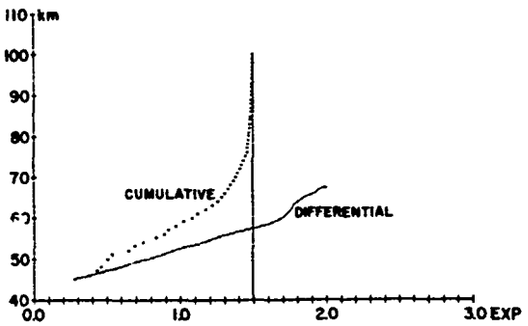


Fig. 11. Computed exponent of frequency dependence of absorption vs. height at noon 23 Sept 1978.

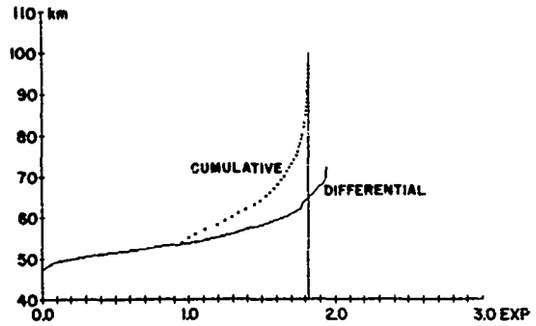


Fig. 12. Computed exponent of frequency dependence of absorption vs. height at noon 26 Sept 1978.

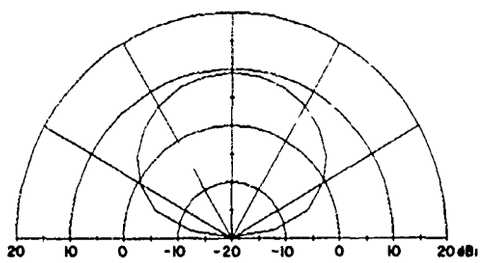


Fig. 13. Computed radiation pattern for the riometer at Thule AB.

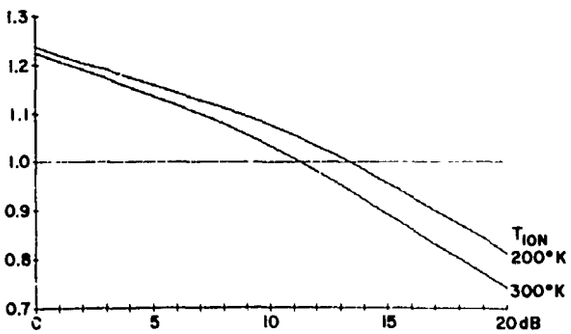


Fig. 14. Ratio of measured absorption and zenith absorption in dB at Thule.

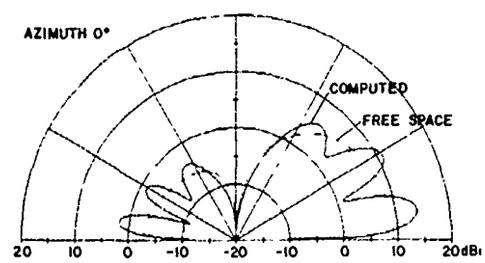


Fig. 15. Computed radiation pattern for the 45 MHz meteor scatter antenna at Thule AB.

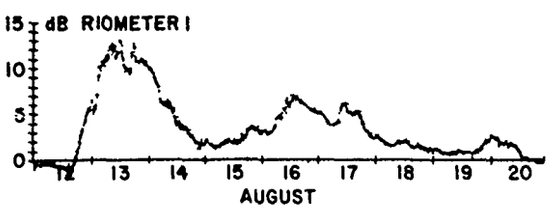


Fig.16. Zenith absorption measured at Thule AB August 10 to 22 1989.

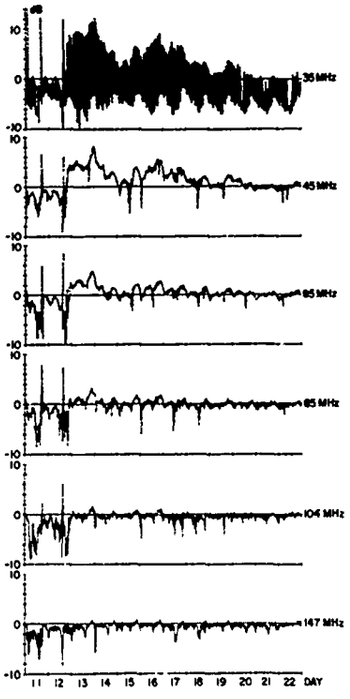


Fig. 17. Measured noise absorption at 35, 45, 65, 85, 104 and 147 MHz Thule AB August 1989.

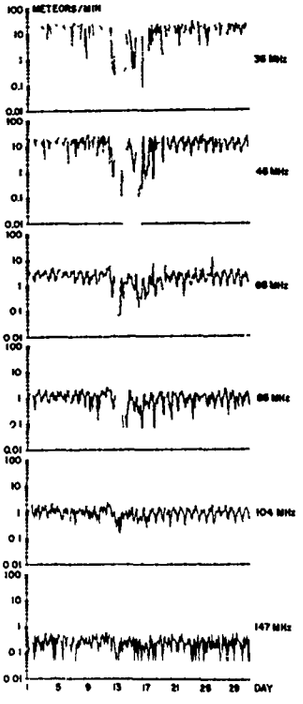


Fig. 18. Meteor trail arrival rate exceeding -116 dBm signal level at Thule AB August 1989.

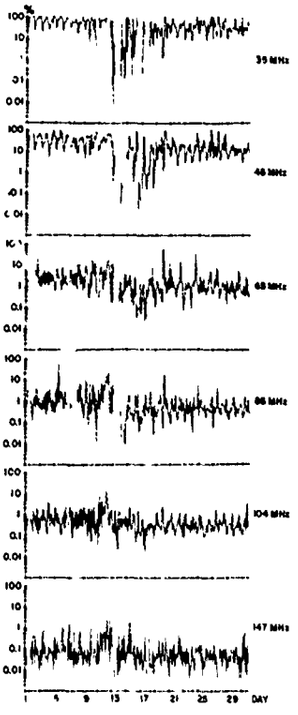


Fig. 19. Duty cycle exceeding -116 dBm at 35, 45, 65, 85,104 and 147 MHz at Thule August 1989.

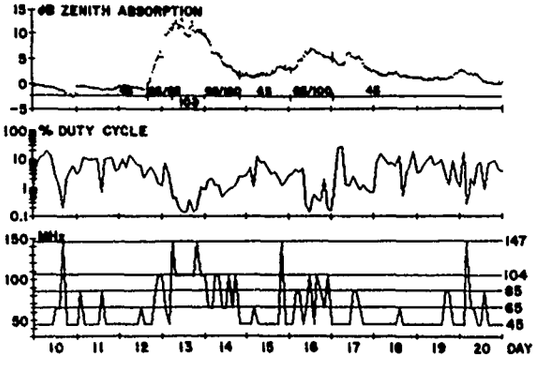


Fig. 20. Duty cycle and frequency of operation for optimum selection of frequency compared to zenith absorption August 10-20 1989.

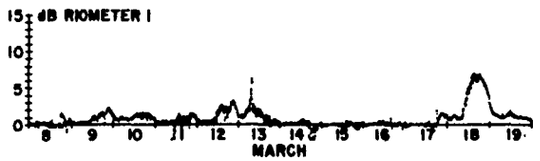


Fig. 21. Zenith absorption measured at Thule March 1989.

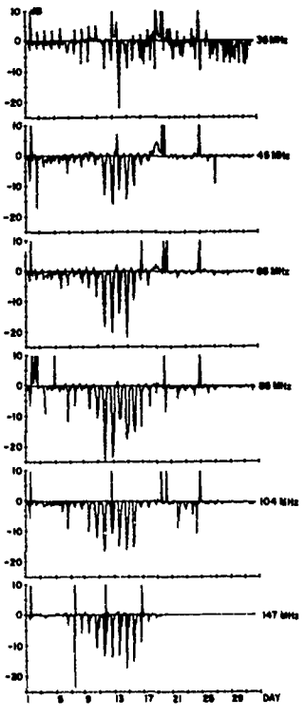


Fig. 22. Measured noise absorption at 35, 45, 65, 85, 104 and 147 MHz Thule March 1989.

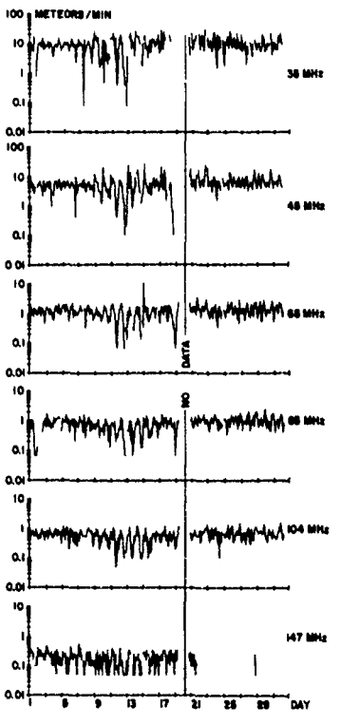


Fig. 23. Meteor trail arrival rate exceeding -116 dBm signal level at 35, 45, 65, 85, 104 and 147 MHz Thule AB March 1989. (27 dB SNR)

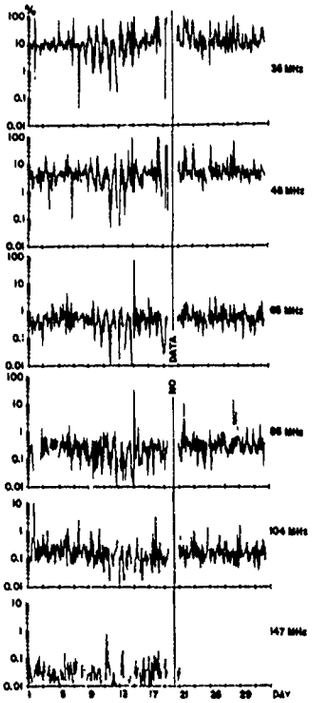


Fig. 24. Duty cycle exceeding -116 dBm signal level at 35, 45, 65, 85, 104 and 147 MHz Thule AB March 1989. (27 dB SNR)

AD-P006 301

EVALUATION OF SIX IONOSPHERIC MODELS AS PREDICTORS OF TEC

Lincoln D. Brown and Robert E. Daniell, Jr.
Computational Physics, Inc.
385 Elliot Street
Newton, MA 02072

Matthew W. Fox
Center for Space Physics
Boston University
725 Commonwealth Ave.
Boston, MA 02215

John A. Klobuchar
Ionospheric Physics Division
Geophysics Laboratory (AFSC)
Hanscom AFB, MA 01731

Patricia H. Doherty
Institute for Space Research
Boston College
Chestnut Hill, MA 02167

91-09701

ABSTRACT

We have gathered TEC data from a wide range of latitudes and longitudes for a complete range of solar activity. This data was used to evaluate the performance of six ionospheric models as predictors of Total Electron Content (TEC). The TEC parameter is important in correcting modern DoD space systems, which propagate radio from the earth to satellites, for the time delay effects of the ionosphere. The TEC data were obtained from polarimeter receivers located in North America, the Pacific, and the East coast of Asia. The ionospheric models evaluated are (1) the International Reference Ionosphere (IRI), (2) the Bent model, (3) the Ionospheric Conductivity and Electron Density (ICED) model, (4) the Penn State model, (5) the Fully Analytic Ionospheric Model (FAIM, a modification of the Chiu model), and (6) the Damen-Hartranft model. We will present extensive comparisons between monthly mean TEC at all local times and model TEC obtained by integrating electron density profiles produced by the six models. These comparisons demonstrate that even though most of the models do very well at representing f_0F_2, none of them do very well with TEC, probably because of inaccurate representation of the topside scale height. We suggest that one approach to obtaining better representations of TEC is the use of f_0F_2 from the CCIR coefficients coupled with a new slab thickness model developed at Boston University.

INTRODUCTION

In order to evaluate the usefulness of publicly available ionospheric models as midlatitude Total Electron Content (TEC) models, we have gathered together an extensive set of TEC data from polarimeter measurements at several locations around the world. We have used this data to evaluate the performance of six ionospheric models. These models are

(1) BENT: The Bent Model [Llewellyn and Bent, 1973; Bent et al., 1976] developed for transionospheric propagation uses

(2) FAIM: The Fully Analytic Ionospheric Model [Anderson et al., 1989], which is a modification of the Chiu model [Chiu, 1975]

(3) ICED: The Ionospheric Conductivity and Electron Density Model [Tascione et al., 1988] developed for the Air Force Air Weather Service

91

(4) IRI: The International Reference Ionosphere [Rawer, 1981]

(5) PSU: The Pennsylvania State University Mark III Model [Nisbet and Divany, 1987]

(6) URSI/DHR: The Damen-Hartranft-Ramsay model [Damon and Hartranft, 1970; Flattery and Ramsay, 1976], using the URSI-88 coefficients to calculate f_oF_2 and h_mF_2, and with a modified F_1 layer.

For each model, we obtained the FORTRAN source code, standardized it to FORTRAN 77 when necessary, and compiled and executed it on a DEC MicroVAX computer. In some cases, we found it necessary to modify the codes to facilitate multiple runs and slant TEC calculations. Only in the URSI/DHR case did we modify any of the subprograms that performed the actual ionospheric calculations. In that case, we found that the F_1 layer was too large. For the purposes of this comparison, we simply eliminated it. In the future, we hope to investigate alternate F_1 models that may give better profile shapes as well as improved TEC estimates. In the case of FAIM, we note that the original Chiu model on which it is based was never intended to compete with statistical models for prediction purposes. We include it here simply because it is widely used.

All of the TEC data were obtained from polarimeter measurements at eight middle and low latitude stations. Slant TEC was converted to vertical equivalent TEC by multiplying the slant TEC by the secant of the zenith angle of the signal path at the ionospheric intersection. This approximation to true vertical TEC is valid in midlatitude regions, where there is little horizontal structure of the ionosphere. Because we were interested primarily in the ability of the models to correctly predict monthly median TEC, we reduced the data to monthly median values at each hour of local time. We also calculated the monthly standard deviations for each hour. Table 1 lists the stations, the years, and the subionospheric coordinates of the data used in these comparisons.

Because we wanted to know how much of the error in calculated TEC was due to inaccurate calculation of f_oF_2 and related parameters, we also obtained ionosonde data from Wallops Island (38.71°N,289.18°E), which is near the ionospheric intersection point for the Hamilton, MA measurements.

Table 1. Polarimeter locations and dates

Location	Year	Solar Activity	Ionospheric Intersection Latitude	Longitude
Goose Bay, Lab.	1981	maximum	47.00 °N	298.00 °E
	1986	minimum	47.00	286.00
Hamilton, MA	1975	minimum	42.63	289.18
	1981	maximum	42.63	289.18
Kennedy AFB, FL	1981	maximum	26.00	280.00
	1986	minimum	26.00	276.00
Osan, Korea	1980	maximum	34.00	127.00
	1986	minimum	34.00	127.00
Palehua, HI	1981	maximum	20.00	203.00
	1986	minimum	20.00	204.00
Ramey, PR	1981	maximum	17.00	293.00
	1986	minimum	17.00	288.00
Shemya, AK	1981	maximum	47.00	164.00
	1986	minimum	46.00	186.00
Taiwan	1980	maximum	23.00	122.00
	1986	minimum	23.00	122.00

MODEL EVALUATIONS

Each model has its own set of input parameters, usually different from the other models. Besides location, date, and local time, each model depends on some index of solar activity, usually sunspot number or $F_{10.7}$. Some models also require a geomagnetic activity index, but we ignored

this parameter since we dealt exclusively with midlatitude data. In place of daily values (whether unsmoothed, smoothed, or running mean), we used monthly medians to facilitate comparison with monthly median data. After verifying for each model that the difference between true and equivalent vertical TEC was negligible for each of the polarimeter stations, we compared true vertical TEC from the models (TEC_{model}) with the equivalent vertical TEC from the data (TEC_{data}). Because we have ionosonde measurements near the ionospheric pierce point of the Hamilton, MA polarimeter measurements, we will concentrate on model comparisons with that data. We will then summarize the model comparisons with the other stations.

The quantity $TEC_{model} - TEC_{data}$ for Hamilton is shown for each hour of local time for January and June 1975 in Figure 1. As an indication of the natural variability of the ionosphere, the standard deviations of the data are also shown. Note that every model misses the data by more than one standard deviation at several local times.

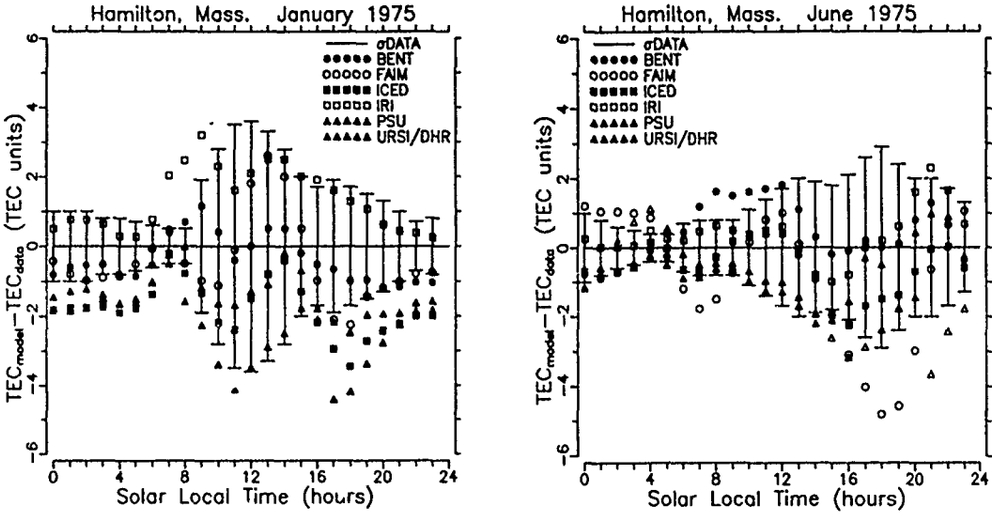


Figure 1. Comparisons of calculated TEC vs. actual median values for each local time during January (left) and June (right) 1975 (solar minimum). The polarimeter was located at Hamilton, MA and the ionospheric pierce point was in the vicinity of Wallops Island, VA. The error bars represent one standard deviation of the data and are used to indicate the relative variability of the ionosphere.

Similar comparisons for January and June 1981, a solar maximum year, are shown in Figure 2. (Note the change in scale.) Not surprisingly, model accuracy seems to be worse than for the solar minimum year. In particular, IRI is consistently high, sometimes by many standard deviations. Figure 3 further illustrates the decrease in performance at high solar activity by displaying the deviation between calculated and observed noon TEC values.

Since we had f_oF_2 data near the ionospheric intersection for the Hamilton data, we decided to investigate the relative importance of f_oF_2 and profile shape in producing the model errors in TEC. The comparisons of model and data f_oF_2 median values for June 1975 and June 1981 are shown in Figure 4. The comparison of model and data noon f_oF_2 monthly median values throughout 1981 are shown in Figure 5. Although the models tend to be low during June of both years, they generally do better at predicting f_oF_2 than TEC. The exception is PSU for which a major source of error in TEC appears to be error in f_oF_2.

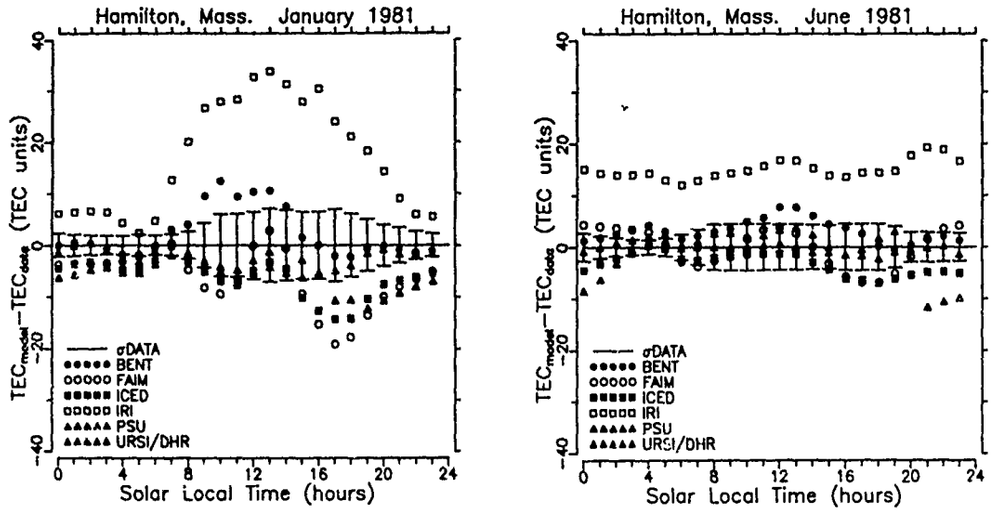


Figure 2. Comparison of calculated and actual TEC values during January (left) and June (right) 1981, a solar maximum year. Note the large change in scale compared to Figure 1.

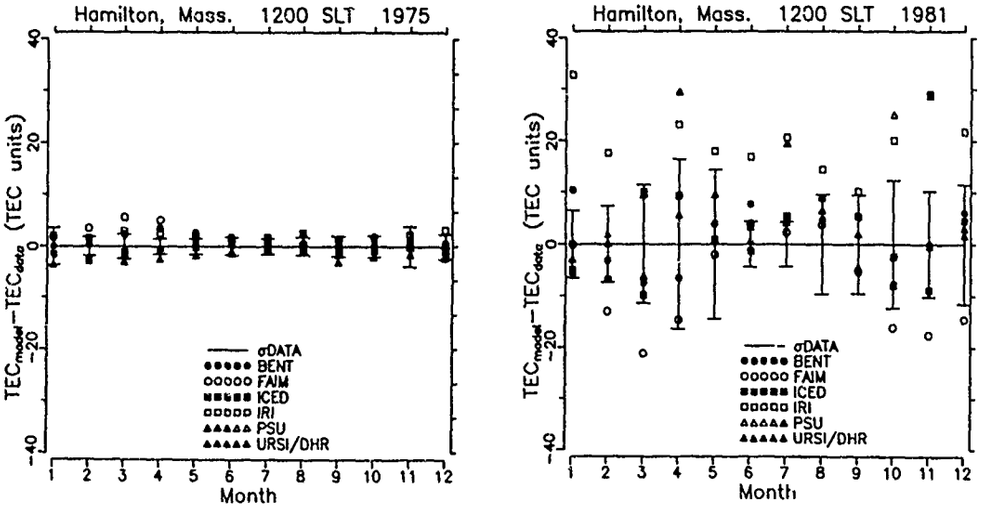


Figure 3. Comparisons of calculated and observed noon TEC monthly median values during a solar minimum year (1975, left) and a solar maximum year (1981, right). Note that the August 1981 PSU value is off the scale: nearly 80 TEC units above the actual value.

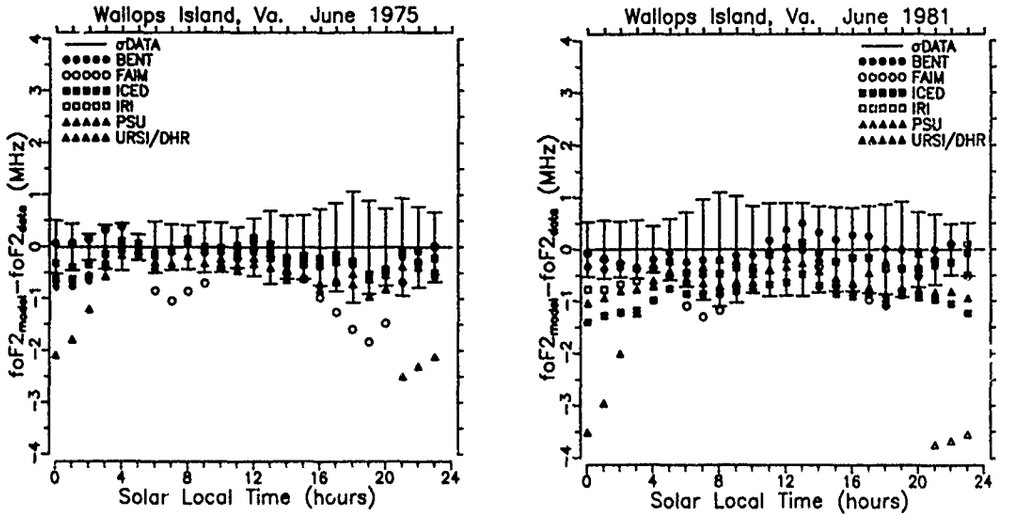


Figure 4. Comparisons of calculated and observed f_0F_2 median values during June 1975 (solar minimum) and 1981 (solar maximum) at Wallops Island, VA. Although most of the models do better predicting f_0F_2 than TEC, there is still considerable error, especially at night.

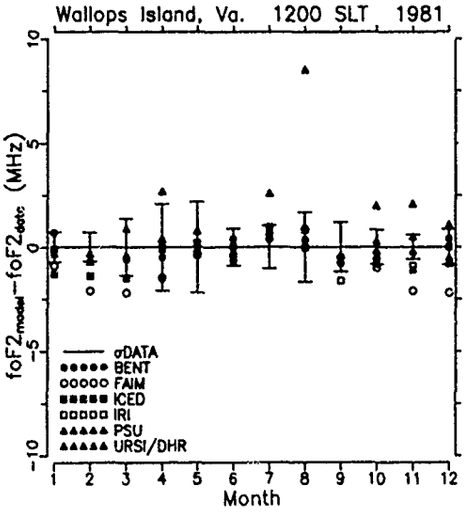


Figure 5. Comparisons of calculated and actual noon f_0F_2 monthly median values at Wallops Island during 1981.

We have carried out similar comparisons for the other seven polarimeter stations (except that we had no ionosonde data for these stations). In order to summarize these comparisons, we have defined two "summary" parameters that characterize the performance of a model over solar local time for a given month, data location, and year:

1. R.M.S. deviation of model from data

$$\sqrt{\sum_{i=0}^{23} \left[TEC_i^{model} - TEC_i^{data} \right]^2}$$

2. R.M.S. normalized deviation of model from data

$$\sqrt{\sum_{i=0}^{23} \left[\frac{TEC_i^{model} - TEC_i^{data}}{\sigma_i^{data}} \right]^2}$$

The index i denotes solar local time, and σ^{data} is the standard deviation of the data, a measure of the natural variation of TEC. To further reduce the volume of data, we averaged each parameter over a year to produce a mean "figure of merit" for each model for that year. These figures of merit are summarized in Tables 2 and 3 for solar minimum and Tables 3 and 4 for solar maximum.

Table 2. Yearly mean RMS TEC Deviations for Solar Minimum (TEC Units)

Station	Year	BENT	FAIM	ICED	IRI	PSU	URSI/DHR
Goosebay	1986	1.40	2.28	1.18	1.59	*	0.71
Hamilton	1975	1.18	1.92	1.22	1.48	1.49	1.30
Kennedy	1986	2.07	2.35	2.22	2.87	*	1.85
Osan	1986	2.14	3.08	2.66	6.45	1.60	3.15
Palehua	1986	3.82	5.73	3.85	10.71	*	4.05
Ramey	1986	2.81	3.28	2.81	5.72	*	3.01
Shemya	1986	1.12	2.36	1.56	1.63	1.61	1.82
Taiwan	N/A						
mean (all stations)		2.08	3.00	2.21	4.35		2.27
mean (midlatitude only)		1.58	2.40	1.77	2.80		1.77

Table 3. Yearly mean RMS Normalized TEC Deviations for Solar Minimum (Standard Deviations)

Station	Year	BENT	FAIM	ICED	IRI	PSU	URSI/DHR
Goosebay	1986	1.76	3.07	1.36	2.20	*	0.71
Hamilton	1975	1.10	1.50	1.06	1.26	1.29	1.30
Kennedy	1986	2.03	2.27	2.01	1.69	*	1.85
Osan	1986	1.53	2.32	1.86	3.55	0.87	3.15
Palehua	1986	0.92	1.95	1.81	4.62	*	4.05
Ramey	1986	1.56	2.39	1.68	3.04	*	3.01
Shemya	1986	0.71	1.42	0.97	0.98	1.00	1.82
Taiwan	N/A						
mean (all stations)		1.37	2.13	1.54	2.48		2.27
mean (midlatitude only)		1.43	2.12	1.45	1.94		1.77

*The PSU model failed under certain low sunspot number conditions because of a problem in its interpolation on the CCIR coefficients.

Table 4. Yearly mean RMS TEC Deviations for Solar Maximum (TEC Units)

Station	Year	BENT	FAIM	ICED	IRI	PSU	URSI/DHR
Goosebay	1981	5.40	10.98	5.97	26.33	15.17	4.95
Hamilton	1981	5.61	11.86	5.44	14.92	15.44	3.42
Kennedy	1981	7.92	14.84	5.99	7.38	17.88	7.04
Osan	1980	5.76	13.30	9.06	0.32	14.07	14.43
Palehua	1981	16.51	26.27	16.27	15.33	23.05	12.80
Ramey	1981	11.30	17.17	6.73	7.41	20.38	7.72
Shemya	1981	5.41	7.90	7.53	12.00	18.80	11.64
Taiwan	1980	15.55	32.46	18.95	13.75	24.45	16.57
mean (all stations)		9.18	16.85	9.49	13.43	18.66	9.82
mean (midlatitude only)		6.02	11.78	6.80	14.19	16.27	8.30

Table 5. Yearly mean RMS Normalized TEC Deviations for Solar Maximum (Standard Deviations)

Station	Year	BENT	FAIM	ICED	IRI	PSU	URSI/DHR
Goosebay	1981	0.77	1.42	0.87	4.20	2.67	0.74
Hamilton	1981	0.96	1.82	0.94	3.12	2.62	0.70
Kennedy	1981	1.18	2.12	0.89	1.33	2.29	1.15
Osan	1980	0.92	2.48	1.53	1.83	2.21	2.25
Palehua	1981	0.98	1.96	1.02	1.34	1.50	1.20
Ramey	1981	1.63	2.33	0.96	1.21	2.47	1.24
Shemya	1981	1.09	1.40	1.56	2.42	3.14	2.42
Taiwan	1980	1.21	2.71	1.91	1.58	2.25	1.58
mean (all stations)		1.09	2.03	1.21	2.13	2.39	1.41
mean (midlatitude only)		0.98	1.85	1.16	2.58	2.59	1.45

In terms of overall performance, BENT, ICED, and URSI/DHR perform about equally well, with BENT doing slightly better than the other two. The remaining models perform very poorly overall. We consider model predictions within one standard deviation of the monthly mean or median value minimally acceptable. None of the models we examined can consistently meet this standard. Even the best one (BENT) barely meets it in the mean — meaning that it is worse than one standard deviation about as often as it is better. We feel that a good statistically based model ought to do better than that.

Tables 6 and 7 show the f_0F_2 performance of each model in comparison with Wallops Island data. At least for this station, most of the models do considerably better at predicting this parameter than they do at predicting TEC. Roughly half the error in TEC appears to be due to errors in f_0F_2, so the remaining error must be due to errors in profile shapes. We suspect that the main problem is inaccurate topside profiles. Preliminary investigations suggest that statistical slab thickness models coupled with ITS-78 or CCIR coefficients can improve TEC predictions. However, for applications requiring the complete profile, improved profiling techniques are needed. Also, we think that regional f_0F_2 coefficient sets and slab thickness models may be helpful where extensive historical data sets are available.

Table 6. YEARLY MEAN R.M.S. FOF2 DEVIATIONS (MHz)
Wallops Island (38°71N,289°18E)

Model	Solar Minimum (1975)	Solar Maximum (1981)
BENT	0.48	0.55
FAIM	0.88	1.88
ICED	0.46	0.80
IRI	0.44	0.85
PSU	0.99	2.32
URSI/DHR	0.59	0.55

Table 7. YEARLY MEAN R.M.S. NORMALIZED FOF2 DEVIATIONS
Wallops Island (38.71N,289.18E)

Model	Solar Minimum (1975)	Solar Maximum (1981)
BENT	0.83	0.70
FAIM	1.47	2.27
ICED	0.81	0.97
IRI	0.73	1.11
PSU	1.75	2.79
URSI/DHR	0.97	0.66

REFERENCES

Anderson, D. N., J. M. Forbes, and M. Codrescu, A fully analytic, low- and middle-latitude ionospheric model, J. Geophys. Res., 94, 1520-1524, 1989.

Bent, R. B., S. K. Llewellyn, G. Nesterczuk, and P. E. Schmid, The Development of a highly successful worldwide empirical ionospheric model and its use in certain aspects of space communications and world-wide total electron content investigations, in Effect of the Ionosphere on Space Systems and Communications, ed. by J. Goodman, NTIS, Springfield, VA, pp. 13-28, 1976.

Chiu, Y. T., An improved phenomenological model of ionospheric density, J. Atmos. Terr. Phys., 37, 1563-1570, 1975.

Damon, T. D., and F. R. Hartranft, Ionospheric electron density profile model, Tech. Memo 70-3, Tech. Support Center, Aerosp. Environ. Support Cent., Scott AFB, IL, 1970.

Flattery, T. W., and A. C. Ramsay, Derivation of Total Electron Content for Real-Time Applications, in Effect of the Ionosphere on Space Systems and Communications, ed. by J. Goodman, NTIS, Springfield, VA 336-344, 1976.

Llewellyn, S. K., and R. Bent, Documentation and Description of BENT Ionospheric Model, AFCRL-TR-73-0657, AD772733, Geophysics Laboratory, Hanscom AFB, MA, 1973.

Nisbet, J. S., and R. Divany, Instructions for running the PC version of the Penn State Mark III Ionospheric Model, Scientific Report CSSL SCI 484, The Pennsylvania State University, University Park, PA, February 1987.

Rawer, K., International Reference Ionosphere - IRI 79, Report UAG-82, ed. by J. V. Lincoln and R. Conkright, World Data Center A, Boulder, CO, 1981.

Tascione, T. F., H. W. Kroehl, R. Creiger, J. W. Freeman, Jr., R. A. Wolf, R. W. Spiro, R. V. Hilmer, J. W. Shade, and B. A. Hausman, New ionospheric and magnetospheric specification models, Radio Sci., 23, 211-222, 1988.

AD-P006 302

A GLOBAL IONOSPHERIC CONDUCTIVITY AND ELECTRON DENSITY (ICED) MODEL

91-09700

Robert E. Daniell, Jr.
Computational Physics, Inc.
385 Elliot Street
Newton, MA 02072

Dwight T. Decker
Institute for Space Research
Boston College
Chestnut Hill, MA 02167

David N. Anderson and John R. Jasperse
Ionospheric Physics Division
Geophysics Laboratory
Hanscom AFB, MA 01731

Jan J. Sojka and R. W. Schunk
Center for Atmospheric and Space Sciences
Utah State University
Logan, UT 84322

ABSTRACT

At the Fifth Ionospheric Effects Symposium (IES '87), a paper was presented and published that described an operational Air Force Air Weather Service (AWS), near real-time, ionospheric specification model called Ionospheric Conductivity and Electron Density (ICED) . That version of the model specified primarily the northern hemisphere midlatitude and auroral zone ionospheric regions and did not cover either the low latitude or high latitude/polar cap regions. We are making extensions and modifications to ICED that provide a global specification — in effect a GLOBAL ICED. Major additions include (1) specification of electron and ion density profiles (90 to 1000 km) in the high latitude trough, auroral oval, and polar cap ionosphere in both northern and southern hemispheres, (2) midlatitude electron and ion density profiles that include magnetic declination effects, and (3) low latitude ionospheric specification that accounts for longitude control through differences in vertical E X B drift patterns. Since the model is driven by real-time data, we describe the various input parameters and how they are incorporated.

INTRODUCTION

The Ionospheric Conductivity and Electron Density (ICED) described by Tascione et al. [1988] provides near real-time specification of the northern midlatitude and auroral ionosphere (20^{o}N - 80^{o}N corrected geomagnetic latitude) based on climatological models. The model is driven by two geophysical parameters: SSN_{eff}, an effective sunspot number (or ionospheric index) determined from the Air Weather Service real-time ionosonde network observations and Q_{eff}, an auroral activity index determined from satellite imagery. ICED was modified to include the low latitude ionosphere (20^{o}S - 20^{o}N) by including the Fully Analytic Ionospheric Model (FAIM) of Anderson et al. [1989], a parameterized version of the Semi-Empirical Low-latitude Ionospheric Model (SLIM) of Anderson et al. [1987] combined with a modified version of the Chiu [1975] model. However, this model is only accurate for the American longitudes where magnetic declination is relatively small.

ICED is now being modified to provide a truly global ionospheric specification (90^{o}S - 90^{o}N) driven by a variety of near real-time data including digital ionosonde data, total electron content data, in situ plasma density, temperature, and composition measurements and satellite based ultraviolet airglow and auroral emission measurements. As part of this modification effort, ICED is being changed from a primarily climatological model to a more physically based model GLOBAL ICED is divided into five regions: Northern High Latitude, Northern Midlatitude, Low Latitude, Southern

Midlatitude, and Southern High Latitude. The boundaries between these regions are dynamic in the sense that they are adjusted according to current conditions. For example, the boundary between the high latitude and midlatitude regions is the equatorward edge of the subauroral trough, which expands and contracts with magnetic activity. The northern high latitude region model is in the testing and validation phase, while the other regions are still in the development phase. The following sections describe the two high latitude models, the two midlatitude models, and the low latitude model.

In all regions, we have based the new models on parameterizations of physically based "first principles" ionospheric models rather than on statistical or climatological models. Although existing physical models are still inadequate for accurately specifying the state of the ionosphere based solely on a few solar and geophysical parameters and require too many computer resources to be used in real time, we believe that they provide more realistic representations of actual ionospheric structures than can be obtained from statistical models that, by their very nature, tend to wash out both temporal and spatial structure. In addition, the physical models give ion composition information unavailable from models based on ionosonde measurements alone. The challenge is to find representations of these physical models that can be driven by real time data using available computer resources.

PARAMETERIZING THE PHYSICAL MODELS

For this version of GLOBAL ICED we have chosen to parameterize the physical models in terms of geophysical environmental parameters by producing databases of ion density profiles on a latitude, longitude (or local time), and universal time grid for a range of environmental parameters. The precise choice of environmental parameters and their values depends on the ionospheric region. Seasonal variations are most important for the middle and high latitudes. Solar activity is an important parameter at all latitudes. Geomagnetic activity is mainly important at high latitudes, while the wind-driven E- and F-region dynamos are important for the formation of the equatorial anomaly. Since the physical models make intensive use of computer resources, we generally produce databases for high, moderate, and low values of each parameter.

Since each database for a given set of geophysical parameters must represent the ionosphere throughout the applicable region during a 24 hour period in universal time, they tend to be unwieldy. Therefore, the next step is the production of a semi-analytic representation based on discrete orthonormal functions. (Discrete orthogonal functions have their orthogonality properties defined by sums on discrete grids rather than by integrals over a continuous interval. The functions themselves may or may not be defined between the grid points.)

The first stage in the process is the approximation of altitude profiles of ion density using Empirical Orthonormal Functions (EOF's). The generation and use of EOF's is described in Secan and Tascione [1984] or Kutzbach [1967] and references therein. In short, the EOF's are the eigenvectors of the covariance matrices of the ion density profiles. If the database consists of N altitude profiles, $f(z_i)$, (i=1,2,...M), the elements of the covariance matrix are

$$C_{ij} = \frac{1}{N} \sum_{n=1}^{N} f_n(z_j) f_n(z_j), \qquad i,j = 1,2,...M \tag{1}$$

For the high latitude databases, M = 37 and N = 5760 (20 latitudes, 24 local times, and 12 UT's).

The eigenvalues (λ_m, m=1,2,...M) of a covariance matrix are real and non-negative, and the eigenvectors, $g_m(z_k)$, are automatically orthogonal and easily normalized. (See, e.g., Hildebrand [1965].) We order the eigenvalues so that $\lambda_m \geq \lambda_{m+1}$. An exact representation of the n^{th} altitude profile, $f_n(z)$, is given by

$$f_n(z_m) = \sum_{k=1}^{M} a_{nk} g_k(z_m), \qquad m = 1,2,...M, \quad 1 \leq n \leq N \tag{2}$$

where

$$a_{nk} = \sum_{m=1}^{M} f_n(z_m) g_k(z_m) \tag{3}$$

It can be shown that the RMS error due to truncating the series (2) at K terms is minimized when the g_k are the eigenvectors of the covariance matrix — rather than any other set of discretely

orthogonal functions (Secan and Tascione [1984] and references therein).

For the high latitude region, we have found it best to generate a separate set of EOF's for each database rather than one set for the entire group of databases. We have also found that the first six (out of 37) EOF's produce good approximations and that further improvement requires many more EOF's. Since the high latitude ionosphere has so much spatial and temporal structure, we expect that the other regions will require no more, and quite probably fewer, EOF's per database. Examples of EOF's for O^+, NO^+, and O_2^+ from one of the high latitude databases are shown in Figure 1.

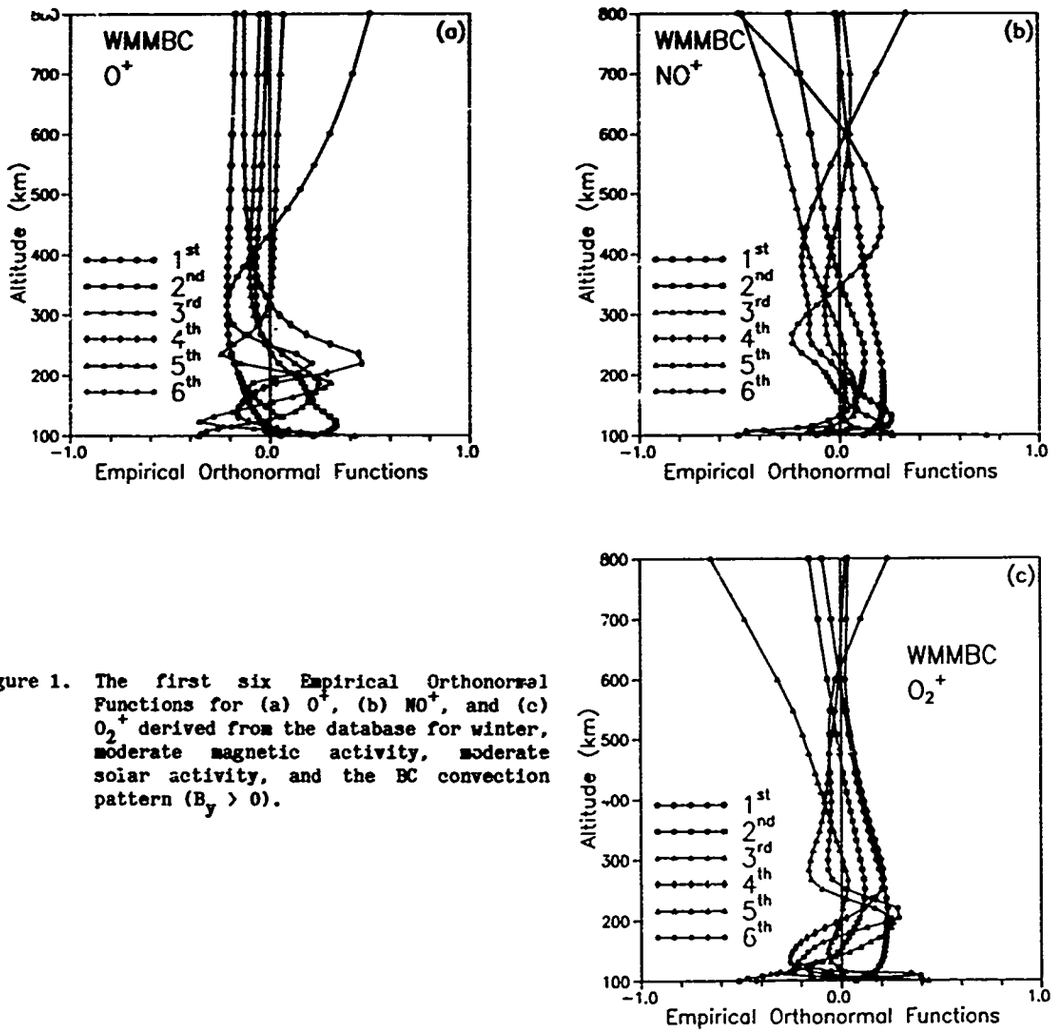


Figure 1. The first six Empirical Orthonormal Functions for (a) O^+, (b) NO^+, and (c) O_2^+ derived from the database for winter, moderate magnetic activity, moderate solar activity, and the BC convection pattern ($B_y > 0$).

The next stage involves the approximation of the horizontal variation of the EOF coefficients. If we were dealing with global databases, we would use some form of spherical harmonic expansion. However, since we are dealing with regions of limited latitude extent, there does not appear to be any special advantage to this approach, so we have chosen to attack each spatial variable in sequence.

For each universal time and magnetic latitude the coefficients, a_{nk}, of the EOF's are represented by Fourier series in magnetic local time (or longitude). (Trigonometric functions have the useful property that they are orthogonal on uniform discrete grids as well as on continuous intervals.) In the high latitude region, we used a nine term series (DC plus four cosine and four sine terms).

For each universal time the Fourier coefficients were represented by linear combinations of discrete orthonormal polynomials in magnetic latitude. These polynomials are generated for a given latitude grid by the algorithm of Beckman [1973]. We found that a nine term series gave a good representation of the latitude variations.

For the high latitude model, the low order polynomial coefficients showed primarily diurnal variations in universal time, but the higher order terms displayed considerably more structure. Consequently, we chose to tabulate the coefficients in UT and use polynomial interpolation rather than to use an orthogonal function expansion. This may not be necessary at middle and lower latitudes.

OPERATIONAL USE OF THE PARAMETERIZED MODELS

Once obtained, the semi-analytic representations of the physical models are used in two different ways. First, when real time ionospheric data are unavailable, they provide "best guesses" for the current state of the ionosphere based on estimates of the current levels of solar and geomagnetic activity. Second, when real time data are available, they are scaled to provide the best (least squares) fit to that data, and are used to give the altitude profile shapes. The way in which the parameterized model is scaled and adjusted by real time data depends on the region. One advantage of working with ion densities (rather than electron densities) is that the separation between the E-layer and F-layer can be made in a natural way: the molecular ions provide the E-layer while the atomic ions provide the F-layer.

The Northern and Southern High Latitude Regions

The high latitude ionosphere is very dynamic and highly structured making near real-time specification of electron density profiles extremely difficult. Fortunately, for many purposes it is the location of the structures that is of greatest interest with electron densities of lesser significance. We treat the E- and F-layers independently. For the F-layer, we divide the high latitude ionosphere into three sub-regions: the Trough, the Auroral Oval, and the Polar Cap.

The dynamic nature of the high latitude ionosphere is a direct result of the dynamic nature of the energy sources that drive it: magnetospheric convection and particle precipitation. Furthermore, the relatively long time scale of the F-layer (hours) implies that the current state of the ionosphere depends on recent history as well as current energy and momentum input. For this reason, even a perfect physical model would fail to accurately specify the current ionosphere unless it had accurate specifications of the recent history of the energy and momentum input. We have adopted a hybrid approach combining both empirical and physical models.

The physical model we adopted is the Utah State University high latitude ionospheric model [Schunk, 1988]. This model calculates ion and electron densities and temperatures by following flux tubes convecting through a moving neutral atmosphere. An analytic version [Rich and Maynard, 1989] of the convection model of Heppner and Maynard [1987] was used to specify the convection pattern for various levels of magnetic activity. Ionization rates due to particle precipitation were calculated using the Strickland et al. [1976] electron transport code and the electron precipitation model of Hardy et al. [1987]. Photoionization rates were calculated using photon fluxes from Hinteregger et al. [1977] and Heroux and Hinteregger [1978] (scaled with $F_{10.7}$) and the weighted cross sections of Torr et al. [1979].

The USU model was used to produce 54 databases containing NO^+, O_2^+, and O^+ densities as functions of altitude, magnetic latitude, magnetic local time, and universal time for a variety of environmental conditions. The specific parameters used were

season: equinox: day 82 (March 23)
June solstice: day 173 (June 22)
December solstice: day 357 (December 23)

solar activity ($F_{10.7}$): high (210), moderate (130), and low (70)

magnetic activity (K_p): high (6.0), moderate (3.5), and low (1.0)

Interplanetary magnetic field: $B_z < 0$ (southward)
$B_y > 0$ (BC convection pattern)
$B_y < 0$ (DE convection pattern)

Because of the uncertainty in convection patterns for $B_z > 0$ (northward), we made no calculations

for this case. Instead, we use the low magnetic activity cases to model B_z northward situations.

The first set of 54 databases were obtained for the northern hemisphere. Because the relationship between the magnetic and geographic poles is different in the two hemispheres, a second set of 54 databases is being generated for the southern hemispheres. This set will be treated in the same manner as the first set. The northern hemisphere databases have been approximated semi-analytically using EOF's, Fourier series, and discrete orthogonal polynomials as described in the previous section. Comparisons of the semi-analytic representations with the original databases are shown in Figure 2.

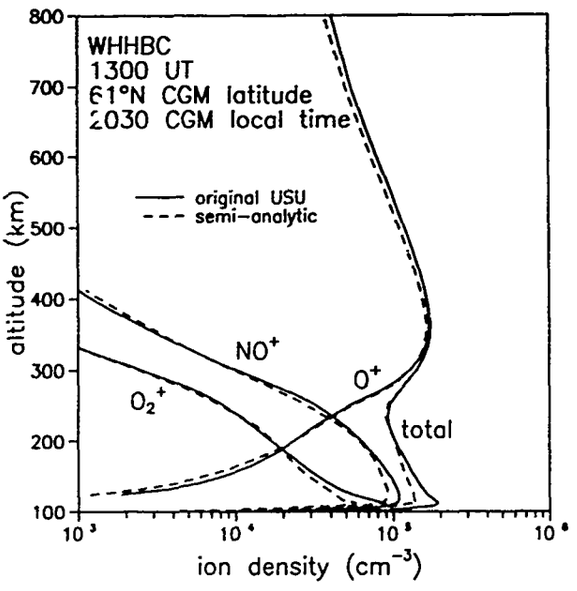


Figure 2. A comparison of ion density profiles from the original USU databases with the semi-analytic fits based on EOF's. These profiles are from the winter, high magnetic activity, high solar activity, B_y positive database. The fit is quite good except near the bottom where there is altitude structure not completely captured by the first six EOF's.

The real time specifications of the high latitude E- and F-layers are handled separately. This is appropriate because of the disparity in time scales associated with each altitude regime. The E-layer can be handled in two different ways. For post event analysis, when computer resources are abundant, a first principles chemical equilibrium model is used. Ion production rates can be determined from a combination of satellite based electron precipitation data, satellite borne nadir viewing UV and X-ray intensity data, and statistical models of precipitation patterns. For real time analysis, when computer resources are scarce, the semi-analytic representations of the molecular ion densities are scaled to provide a least squares fit to the relevant real time data (mainly ionosonde measurements of f_oE). In addition, if information on the location of the auroral oval is available (e.g., from precipitating particle measurements or from visible or UV imagery) the latitude scale is adjusted accordingly.

Unlike the mid- and low-latitude regions, the semi-analytic representation of the high latitude physical model is not directly scaled to fit the real-time F-layer data. Instead, simple semi-empirical models of the three F-layer subregions are used. These are models of f_oF_2 only, so the parameterized physical model is used to provide the shapes of the O^+ altitude profiles.

The F-layer trough. Our semi-empirical trough model is based on the studies of Whalen [1989]. There appear to be three different local time features that together make up the complete trough. In the afternoon sector from shortly after magnetic local noon and fading out between the terminator and magnetic local midnight is a stable feature closely related to the ionospheric convection pattern. Because of the long time scale associated with F-layer phenomena, this feature does not respond to the instantaneous convection pattern, but "averages" over conditions prevailing during the previous several hours. From the end of the afternoon trough until sunrise in the F-layer the observed trough appears to be the result of superimposing auroral density enhancements on top of the naturally decaying nighttime ionosphere. However, the equatorward edge of the auroral density enhancement does not necessarily coincide with the equatorward edge of the precipitation because of convection. Finally, a morning trough sometimes develops between local dawn and local noon. The

precise conditions under which this trough appears are not clear at this time.

In our model, the equatorward edge of the trough is continuous in local time and is described as a distorted circle with the same shape as the convection boundaries of Heppner and Maynard [1987]. The radius of the circle is treated as an adjustable parameter. The width and depth of each trough feature (afternoon, nighttime, and morning) are independently adjustable. In the absence of sufficient data to fix these parameters, the afternoon trough is assigned a nominal depth of 2.5 MHz and a nominal width of 5^{o} latitude, while the nighttime and morning trough parameters are set to reproduce the midlatitude ionosphere. The O^{+} densities calculated by the midlatitude models provide "boundary conditions," and the depth parameters are calculated relative to those densities.

The auroral F-layer. Because the auroral F-layer is controlled by several different and competing processes (e.g., precipitation and convection), and because it is of less interest than the auroral E-layer, we use a very simple model. The equatorward edge is simply the poleward edge of the trough, and the poleward edge is simply the equatorward edge of the polar cap. A single adjustable parameter describes the peak f_oF_2 relative to the f_oF_2 value at the polar cap boundary. Although the potential to adjust the location of the peak f_oF_2 is retained, under normal circumstances it is simply centered between the two boundaries.

The polar cap F-layer. The polar cap is the most difficult region to model, and in the absence of a dense network of high time resolution ionosondes, impossible to specify completely. Under B_z south conditions, convection brings high density daytime plasma into the polar cap. Because convection is rarely a steady process, the usual result is patches of varying size that move through the polar region. Under B_z north conditions, the polar cap ionosphere is more uniform, but density irregularities at various scales may still occur. We attempt only to provide a crude model of the background ionosphere and to give statistical information on the occurrence of patches. If possible, high time resolution digital ionogram data will be used to estimate the interpatch (background) f_oF_2 values, intrapatch f_oF_2 values, and the size and spacing of the patches. The background plasma density will be obtained by scaling the CCIR coefficients to digital ionosonde observations in the polar cap. In the absence of such measurements, the current effective sunspot number will be used.

We are currently in the process of testing and refining this part of the model. Two preliminary tests of the model using actual data are shown in Figures 3 and 4. In both cases, only ground based ionosonde data was used in the parameter adjustment process. Independent data was used to check the accuracy of the adjustment in reproducing ionospheric features.

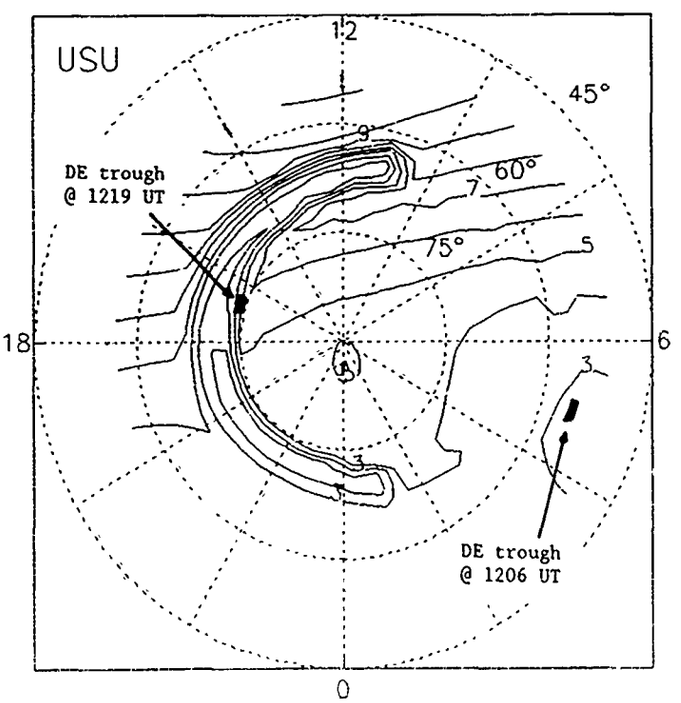


Figure 3. A contour plot of f_oF_2 (MHz) in geographic coordinates based on ionosonde data taken at 1200 UT on 21 December 1981. Also shown are the locations of the afternoon and morning trough minima as determined from DE-2 in situ density measurements by Santimay Basu (personal communication). There were a total of five ionosonde stations reporting data in the high latitude region. The model required four iterations to converge to the configuration shown here.

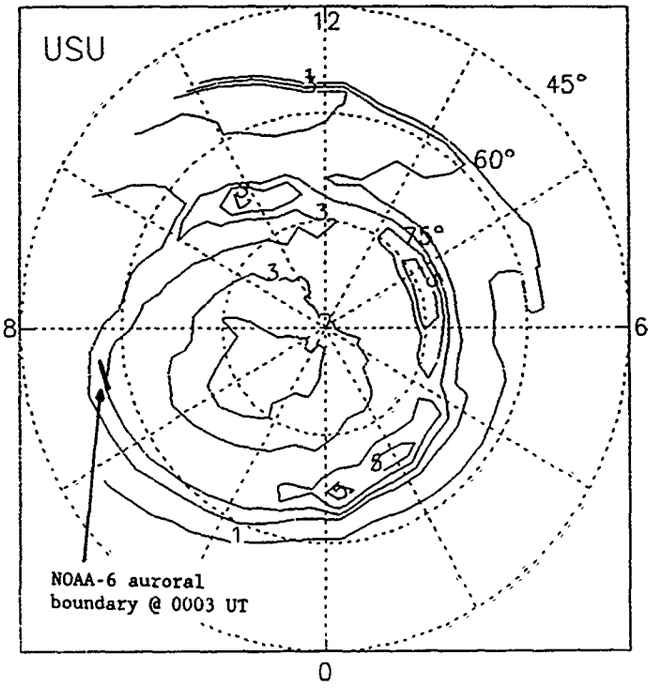


Figure 4. A contour plot of f_0E (MHz) at 0000 UT on 21 December 1981. Also shown is the equatorward boundary of the auroral oval (0.3 erg cm^{-2} s^{-1}) as determined from NOAA-6 data by H. W. Kroehl (personal communication). There were a total of seven ionosonde stations reporting high latitude information, of which only one reported E-layer data.

In both of these cases, reasonable agreement was obtained despite the limited data set used to drive the model. We anticipate even better accuracy in normal operation with a variety of near real time data available. After further testing and refinement, this model will be validated against an extensive set of data intended to resemble as closely as possible the expected complement of operational data.

The Northern and Southern Midlatitude Regions

The midlatitude ionosphere is less dynamic and less structured than the high latitude ionosphere. Consequently, it is not necessary to resort to semi-empirical subregion models. Furthermore, convection is unimportant so the IMF orientation may be ignored. Otherwise, the procedure is very similar to that used for high latitudes.

The midlatitude physical model is a combination of the photoelectron model of Jasperse [1981], a photochemical equilibrium E-layer model, and a midlatitude version of the F-layer (O^+) model of Anderson [1973].

The semi-analytic representation of the midlatitude physical model is being produced by the same method used for the high latitude model. The real time operation will be somewhat different, however. A global ionospheric index, roughly corresponding to the SSN_{eff} of the original ICED, will be determined from the complete set of Air Weather Service digital ionosondes. This index will be used to scale the global model. However, regional indices will also be generated using ionosonde and other data (TEC, in situ plasma densities and temperatures, UV imagery, etc.). These indices will allow the ion density scalings to vary with spatial location, resulting in a more accurate representation of the ionosphere where data is abundant without introducing extraneous changes where data is absent. As with the high latitude data, the profile shapes come from the theoretical model, although the topside shapes will be modified where appropriate data is available. In particular in situ temperature, density, and composition measurements in the topside F-layer will be used to adjust the theoretical topside profiles to provide a more accurate representation.

The Low Latitude Region

The low latitude ionosphere has many of the characteristics of the high latitude ionosphere: it is dynamic and exhibits spatial structure. As for the polar cap, we do not attempt to specify the details of the small scale structure (bubbles, plumes, etc.). The incorporation of FAIM into ICED provided limited low latitude capability, primarily in the region of low magnetic declination. FAIM includes an analytic representation of the low latitude O^+ model of Anderson [1973, 1981] using ion drift data from Jicamarca [Woodman, 1970; Fejer et al., 1979] and assuming no magnetic declination effects. No additional real-time information (beyond SSN_{eff} and Q_{eff}) was incorporated in the model. We are using the same model (combined with our midlatitude E-layer model) along with additional data and a more realistic magnetic field as the basis for the GLOBAL ICED low latitude model. As in the high and middle latitude regions, the ability to adjust itself on the basis of near real time data will be an integral part of the model.

The semi-analytic representation of the physical model is obtained using the same methods as for the high and middle latitude regions. A series of databases is being produced for a range of environmental conditions. The altitude profiles will be represented by a linear combination of EOF's, and the coefficients will be represented by Fourier series in local time and discretely orthogonal polynomials in latitude.

The real time operation will make use of in situ electron density measurements and UV airglow images to determine the location of the equatorial anomaly peaks (or the width of the equatorial bulge when the anomaly is absent) to set the scale of the electric field. TEC data and/or UV and visible airglow measurements will be used to set the scale of the density. As in midlatitudes, adjustments will be made regionally as well as globally to produce the best least squares fit to the available data.

THE INTEGRATED MODEL: GLOBAL ICED

The five regional models are to be combined to produce a single GLOBAL ICED. The midlatitude model will be used to define the values of f_oF_2 at the equatorward edge of the trough, and these are fixed as "boundary conditions" on the high latitude F-layer model. Continuity of the O^+ profiles is insured by taking a weighted average of the mid-latitude and high-latitude profiles in the trough wall. Continuity of the E-layer (molecular ions) is insured by doing a similar weighted average over the same latitude range.

The boundary between the low and middle latitude regions is less sharply defined. Because the midlatitude and low latitude models are basically the same, the principle difference being the absence of electric fields in the midlatitude model, the two models must give the same results far enough from the magnetic equator. The latitude at which the two models agree will become the boundary between the two regions and will move according to actual conditions as detected in the real time data. Any minor discrepancy between the two models will be smoothed out by using a weighted average in the region of overlap.

Integration will proceed as soon as all the regional models have been tested, refined, and validated against actual data. Once integration is complete, GLOBAL ICED will undergo a final process of testing, refinement, and validation before being turned over to the Air Weather Service for conversion to operational use. We expect GLOBAL ICED to be complete by the end of 1991.

FUTURE DEVELOPMENT

We see GLOBAL ICED as the first step in the continued evolution of ICED. As scientific understanding of the physical processes that control the ionosphere continues to improve, and as computers continue to become more powerful, we expect the physical basis of the model to become stronger and the reliance on semi-empirical models and semi-analytical parameterizations to be reduced. The ultimate goal is a unified first principles model running in real time with periodic adjustments from real time measurements to compensate for the uncertainties inherent in any attempt to specify the global energy and momentum inputs to the ionospheric system. We also have begun the process of developing a limited forecast model whose evolution will parallel that of ICED.

REFERENCES

Anderson, D. N., A theoretical study of the ionospheric F region equatorial anomaly, 1, Theory, Planet. Space Sci., 21, 409-419, 1973.

Anderson, D. N., M. Mendillo, and B. Herniter, A Semiempirical Low-latitude Ionospheric Model, Radio Sci., 22, 292-306, 1987.

Anderson, D. N., J. M. Forbes, and M. Codrescu, A Fully Analytic, Low- and Middle-Latitude Ionospheric Model, J. Geophys. Res., 94, 1520-1524, 1989.

Chiu, Y. T., An Improved Phenomenological Model of Ionospheric Density, J. Atmos. Terr. Phys., 37, 1563-1570, 1975.

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, 5792-5796, 1979.

Hardy, D. A., M. S. Gussenhoven, R. Raistick, and W. J. McNeil, Statistical and Functional Representation of the Pattern of Auroral Energy Flux, Number Flux, and Conductivity, J. Geophys. Res., 92, 12275-12294, 1987.

Heppner, J. P., and N. C. Maynard, Empirical High-Latitude Electric Field Models, J. Geophys. Res., 92, 4467-4489, 1987.

Heroux, L., and H. E. Hinteregger, Aeronomical reference spectrum for solar UV below 2000 Å, J. Geophys. Res., 83, 5305-5308, 1978.

Hildebrand, F. B., Methods of Applied Mathematics, Prentice-Hall, Englewood Cliffs, pp. 30-34, 1965.

Hinteregger, H. E., D. E. Bedo, J. E. Manson, and D. R. Skillman, EUV flux variations with solar rotation observed during 1974-1976 from the AE-C satellite, Space Res., 17, 533-544, 1977.

Jasperse, J. R., The photoelectron distribution function in the terrestrial ionosphere, in Physics of Space Plasmas, ed. by T. S. Chang, B. Coppi, and J. R. Jasperse, Scientific Publishers, Cambridge, MA, pp. 53-84, 1982.

Kutzbach, J. E., Empirical Eigenvectors of Sea-Level Pressure, Surface Temperature, and Precipitation Complexes over North America, J. Appl. Meteor., 6, 791, 1967.

Rich, F. J., and N. C. Maynard, Consequences of Using Simple Analytical Functions for the High-Latitude Convection Electric Field J. Geophys. Res., 94, 3687-3701, 1989.

Schunk, R. W., A Mathematical Model of the Middle and High Latitude Ionosphere, Pageoph, 127, 255-303, 1988.

Secan, J. A., and R. F. Tascione, The 4D Ionospheric Objective Analysis Model, in Proceedings of the 1984 Ionospheric Effects Symposium, ed. by Goodman, Klobuchar, and Soicher, 336-345, 1984.

Strickland, D. J., D. L. Book, T. P. Coffey, and J. A. Fedder, Transport Equation Techniques for the Deposition of Auroral Electrons, J. Geophys. Res., 81, 2755-2764, 1976.

Tascione, T. F., R. W. Kroehl, R. Croiger, J. W. Freeman, R. A. Wolf, R. W. Spiro, R. V. Hilmer, J. W. Slade, and B. A. Hausman, New Ionospheric and Magnetospheric Specification Models, Radio Sci., 23, 211-222, 1988.

Torr, M. R., 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., 10, 771-774, 1979.

Whalen, J. A., The Daytime F Layer Trough and Its Relation to Ionospheric-Magnetospheric Convection, J. Geophys. Res., 94, 17169-17184, 1989.

Woodman, R. F., Vertical drift velocities and east-west electric fields at the magnetic equator, J. Geophys. Res., 75, 6249-6259, 1970.

AD-P006 303

IONOSPHERIC ERRORS AT L-BAND FOR SATELLITE AND RE-ENTRY OBJECT TRACKING IN THE EQUATORIAL ANOMALY REGION

W. A. Pakula
Cadet Squadron 33
USAF Academy, Colorado 80841

J.A. Klobuchar and D. N. Anderson
Ionospheric Physics Division
Hanscom AFB, MA 01731

P. H. Doherty
Institute for Space Research
Boston College
Newton, MA 02159

ABSTRACT

The ionosphere can significantly limit the accuracy of precise tracking of satellites and re-entry objects, especially in the equatorial anomaly region of the world where the electron density is the highest. To determine typical changes induced by the ionosphere, the Fully Analytic Ionospheric Model, (FAIM), was used to model range and range-rate errors over Kwajalein Island, located near the equatorial anomaly region in the Pacific. Model results show that range-rate errors of up to one foot per second can occur at L-band for certain, near-vertical re-entry object trajectories during high solar activity daytime conditions.

Since the variability of the location and strength of the equatorial anomaly peak electron density region is large, models which fail to take into account the actual position of the anomaly, at the time when the space object is being tracked, can result in erroneous range and range-rate corrections. Even relatively small velocity errors can be significant if they have the cumulative effect of producing large total range errors over the entire space object flight path.

INTRODUCTION

The purpose of this investigation was to model the effects the ionosphere can have on precise tracking of a satellite by an L band radar, as well as the effects on tracking a re-entry object, either being tracked by a radar, or carrying an L-band transponder, in producing target range and velocity errors. In order to demonstrate "worst case" conditions the equatorial anomaly region of the world was chosen because the highest electron densities in the world occur in this region. The equatorial anomaly produces the world's highest ionospheric electron densities in regions displaced approximately plus and minus 15^{o} either side of the earth's magnetic equator, due to an electric field at the magnetic equator which moves ionization upwards at the equator, then across magnetic field lines to the anomaly latitudes. Since Kwajalein Island is located nearly under the equatorial anomaly it was chosen as a representative location to study the effects of ionospheric range and range-rate errors on precise tracking of re-entry objects. The ionospheric model used was the Fully Analytic Ionospheric Model, FAIM, developed by Anderson, et. al., (1989), which does a realistic job of representing the low latitude F region for various seasonal and solar cycle activity conditions.

91-09699

IONOSPHERIC RANGE AND RANGE-RATE ERRORS

The additional time delay, over the free space transit time, of a signal transmitted from above the ionosphere to a user on, or near, the earth's surface was given by Millman (1965) and can be expressed as:

$$(\text{Delta})\ t = \frac{40.3}{c * f^2} * \text{TEC} \qquad \text{(seconds)} \qquad (1).$$

where TEC is the total number of electrons along the path from the radar to the target, c is the velocity of light in meters/second, and f is the system operating frequency in Hertz. The TEC is generally expressed as the number of electrons in a unit cross section column of one square meter area along the radar signal path.

For a radio or radar signal at 1.6 GHz, a TEC value of 10^{18}, a value commonly exceeded in the equatorial anomaly region, would produce a one-way range error due to the ionosphere of 15.7 meters, or 51.6 feet.

The one way error in range-rate, or velocity, which the ionosphere produces on radio or radar signals can be expressed as:

$$\text{Delta (phi)} = \frac{1.34 \times 10^{-7}}{f} * \text{TEC} \qquad \text{(cycles)} \qquad (2).$$

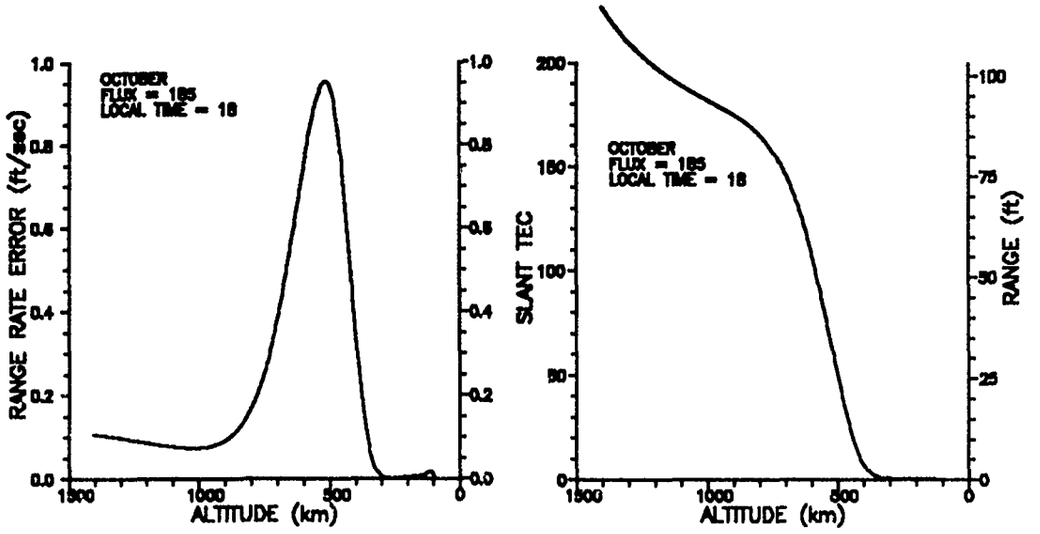
where Delta (phi) is the number of RF carrier cycles of phase path decrease suffered by the signal. At 1.6 GHz, one cycle, (wavelength), corresponds to 0.62 feet of distance, and for a rate of change of TEC of 10^{16} el/m^2 per second, the range-rate, or velocity, error would be 0.52 feet per second. If that rate of change continued for 30 seconds, the cumulative range error would be 4.76 meters or 15.6 feet.

Range-rate errors are produced by the rate of change of the TEC over the radio or radar target velocity measurement interval. Since this measurement interval can be as short as one second, the change in TEC over a one second interval is the important parameter in computing ionospheric range-rate errors. The change in TEC, looking in a fixed direction, normally does not change rapidly, since only the slow diurnal changes in TEC are involved. However, a radar tracking a space object, especially an object re-entering the atmosphere, can see relatively large changes in TEC along the target trajectory, especially as it goes through the region of ionospheric peak electron density. Thus, the effective rate of change of TEC which the radar sees, over its velocity tracking interval, is the important parameter, in range-rate calculations, and not the absolute value of TEC itself.

RANGE AND RANGE-RATE CALCULATIONS FOR A RE-ENTRY OBJECT

Since the operating frequency of the tracking system is known, the fundamental problem becomes one of calculating the ionospheric TEC along a Line Of Sight (LOS) from the transmitter to the re-entry object. Mathematically, this is equivalent to integrating an electron number density function (electrons per cubic meter) over a certain path (meters) to get units of electrons per square meter. Rates of change were then calculated by the difference of TEC over successive one second intervals as the re-entry object moved downwards through the ionosphere.

First, an orbital generation program was used to determine a typical trajectory of a re-entry object. The prepared input file contained the geographic latitude, longitude, and altitude of the re-entry object for each second of a typical flight path following apogee. It also contained the x, y, z, Earth-Centered, Earth-Fixed (ECEF) coordinates of the re-entry object for the same times. The ECEF system is simply that right handed, orthogonal coordinate system with the x-axis pointing towards the Greenwich meridian and the z-axis pointing towards the north geographic pole. The other coordinate systems used in the program are geographic and geomagnetic Local Level (LL) coordinates. LL

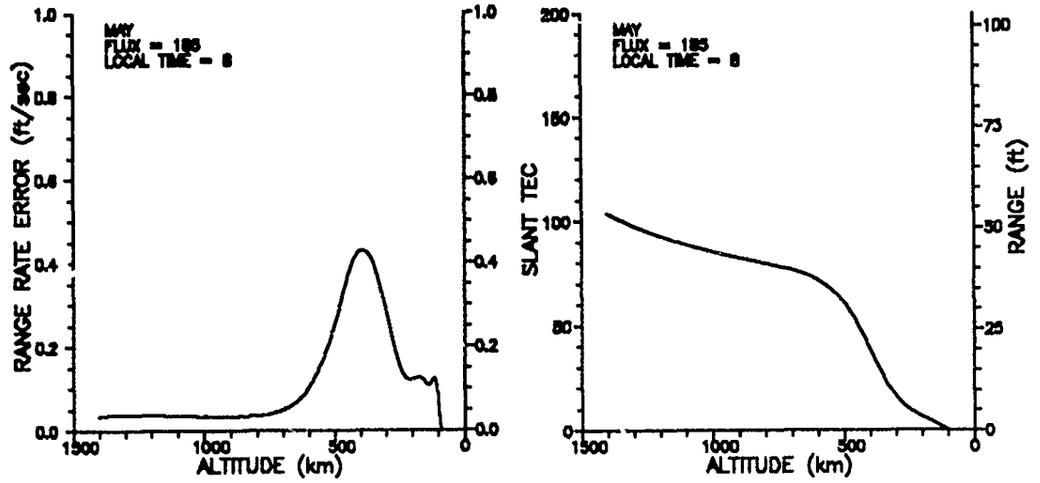


a) Range-rate error versus altitude for October, solar flux = 185, 18 hours local time.

b) Range error versus altitude for October, solar flux = 185, 18 hours local time.

FIGURE 1.

For the month of May, for the same solar flux, but at 0800 local time, when the TEC is much lower, the maximum range-rate error is 0.430 feet per second and occurs at an altitude of 394 kilometers, as illustrated in Figure 2A. The corresponding slant TEC, corresponding to radar range error as a function of re-entry object height is shown in Figure 2B.

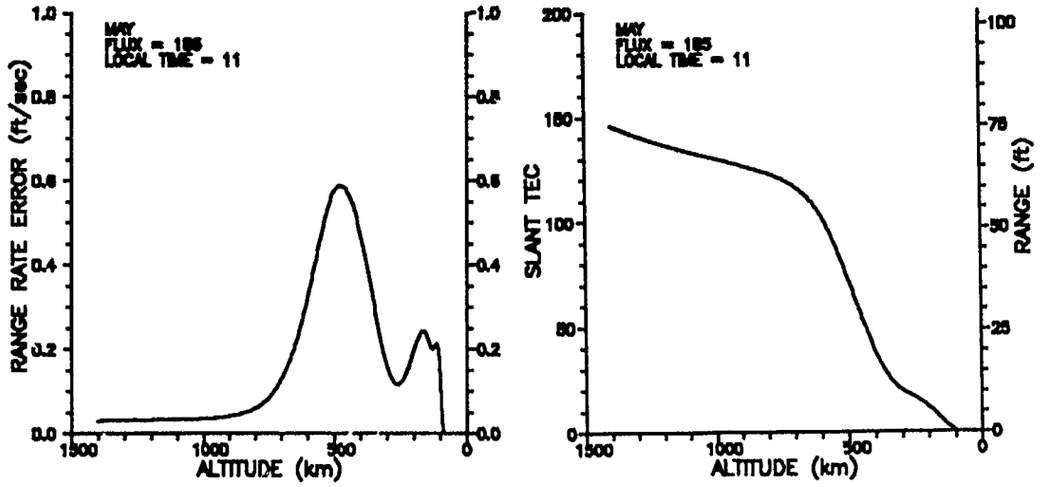


a) Range-rate error versus altitude for May, solar flux = 185, 08 hours local time.

b) Range error versus altitude for May, solar flux = 185, 08 hours local time.

FIGURE 2.

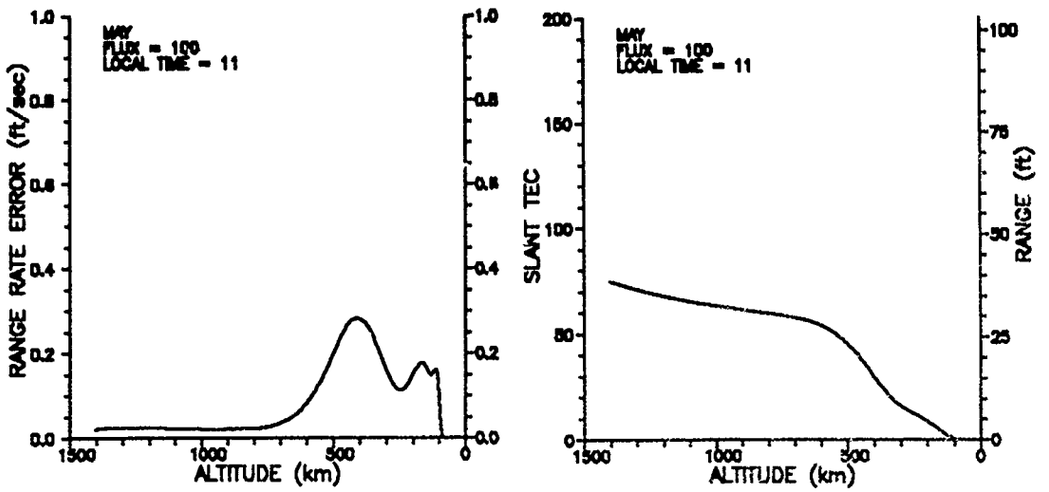
For the same conditions as those in Figure 2, namely May, with a solar flux of 185, but at 1100 hours local time, the maximum range-rate is 0.587 feet per second which occurs at an altitude of 476 km. The range-rate and TEC for this case as shown in Figures 3A and 3B respectively. To contrast this case against one with a low value of solar flux, Figures 4A and 4B illustrate the range-rate and slant TEC, respectively, for May at 1100 hours local time, but with a solar flux of 100. The largest range-rate error for this case was 0.282 feet per second.



a) Range-rate error versus altitude for May, solar flux = 185, 11 hours local time.

b) Range error versus altitude for May, solar flux = 185, 11 hours local time.

FIGURE 3.



a) Range-rate error versus altitude for May, solar flux = 100, 11 hours local time.

b) Range error versus altitude for May, solar flux = 185, 11 hours local time.

Figure 4.

coordinates consist of latitude, longitude, and altitude. Once the input file was prepared, the actual velocity error calculations were made. All test cases used Kwajalein Island as the simulated tracking site, due to its location nearly under the equatorial anomaly.

Second, the ECEF coordinates of forty points along the direct LOS from the tracking station to the first point in the re-entry object trajectory were calculated by means of an iterative loop. Each coordinate triplet was also converted, within the loop, to geographic LL coordinates. As each LOS point was calculated, the Local Time (LT) for each point was determined by finding the difference in longitude between the LOS point and the splashdown point of the re-entry object, and adding/subtracting the equivalent number of hours to the splashdown time. Once the LOS coordinate array and corresponding LT array were filled for all forty points, the entire LOS array was converted from geographic to geomagnetic LL coordinates.

Next, the electron number densities at each point in the LOS array were found through repeated calls to the FAIM model ionospheric code. Important inputs to FAIM consist of the geomagnetic LL coordinates and the LT for the point and time of interest, the current month, and solar flux. The model then returns the number electron density for the given point and the specified conditions. FAIM was used to calculate the electron density for each of the forty points along the LOS. These number densities were then integrated along the LOS, resulting in the TEC of the ionosphere from the tracking station to the re-entry object. This overall process of finding the TEC was repeated for each LOS, one per each second of flight time, from apogee to splashdown.

Once the TECs along each slanted LOS were known, the phase advance of the tracking signal for each point in the trajectory were easily found using equation (2). Using a curve fitting package, a phase advance versus re-entry object flight time was constructed. Finally, the derivative of this curve was calculated for each point in the re-entry object trajectory. These were used in accordance with (2) to calculate the range- rate errors produced at each point in the trajectory. To simplify the calculations the location of the re-entry object splashdown and the ground radar were assumed to be coincident, though small differences in these locations do not significantly change the results.

Among the assumptions made by the program are the following:

- spherical earth coordinate transformation from ECEF to LL coordinate system
- splashdown location and radar location are the same
- the approx. 6 minute travel time of RV is negligible in inputting time to FAIM model, that is any diurnal changes in electron density were small during that time interval
- electron number density is set to zero below 50 kilometers re-entry object altitude
- solar flux activity levels greater than 185 are not accepted by FAIM, and are thus not used

RESULTS

Ten test cases were run for different seasons, solar activity and local times of day. The maximum range-rate error for all the test cases was for October at 18:00 hours Kwajalein time, with a solar 10.7 cm. flux of 185. For those conditions the ground radar will experience a maximum range-rate error of -0.955 feet per second, at the point when the re-entry object is at an altitude of 516 kilometers. This is shown in Figure 1A, along with the corresponding slant TEC curve on Figure 1B.

In all test cases the maximum range rate errors were produced when the slant TEC curves had the greatest slope, or first derivative. This was to be expected, since the range rate error equation includes the first derivative of the phase advance, which is dependent upon the TEC. This can be seen by referring, for example, to figures 1A and 1B. Note that the maximum range-rate error, in Figure 1A, occurs at the point, in Figure 1B, where the rate of change of TEC is greatest. Thus, in the case of an object traveling rapidly through the ionosphere, it is the rate of change of TEC, and not the TEC itself, which dictates the magnitude of the range-rate error. Table 1 summarizes the maximum range rate errors calculated for all test cases, along with the altitude at which they occur.

TABLE 1.

Case	Month	Flux	Time (hrs)	....MAXIMUM ERROR.... Range (ft)	Range Rate (ft/sec)	Altitude (km)
1	MAY	185	11	75	-0.59	476
2	MAY	185	8	54	-0.43	394
3	OCT	185	11	106	-0.87	440
4	MAY	100	11	39	-0.28	412
5	OCT	185	18	117	-0.96	516
6	JAN	185	11	72	-0.68	339
7	JAN	100	11	38	-0.38	287
8	MAR	185	11	101	-0.82	434
9	MAR	18[illegible]	6	18	-0.197	307
10	MAY	185	14	101	-0.798	480

Table 1: Maximum range and range-rate errors. The maximum range error occurs when the object is at its greatest range. The maximum range-rate error occurs at the altitude indicated.

CONCLUSIONS

The range and range-rate errors produced by the ionosphere on an L-band radar or tracking beacon can be significant if precision tracking is required. Analysis of the problem using the FAIM ionospheric model predicts range-rate errors to be on the order of 1/2 to 1 foot per second for the conditions used in this study. Due to the limitations of the FAIM model solar flux values only up to 180 could be used, however, the smooth 10.7 cm. solar flux for the present solar cycle maximum is approximately 250. The values for ionospheric range and range-rate error presented here are certainly lower than those which would be encountered during the present solar maximum conditions.

REFERENCES

Anderson, D. N., J. M. Forbes, and M. Codrescu, "A Fully Analytic, Low-and Middle-Latitude Ionospheric Model", J. Geophys. Res., V. 94, No. A2, pp 1520-1524, Feb. 1, 1989.

Millman G. H. in Modern Radar, Edited by R. S. Berkowitz, John Wiley & Sons, New York, (1965).

AD-P006 304

91-09698

THE LATITUDINAL EXTENT OF THE EQUATORIAL ANOMALY

J. A. Klobuchar and D. N. Anderson
Ionospheric Physics Division
Air Force Geophysics Laboratory
Hanscom AFB, MA 01731

P. H. Doherty
Institute for Space Research
Boston College
Newton, MA 02159

ABSTRACT

The latitudinal extent of the equatorial anomaly has been studied using a theoretical model of the ionosphere which incorporates measured values of vertical E x B drift at the earth's magnetic equator. Realistic values of neutral winds are also included. The equatorial anomaly region, typically between plus and minus 20 degrees magnetic latitude, is that part of the world where the highest values of electron density and Total Electron Content, (TEC), normally occur, and hence is very important to high frequency propagation and to trans-ionospheric propagation effects. During the daytime upward E x B drift at the magnetic equator drives the ionization across field lines to higher latitudes, causing crests in ionization to occur at approximately plus and minus 15° dip latitude.

The latitude range over which the anomaly makes a significant difference in values of foF2 and TEC is calculated as a percent departure from the case with no equatorial electric field. Results from the model studies with different values of realistic electric fields show that the effects of the anomaly can be highly variable and widespread in latitude and local time.

INTRODUCTION

The equatorial anomaly region of the earth's ionosphere is characterized by the nighest values of peak electron density, N_{max}, and total electron content, (TEC), as well as the greatest amplitude and scintillation effects, in the worldwide ionosphere. It is well known that these large values of peak electron density and total electron content are due directly to the daytime upward E x B drift at the magnetic equator which drives the ionization across field lines to higher latitudes, causing crests in ionization to occur at approximately plus and minus 15° dip latitude, Hansom and Moffett, (1966). What may not be generally recognized, is the extent to which the equatorial "fountain" really modifies the otherwise directly solar-controlled ionosphere at low latitudes. The expression "fountain effect" apparently was first used by Martyn (1955), in describing the equatorial anomaly, and, as shown in this paper, the term is very appropriate.

There have been numerous papers in the literature illustrating the experimental effects of the equatorial anomaly, (sometimes called the Appleton anomaly), in foF2, e. g., Appleton, (1946), Rastogi, (1959), Rush and Richmond (1973), as well as in TEC, e.g. Deshpande, et al., (1977), Huang, et. al., (1989). However, no paper has yet attempted to determine 1) the latitude and local time difference between the equatorial anomaly as observed in N_{max} and in TEC, 2) the sensitivity of these two quantities to variations in vertical E x B drift patterns, and 3) the dependence of this difference on solar activity. In addition, the overall

effects of the anomaly, and its variability, on the entire distribution of electron density and TEC in the low latitude ionosphere have not been investigated.

In this paper a theoretical model of electron density developed by Anderson (1973) has been used, with realistic values of vertical drift and neutral winds, to calculate the behavior of the F region of the earth's ionosphere over a range of latitudes which covers the effects of the equatorial anomaly, even under extreme conditions, to determine the overall geographic and temporal effects of the anomaly.

MODEL CALCULATIONS

Equation 1. presents the ion (O^+) continuity equation which is solved numerically to find ion and electron densities as a function of altitude, latitude and local time.

$$\frac{\partial N_i}{\partial t} + \nabla \cdot (N_i \mathbf{V}_i) = P_i - L_i \tag{1}$$

N_i is the ion density, V_i the ion velocity, P_i the ionization rate by solar ultraviolet radiation, and $_i$ the loss rate through charge exchange with N_2 and O_2. The ion velocity, V_i includes plasma diffusion and collisions with neutrals along geomagnetic field lines and E X B drift perpendicular to the field lines. Solving (1) at low latitudes is facilitated by transforming the independent coordinates r, , and to a coordinate system parallel and perpendicular to B, (Anderson, 1973).

Equation (1) can then be written :

$$\frac{\partial N_i}{\partial t} + \mathbf{V}_{i_\perp} \cdot \nabla N_i = P_i - L_i - \nabla \cdot (N_i \mathbf{V}_{i_\parallel}) - N_i \nabla \cdot \mathbf{V}_{i_\perp} \tag{2}$$

The right hand side of (2) contains terms which are second order in the coordinate parallel to B while the left hand side gives the time-rate-of-change of the ion density in a frame of reference which drifts with the E X B velocity. In this study we include only the vertical component of E X B drift and allow the plasma to co-rotate with the earth. For a review of the transformations and numerical procedures see Moffett (1979).

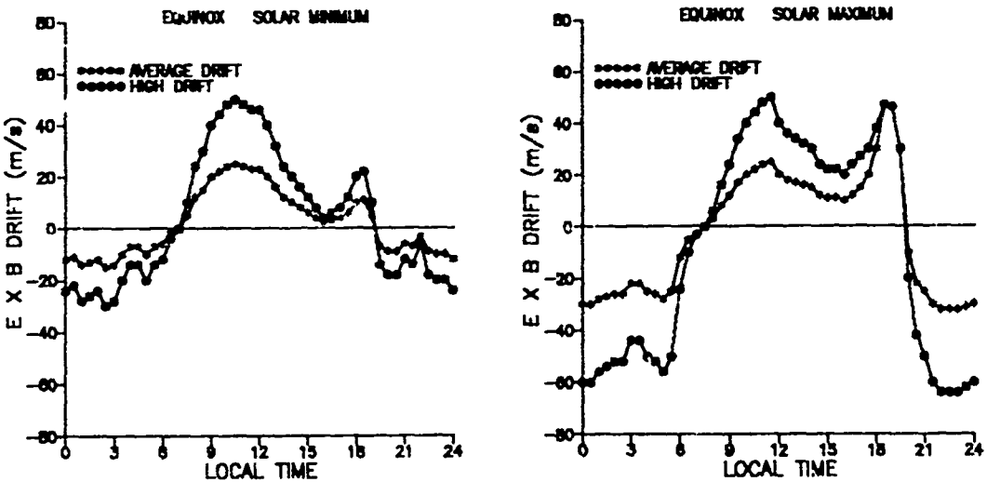
In order to solve the continuity equation models for the neutral atmosphere, ion and electron temperature, production, loss and diffusion rates must be assumed as well as models for both the neutral wind and the vertical E X B drift pattern. We have adapted the MSIS neutral atmosphere, (Hedin, et. al., 1977), the same production and diffusion rates as given by Anderson (1973), and the ion and electron temperatures presented in Anderson, et. al., (1987). For the photoionization rate at the top of the atmosphere, P , a value of $5.5 * 10^{-7}$ sec^{-1} represents solar cycle maximum conditions, while a value of 3.0×10^{-7} sec^{-1} represents solar minimum conditions. The loss rates are adopted from Torr and Torr (1979). Vertical drifts used were those measured by Fejer (1981) and Fejer, et. al. (1989) using the Jicamarca, Peru radar.

TEST CASES USED

Since the available vertical drifts were taken at Jicamarca, Peru, the model was run for a longitude corresponding to that of Jicamarca, 70^o west geographic longitude. The magnetic equator was assumed to be 11^o south of the geographic equator, and the magnetic declination was assumed to be zero, in order that any effects of off-meridian drifts and zonal winds would be minimal. The model was run for both average solar maximum, $F_{10.7} = 180$, and solar minimum conditions, $F_{10.7} = 75$. Equinoctial conditions were assumed to minimize the differential effects of solar heating in the two hemispheres. Quiet magnetic activity, $A_p = 10$, was also assumed.

The model calculates electron densities along magnetic field lines for L shells up to a specified maximum. To determine the effects of the vertical drifts independently, equivalent vertical electron density profiles were calculated as a function of latitude and time for each of the solar activity conditions studied. The parameters N_{max}, TEC, and h_{max} were determined for each latitude, for each case. To separate the effects of the vertical drift differences were computed between modeled results using various values of vertical drift, normalized by cases having zero drift. The results are then presented both in the form of percentage changes, and in absolute value of each of the three F region parameters studied, N_{max}, TEC, and h_{max}.

Figures 1A and 1B illustrate the vertical drifts used in our calculations. Those for solar minimum conditions were taken from Fejer, (1981), and those for solar maximum conditions were from Fejer, et. al., (1989). The average drifts were determined from the quiet day equinox curves in both Fejer papers. The high drifts illustrated in Figures 1A and 1B are simply a factor of 2 times the average drifts, with the exception of the evening sector of the solar maximum drift where no increase from the average drift measured by Fejer, et. al., (1989) was assumed.



a) Vertical drift, solar minimum equinox b) Vertical drift, solar maximum equinox

Figure 1.

In order to test the effects of varying the magnitude of the drift on the resultant peak electron density, N_{max}, and TEC the average drifts, illustrated in Figures 1A and 1B, were multiplied by 0.5, and 1.5, as well as the factor of 2.0 shown in the figures. The differences between the cases with various drift values minus the case with no drift were calculated, and the percentage ratios from the no-drift case were also calculated for those assumed extremes of drift. For solar minimum conditions there is only a slight post-sunset maximum in drift, while during solar maximum conditions, Figure 1B, a large post-sunset vertical drift can be seen. In parameterizing the effects of varying drifts during solar maximum conditions the post-sunset maximum drift was not changed. In the initial attempts at parameterizing solar maximum vertical drifts, changing the post-sunset drifts resulted in unrealistically high values of F region parameters in the anomaly regions; therefore, only the daytime values were modified by the factors of 0.5, 1.5 and 2.0 to determine the effects of varying drifts on the low latitude F region.

The NASA neutral wind model provides meridional winds as illustrated in Figure 2. Note that the daytime winds blow generally away from the geographic equator, while nighttime winds blow toward the equator, and are only generally symmetric about the geographic equator.

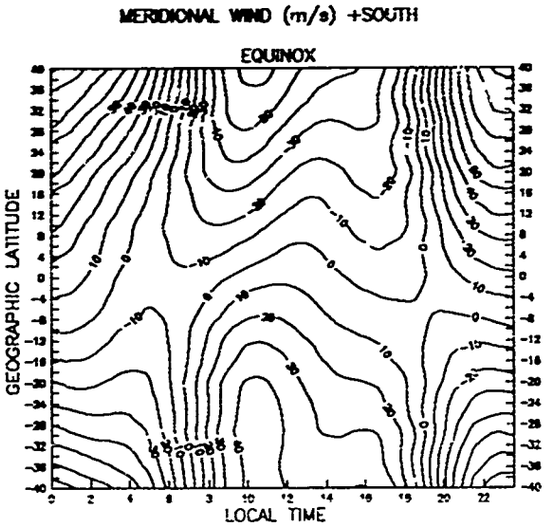


Figure 2. Meridional wind for equinox conditions.

MODEL RESULTS - SOLAR MINIMUM

The results of model calculations for N_{max} during solar minimum equinoctial conditions with no vertical drift, average drift, and with two times the average drift, are shown in Figures 3A, B and C. For the no drift conditions, Figure 3A, no evidence of an anomaly in N_{max}, is seen, though the peak values occur at the magnetic equator, rather than at the geographic equator, as would be expected from a simple solar control of the F2 region. This difference is due to the effects of diffusion of the F region ionization along magnetic field lines.

For average vertical drift conditions, shown in Figure 3B, the anomaly begins to develop by the late morning hours, but doesn't reach a peak until the evening hours of 19-20 local time. For the assumed vertical drift of twice the average, shown in Figure 3C, the anomaly has still larger values of N_{max}, and their displacements from the magnetic equator are greater. The effects of the anomaly are seen more vividly when the modeled results using average and high drifts have the zero drift values subtracted from them. Figure 4A illustrates the calculated values of N_{max} for average drift conditions minus the no drift calculations. Note that an excess peak electron density over the no drift case reaches a maximum of 10^6 el/cc at dip latitudes of approximately plus and minus 16^o, and that the excess peak electron density extends to over plus and minus 26^o dip latitude.

The difference of the model for the assumed high drift calculation minus the zero drift case is shown in Figure 4B. For a high drift day during solar minimum the effects of the drift are over $1.4*10^6$ el/cc at the anomaly peaks and the locations of the anomaly peaks have shifted to plus and minus 19^o dip latitude. Both the average drift and the high drift cases produce noticeable increases in peak electron density in the afternoon and evening time periods at dip latitudes extending up to over 25^o. The predawn and early morning hours have small negative changes in N_{max} for both the average and the high drift calculations, when compared with the no drift case. These negative values are likely due to the downward E X B drift which drives the ionization to lower heights where the loss is greater.

One surprising feature of the calculated effects of the anomaly is the relatively late time of the anomaly peak as compared with the time of peak drift for solar minimum conditions. In Figure 1A the peak drift occurs, during solar minimum conditions, at approximately 10-11 hours local time, yet the anomaly maxima, illustrated in Figures 3A and 3B occur, not with the expected 2 to 3 hour difference, but at approximately 20 hours local time, approximately a 10 hour delay. The effects of the anomaly do begin to show up in the difference contours, Figures 4A and 4B, much earlier than 20 hours, and are obvious as early as 12 hours local time. The cumulative effects of the continuous upward drift over the entire period from the early morning hours throughout the day produce the late evening peak in the anomaly in N_{max}, but, as will be shown, not in TEC.

The TEC for solar minimum conditions for the assumed case of no vertical drift is shown in Figure 5A. The anomaly, as expected, does not occur with no vertical drift. Figures 5B and 5C illustrate the computed TEC contours for the average drift and two times the average drift, respectively, for solar minimum conditions. For average drift the TEC exceeds the no drift case by more than a factor of two over a large time and latitude region, while for the assumed case of 2 times the average drift, the highest TEC anomaly values are almost a factor of 4 greater than those with no drift. The time of the maximum anomaly in TEC for the average drift calculation occurs at approximately 14-15 hours, significantly earlier than the 20 hours local time peak of N_{max} for the same vertical drifts.

For the assumed high drift case the anomaly in TEC, though later than that for the average drift case, still occurs significantly earlier than the maximum in N_{max}. However, when the difference of the TEC is computed from the model with average drifts minus the no drift case, as illustrated in Figure 6A, the peak difference in the anomaly time shifts to 20 hours local time. Similarly for the 2 times average drift case minus the no drift case, shown in Figure 6B, the anomaly peak occurs later, though not as late as for the average drift. The differences in the times of the peak anomaly in TEC, for the cases with drift minus that of no drift, are due to the relatively rapid decay of TEC in the normal anomaly latitude region for the no drift case. Thus, the time of the TEC anomaly peak is shifted towards later hours.

Also, it is worth noting that the TEC for the case of average drift minus the no drift, there is no region of negative values (Figure 6B), and for the high drift minus the no drift case there is only a small region near the equator, centered near dawn, where the TEC is less during high vertical drift conditions than it would be with no vertical drift. Again, this is likely due to the downward drift during the nighttime hours.

Similar calculations were made for H_{max}, and the results for no drift, average drift and twice the average drift are illustrated in Figures 7A, B and C, respectively. For the no drift case there is not much latitude dependence of H_{max} in the late afternoon and evening time period; that is, after 1800 hours the H_{max} values range from approximately 300 to 325 km. For the average drift conditions there is a equatorial peak of over 375 km in H_{max} at approximately 19 hours local time, which increases to over 400 km when 2 times the average drifts are used, Figure 7C.

When the contours of differences in H_{max} from the average drift case and the 2 times average case, minus the no drift case, are computed the results, plotted in Figures 8A and 8B respectively, show increased heights in the early afternoon time. These height increases occur at latitudes at, or slightly lower, and certainly at much earlier times, than the latitudes where the peak anomalies in N_{max} and TEC occur. Lower values of H_{max} occur in the early morning hours, again due to the downward vertical drifts during those times.

SOLAR MAXIMUM CONDITIONS

The model was run for somewhat higher than average solar maximum conditions, $F_{10.7}$ = 180 flux units, again for the cases of no vertical drift, the average vertical drifts shown in Figure 1B and two times the average drift. For the case

or two times the average drift, indicated by the curve with open circles in Figure 1B, the drifts in the evening time period were not increased from the average drifts of Fejer, et. al. (1989) because they gave unrealistically high values of the parameters N_{max} and TEC. N_{max} for the no drift, average drift and twice the average drift model calculations are illustrated in Figures 9A, B and C, respectively. For average drift conditions the value of N_{max} at the anomaly peak regions is greater than $5*10^6$ el/cc, which corresponds to a critical frequency, foF2, of 20 MegaHertz. The plots of the difference of the model calculations with average and high drifts, minus the no drift case are shown in Figures 10A and B, respectively. For the high drift minus zero drift case, Figure 10B, the effects of the anomaly in N_{max} can be seen to extend to over plus and minus 38 degrees dip latitude, the limit of our calculations. For average drift minus zero drift conditions, Figure 10A, there is a significant region from approximately plus and minus 8^o dip latitude where N_{max} is lower than the average drift case, but up to at least plus and minus 26^o dip latitude N_{max} is larger than the no drift case.

TEC shows no anomaly for the case of no drift, as expected, Figure 11A, and an anomaly which increases both in magnitude and dip latitude as the drift is increased from average values, Figure 11B, to two times the average values, Figure 11C. The contours of differences in the modeled values from the average minus the no drift case, Figure 12A, show lower values than the no drift case during the post midnight to 10 hour local time period over a fairly large range of dip latitudes from at least plus and minus 15^o. Similar negative effects of drifts are seen during the late night hours for the high drift case minus the zero drift, Figure 12B, but the negative effects continue until well after local noon near the dip equator. Note that the effects of the assumed strong drifts make a large positive contribution of additional TEC up to the limit of our calculations at 38 degrees dip latitude. Also, as expected, the higher drift makes both the magnitude and the dip latitude of the anomaly greater.

The effects of drifts on H_{max} during solar maximum conditions are to increase the height over the dip equator, by well over 200 km, from a nominal height of approximately 400 km during the evening hours, peaking at approximately 20 hours local time. Significant values of increased H_{max} of at least 50 km over the no drift case, occur over a latitude range of plus and minus 15 degrees for several hours from local noon until approximately 21 hours local time, following the upward drift.

CONCLUSIONS

Theoretical calculations of the location and strength of the anomaly in the F2 region, as shown in N_{max}, TEC and H_{max} all illustrate effects from the equatorial anomaly which extend, in extreme cases, to over plus and minus 38^o in dip latitude. When one considers that 50% of the earth's surface area is contained within plus and minus 30^o latitude, (it makes no difference whether geographic latitude of dip latitude is used as the measure), the equatorial anomaly has very large effects indeed. The name "fountain effect" was aptly chosen, Martyn, (1955), because on days when the drifts do not occur, (e.g. counter-electrojet days), and the fountain effectively is turned off, as determined by the Anderson (1973) model, the F region is much lower in both its peak electron density and TEC than during the normal fountain days. On days when the fountain is strong, its contribution to the total number of additional electrons on a world-wide basis is very large.

The theoretical model can be used with either observed vertical drifts, or with estimates of the drift, to make comparisons against observed values of N_{max}, TEC and H_{max} made over a latitude region sufficient to clearly delineate the location and strength of the anomaly in those parameters. Unfortunately, direct measurements of vertical drifts have only been made using the Jicamarca incoherent scatter radar in the American sector, but no corresponding latitudinal contours of the F2 region parameters are available from the American longitudes. In the future model comparisons will be made against available F2 region latitude contours for other longitude sectors. The potential use of tomographical techniques to determine two dimensional electron density contours from a few stations making measurements of TEC along one meridian (Asten, et. al., 1988) would make a significant contribution to the understanding of the day-to-day variability effects of the equatorial anomaly region.

REFERENCES

Anderson, D. N., "A Theoretical Study of the Ionospheric F-Region Equatorial Anomaly, II. Results in the American and Asian Sectors", Planet. Space Sci., V 21, pp 421-442, 1973.

Anderson, D. N., M. Mendillo and B. Herniter, "A Semi-Empirical Low-Latitude Ionospheric Model", Radio Science, V 22., No. 2, pp 292-306, March-April 1987.

Appleton, E. V., "Two Anomalies in the Ionosphere" , Nature, V. 157, p. 691, 1946.

Austen, J. R., S. J. Franke, and C. H. Liu, "Ionospheric Imaging using computerized Tomography", Radio Science, V. 23, No. 3, pp 299-307, May-June, 1988.

Deshpande, M. R., R. G. Rastogi, H. O. Vats, J. A. Klobuchar, G. Sethia, A. R. Jain, B. S. Subbarao, V. M. Patwari, A. V. Janve, R. K. Rai, Malkiat Singh, H. S. Gurm, B. S. Murthy, "Effects of Electrojet on the Total Electron Content of the Ionosphere over the Indian Subcontinent", Nature, V. 267, No. 5612, pp. 599-600, June 16, 1977.

Fejer, B. G., "The Equatorial Ionospheric electric Fields, A Review", J. Atmos. Terr. Phys., V. 43, No. 5/6, pp 377-386, 1981.

Fejer, B. G., E. R. de Paula, I. S. Batista, E. Bonelli, and R. F. Woodman, "Equatorial F Region Vertical Plasma Drifts During Solar Maxima", J. Geophys. Res., V. 94, No. A9, pp12,049-12054, Sept. 1, 1989.

Hanson, W. B., and R. J. Moffett, "Ionization Transport Effects in the Equatorial F Region", J. Geophys. Res., V. 71, No. 23, Dec. 1, 1966.

Hedin, A. E., C. A. Reber, G. P. Newton, J. W. Spencer, H. C. Brinton, H. G. Mayr, and W. E. Potter, "A Global Thermospheric Model based upon Spectrometer and Incoherent Scatter Data, MSIS 2, Composition", J. Geophys. Res. V. 82, p 2148, 1977.

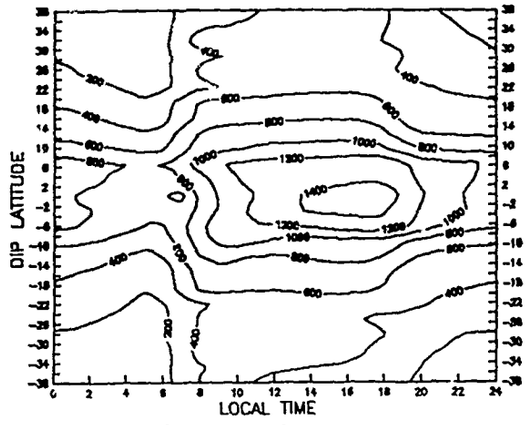
Huang, Yinn-Nien, Kang Cheng, and Sen-Wen Chen, "On the Equatorial Anomaly of the Ionospheric Total Electron Content Near the Northern Anomaly Crest Region", J. Geophys. Res., V. 94, No. A10, pp 13,515-13,525, Oct. 1, 1989.

Moffett, R. J., "The Equatorial Anomaly in the Electron Distribution of the Terrestrial F-Region", Fundam. Cosmic Phys. V. 4, pp 313-391, 1979.

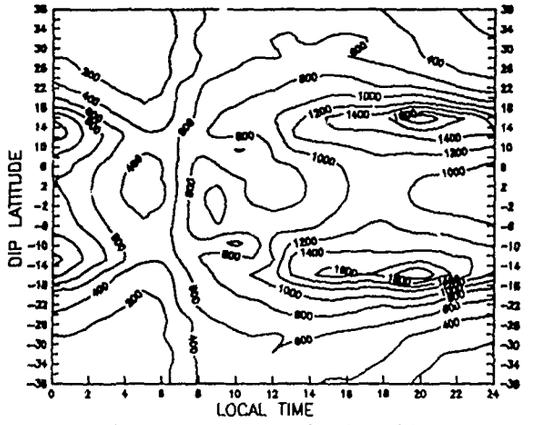
Rastogi, R. G., "The Diurnal Development of the Anomalous Equatorial Belt in the F2 Region of the Ionosphere", J. Geophys. Res., V. 64, No. 7, pp 727-732, July, 1959.

Rush, C. M. and A. D. Richmond, "The Relationship between the Structure of the Equatorial Anomaly and the Strength of the Equatorial Electrojet", J. Atmos. Terr. Phys., V. 35, pp 1171-1180, 1973.

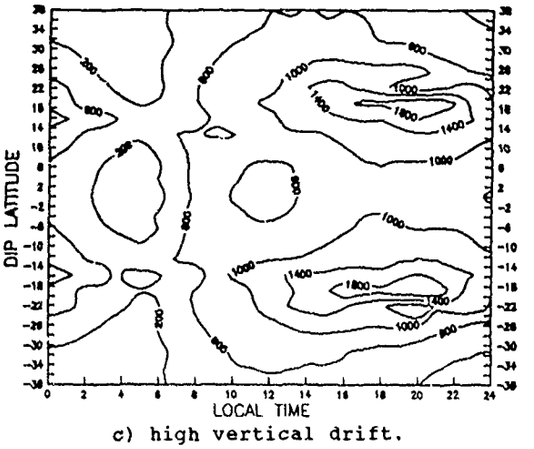
Torr, M. R., and D. G. Torr, "Chemistry of the Thermosphere and Ionosphere", J. Atmos. Terr. Phys. V 41, pp 797-839, 1979.



a) no vertical drift.

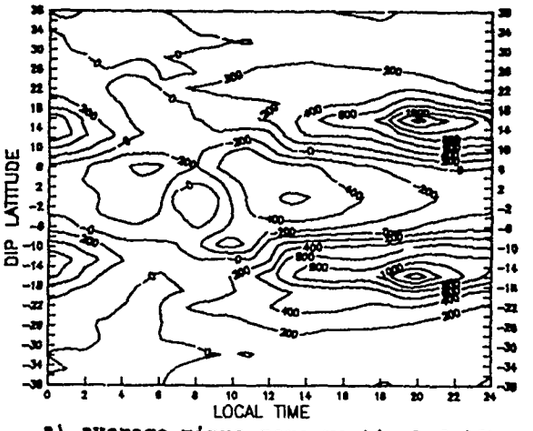


b) average vertical drift.

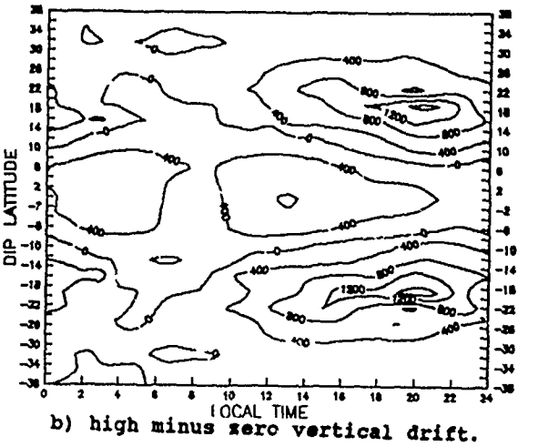


c) high vertical drift.

Figure 3. N_{max} for equinox solar minimum

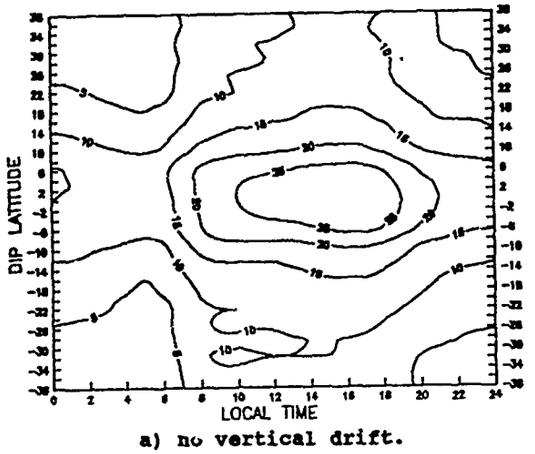


a) average minus zero vertical drift.

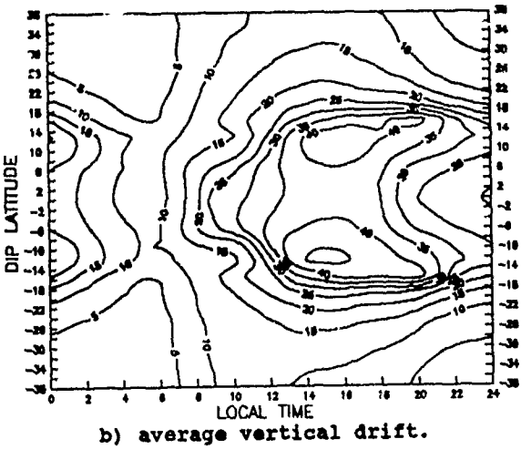


b) high minus zero vertical drift.

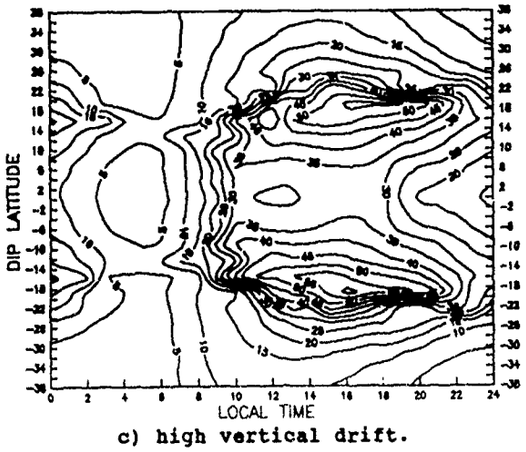
Figure 4. N_{max} difference of



a) no vertical drift.

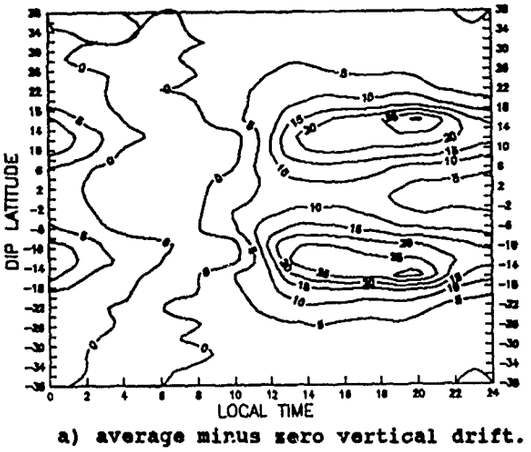


b) average vertical drift.

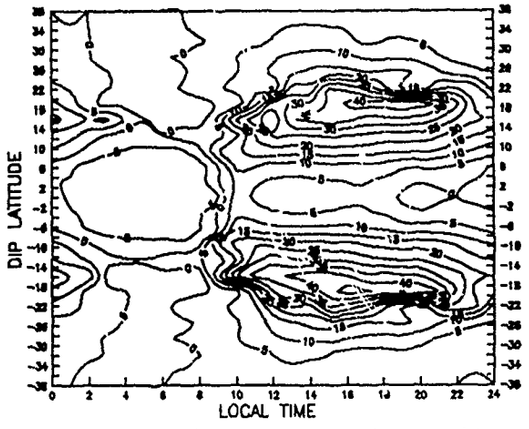


c) high vertical drift.

Figure 5. TEC for equinox solar minimum

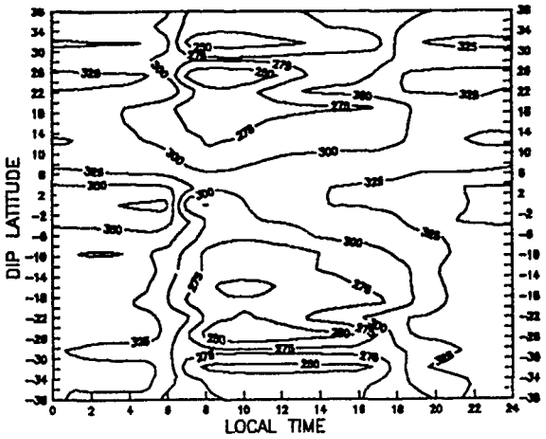


a) average minus zero vertical drift.

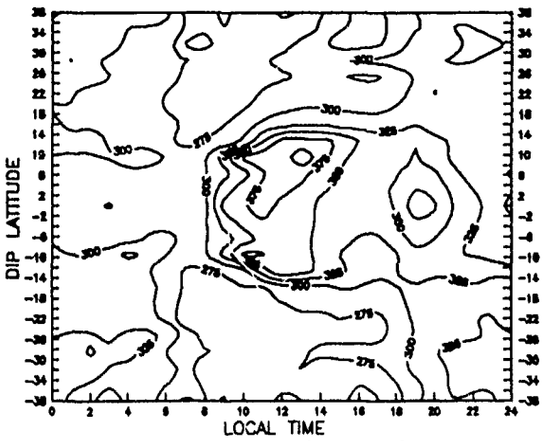


b) high minus zero vertical drift.

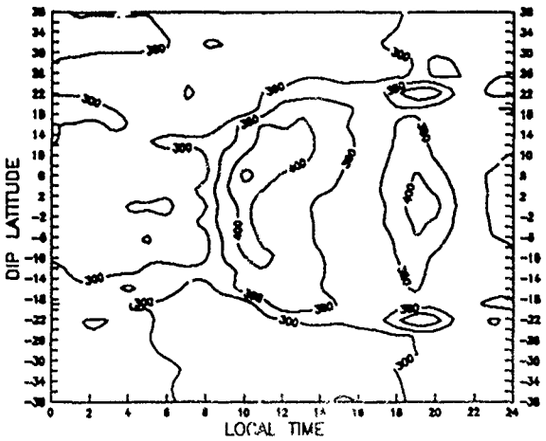
Figure 6. TEC difference



a) no vertical drift.

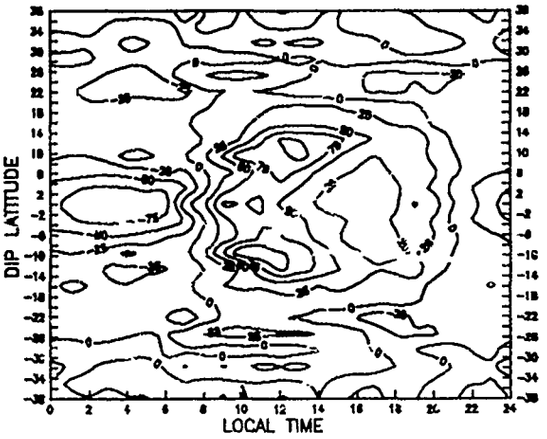


b) average vertical drift.

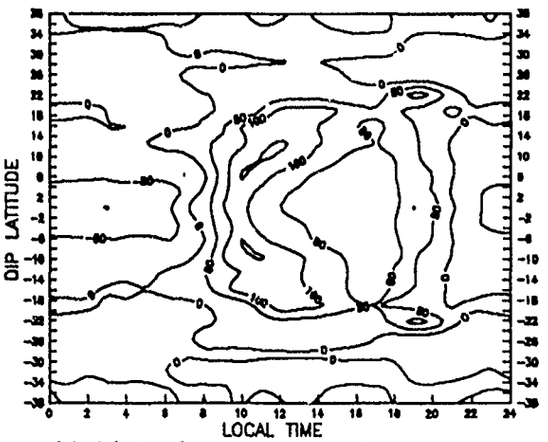


c) high vertical drift.

Figure 7. H_{max} for equinox solar minimum

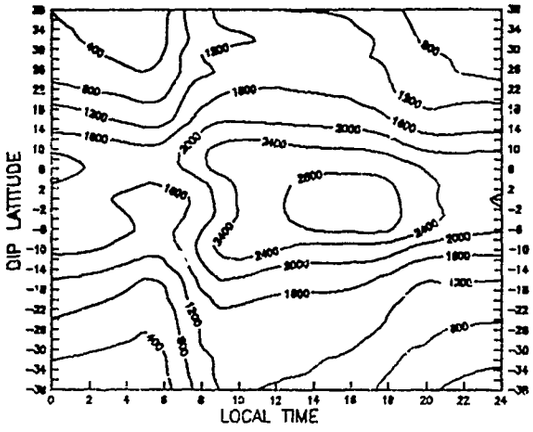


a) average minus zero vertical drift.

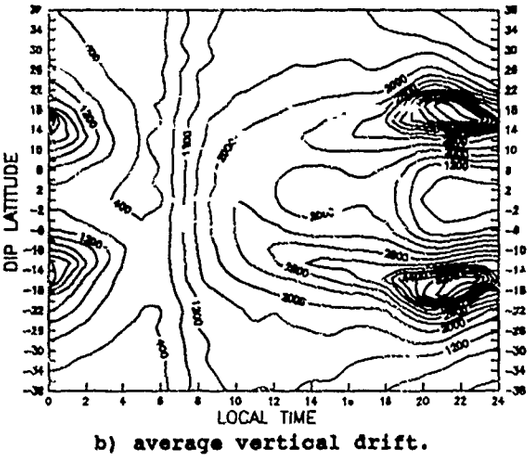


b) high minus zero vertical drift.

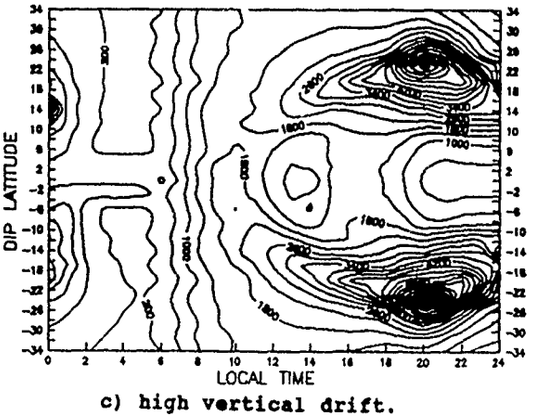
Figure 8. H_{max} difference of



a) no vertical drift.

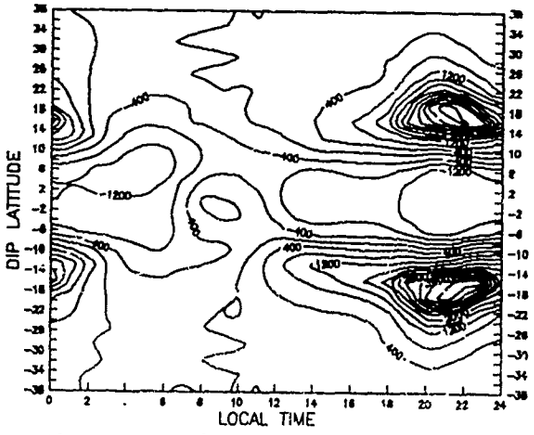


b) average vertical drift.

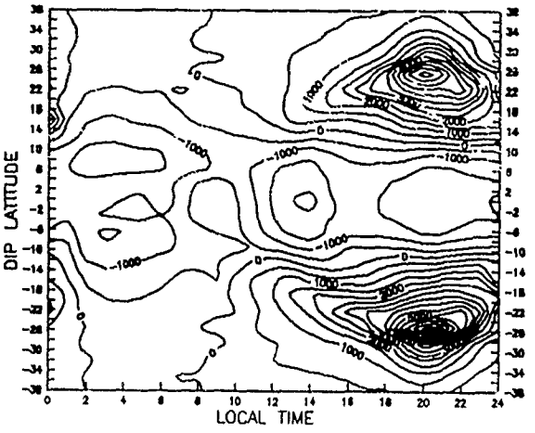


c) high vertical drift.

Figure 9. N_{max} for equinox solar maximum

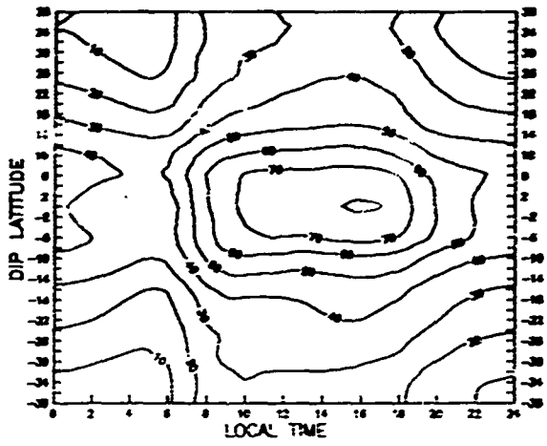


a) average minus zero vertical drift.

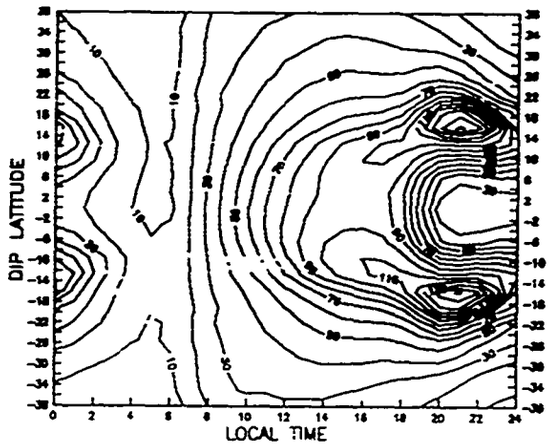


b) high minus zero vertical drift.

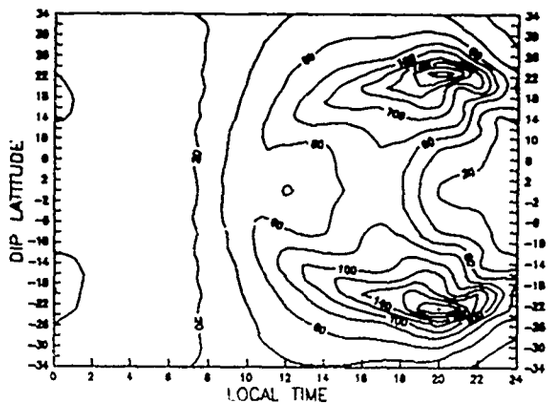
Figure 10. N_{max} difference of



a) no vertical drift.

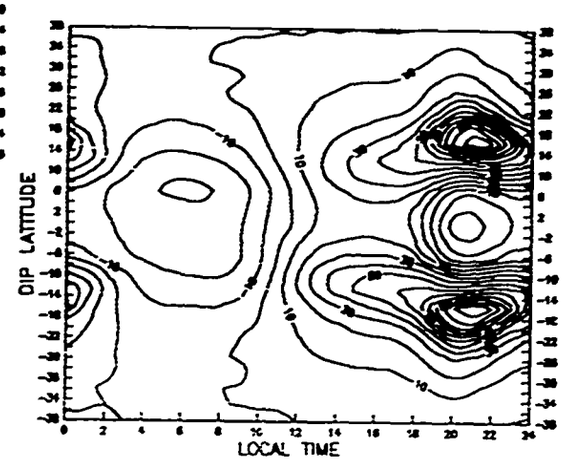


b) average vertical drift.

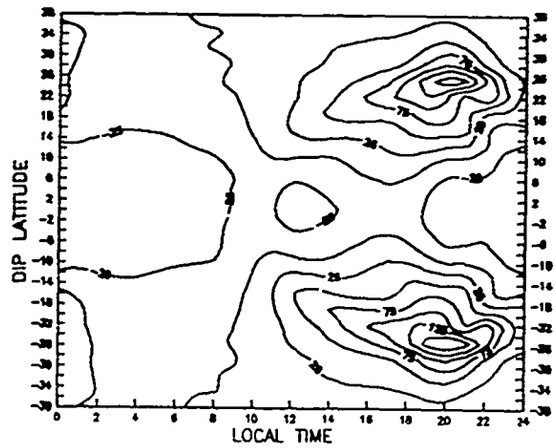


c) high vertical drift.

Figure 11. TEC for equinox solar maximum for



a) average minus zero vertical drift.



b) high minus zero vertical drift.

Figure 12. TEC difference of

AD-P006 305

A COMPARISON OF IONOSPHERIC SLAB THICKNESS AT WIDELY SPACED STATIONS IN MIDDLE LATITUDES AND LOW LATITUDES

K. Davies and X. M. Liu
NOAA/ERL/SEL
Boulder, CO 80303

1. Introduction

Ionospheric slab thickness τ is a useful ionospheric parameter from the point of view of satellite-to-ground radio communications. From a knowledge of the slab thickness and the F_2 layer penetration frequency ($f_o F_2$), we obtain the total electron content in a vertical $1m^2$ column which, in turn, determines: ionospheric time delay, refraction, and rotation of the electric field.

$$\text{TEC} = 1.24 \times 10^{10} (f_o F)^2 \tau \ el.m^{-2} . \tag{1}$$

Various computer programs are available for predicting the critical frequency (e.g., IONCAP), so, with a knowledge of slab thickness, predictions of total electron content are possible (Klobuchar 1987).

With a view to improving knowledge of the geographical, seasonal, and sunspot variations of slab thickness, data on TEC and $f_o F_2$ have been analyzed from a number of locations around the globe.

2. The data base

The data on total electron content (TEC) used here have been obtained mostly from the Faraday rotation of radio beacon signals from geostationary satellites at 17 locations as listed in Table 1. In most cases the total electron content data were obtained from tabulations of the station data, but in a few cases the data were derived from published Fourier coefficients (e.g., Poulter and Hargreaves 1981, Feichter et al. 1989, Bhuyan and Tyagi 1986). In a few cases TEC data are from measurements of differential phase on signals from orbiting satellites (Feichter et al. 1989). The $f_o F_2$ (for a particular station) was obtained from ionosonde measurements at a station closest to the sub-ionosphere point, i.e., the point directly below the intersection of the satellite-to-ground ray path and the 400-km height. Otherwise, $f_o F_2$ data at the station itself was used. In a few cases, where no suitable measured data were available, $f_o F_2$ was calculated from the IONCAP program using the 12-month smoothed sunspot index at the time of observation.

TEC data are subject to errors caused by: (1) uncertainty in the conversion of Faraday rotation measurements to TEC, (2) error in the zero calibration, (3) an uncertainty in the cycle (or ramp) number. These concerns will be discussed below.

The conversion of Faraday rotation to TEC, or Faraday, content N_F is

$$\overline{M} \, N_F = \Omega , \tag{2}$$

where $\overline{M} = B_0 \cos\theta \sec\chi$. Here B_0 is the strength of the earth's magnetic field (in teslas), θ is the angle between the ray and the magnetic field, and χ is the angle between the ray and the vertical. Following Titheridge (1972), the value of $\overline{M}$ has been taken as the value of $B_0 \cos\chi \sec\theta$ at a height of 420 km on the ray path. While the most appropriate value of $\overline{M}$ varies with electron density profile and, therefore, with time and location, the value at 400 km is convenient and accurate to within ±3%

91-09695

Table 1. List of stations for which data on total electron content were available; F = 10.7-cm solar flux.

Station	Latitude(°N)	Longitude(°E)	Dipole Lat°	F	Year
MANILA	14.6	122.0	3.6	70	1986
HONG KONG	21.0	116.0	11.0	70-155	1965,68
TAIWAN	23.0	122.0	13.8	70	1986
DELHI	28.6	77.2	18.9	86-155	1974,68
HAWAII	20.8	204.0	21.2	70	19-6
TOKYO	35.7	39.5	25.7	83-200	1977-79
JAMAICA	18.0	283.2	29.2	150	1970
RAMEY	17.0	288.0	29.8	70-190	1986,81
KENNEDY	26.0	276.0	39.5	70	1986
SHEMYA	46.0	186.0	40.6	70-190	1986,81
GRAZ	47.0	15.5	46.7	75-200	1975,80
BOULDER	40.0	255.0	48.9	80	1974-75
LANCASTER	50.1	4.6	51.7	75	1975-76
LINDAU	51.7	10.0	52.1	72-200	1975,80
SAG. HILL	42.6	288.5	53.9	72-200	1975-86
ABSTWYTH	52.4	4.1	55.6	70-75	1975,86
GOOSE BAY	47.0	286.0	64.4	70-190	1981,86

The zero calibration refers to the direction of the electric field of the received radio wave in the absence of the ionosphere. Unfortunately, only in the case of a very few beacons (e.g., ATS 6, Davies 1980) has this direction been accurately known. Empirically, it can be estimated from measurements of $f_o F_2$ and an assumed value of slab thickness. This can be done only at night when the rotation is less than one half cycle.

The half-cycle ambiguity introduces the greatest uncertainty into these measurements. The half-cycle number has to be determined at night when the total electron content is small and, hence, (it is hoped) the half-cycle ambiguity is zero. Thereafter, the number of cycles has to be carefully tracked as the Faraday rotation increases and decreases during the day. However even at night, and especially during summer nights at high sunspot numbers, there could be two or three half cycles, and this error is reflected in the level of the diurnal curve. With some equipment, the cycle number changes from one night to the next without a discontinuity.

In the present study it is assumed that the above uncertainties have not materially affected the data.

3. Diurnal and seasonal variations

Because of the relatively large variations in the diurnal curves from one month to another, the data have been grouped into four seasons as follows:

Equinox:	March, April, September, October
Summer:	May, June, July, August
Winter:	November, December, January, February

Some curves showing diurnal variations at widely separated locations are shown in Figure 1 for a low value of the 10.7-cm solar radio flux and for the three seasons listed above. In middle latitudes at night the slab thickness is around 260 km (Aberystwyth, U.K.) and 300 km (Sagamore Hill, U.S.A.). Also, in middle latitudes the winter slab thickness decreases from night to day, whereas during summer it increases from night to day. There is a predawn maximum, around 0500 LT, that is more pronounced in low latitudes than in middle latitudes. In middle latitudes the midday values are greater in summer than in winter.

Contours of the diurnal variations of slab thickness for different months are shown in Figures 2 and 3 for a sunspot minimum year and a sunspot maximum year, respectively, for a middle-latitude site. These figures confirm the statements made above. Note that for high values of solar flux the predawn increase is more conspicuous

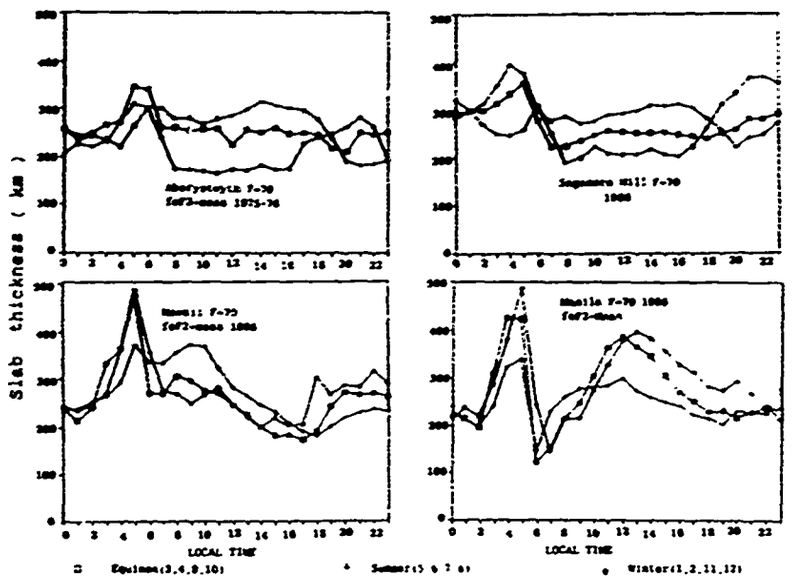


Figure 1. Hourly mean values of the slab thickness for equinox, summer, and winter during solar cycle minimum.

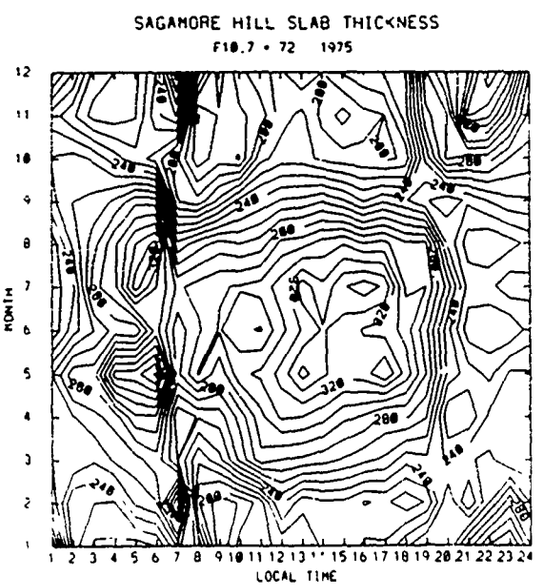


Figure 2. Contours of slab thickness (in km) at Sagamore Hill during a solar-cycle-minimum year.

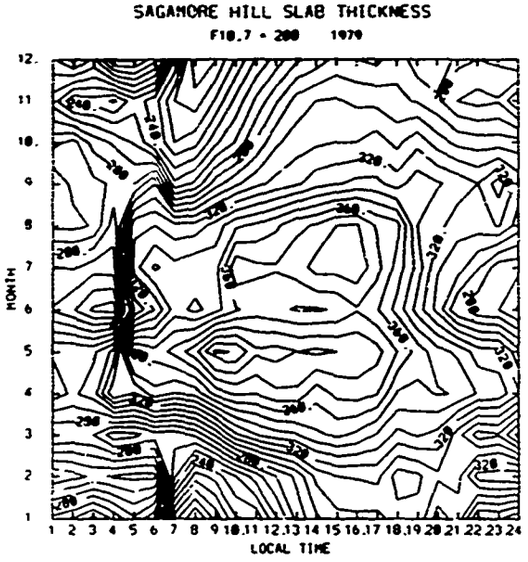


Figure 3. Contours of slab thickness (in km) at Sagamore Hill during a solar-cycle-maximum year.

in summer than in winter. Care must be taken in detailed interpretation of the predawn increase because it is sensitive to possible errors, discussed in Section 2.

At certain locations the predawn increase is more pronounced in some months than in others, as seen in Hawaii, in Figure 4, where the maxima are more evident around March and October/November. The predawn increase is caused by a decrease in the maximum density rather than by an increase of total electron content, as is illustrated in Figure 5. That is, the ionization near the layer peak erodes relatively more than at higher altitudes. The predawn increase is more conspicuous in low latitudes than in middle latitudes (Ma and Long 1988, Rao et al. 1988).

The magnitude of the predawn increase is defined as the excess of the peak value over the value to be expected in the absence of any increase, as depicted in Figure 6. The expected value is obtained by linear interpolation between the hourly values on either side of the increase. Some data for Sagamore Hill on the seasonal and local time occurrence, durations of the increases are shown in Figure 6. In general there is a maximum of occurrences in northern winter, and a minimum in summer. The maximum daily value occurs around 0400 LT at Sagamore Hill, with a duration between 3 and 4 hours.

4. Geographical and/or geomagnetic variations

Figure 7 shows plots of the noon slab thickness versus dipole latitude for the three seasons, summer, equinox, and winter, and for years of high and low solar radio fluxes. While there is some evidence to suggest a small decrease in slab thickness with dipole latitude, Φ, in winter when the solar flux is low for the northern hemisphere, there appears to be no significant dependence of τ versus Φ on either equinox or summer. Thus there is little or no dependence on dipole latitudes, to within the scatter of the data points. The low values of slab thickness for Hong Kong are probably erroneous.

5. Solar cycle variations

The long-term variation of slab thickness at Sagamore Hill is shown in Figure 8 for noon and midnight. The average of all the monthly noon (11, 12, 13 LT) values from January 1968 through December 1986 is 282 km, which does not differ significantly from the corresponding midnight (23, 00, 01 LT) value of of 279 km. The noon values show a marked seasonal variation, with high summer values and low winter values.

In some ionospheric prediction systems the 12-month smoothed solar index is used. The smoothed noon values of slab thickness τ_{12} are related to the smoothed $F10.7$ (F) approximately as

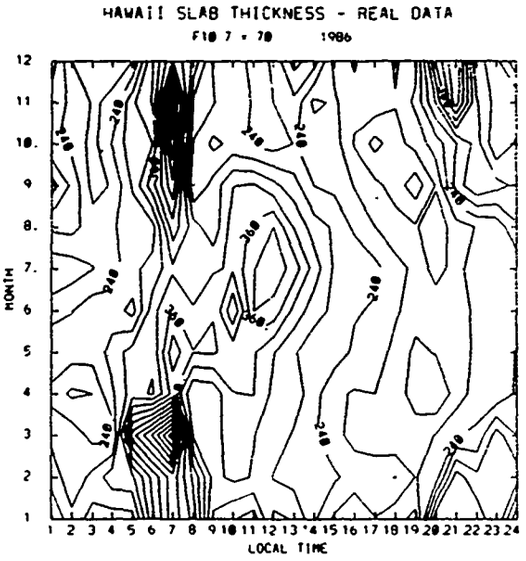


Figure 4. Contours of slab thickness at Hawaii during a solar-cycle-minimum year.

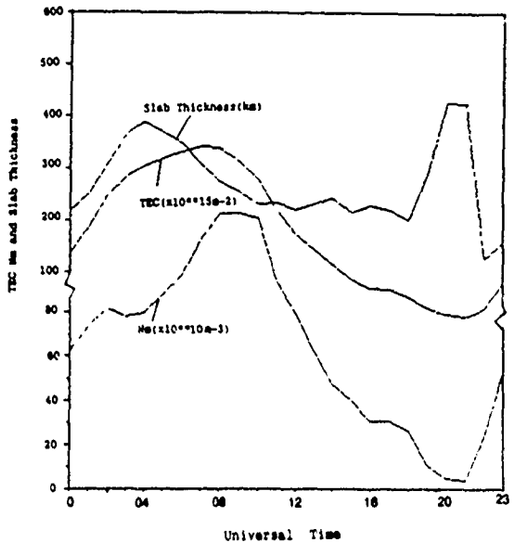


Figure 5. Seasonally averaged hourly values of slab thickness, at Manila, of total electron content, maximum electron density, and slab thickness. This figure shows that the predawn increase in slab thickness is caused by low values of maximum density and not by increase of total electron content.

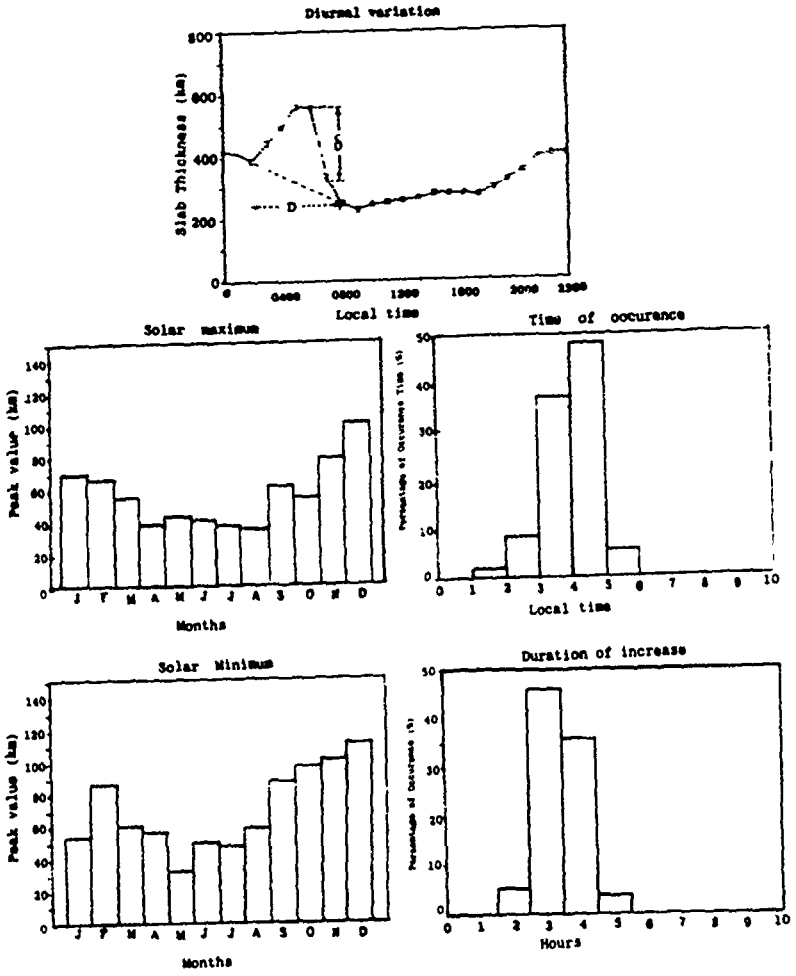


Figure 6. Variation of the magnitude of the predawn increase in slab thickness at Sagamore Hill, with local time and season, time of occurrence, and duration.

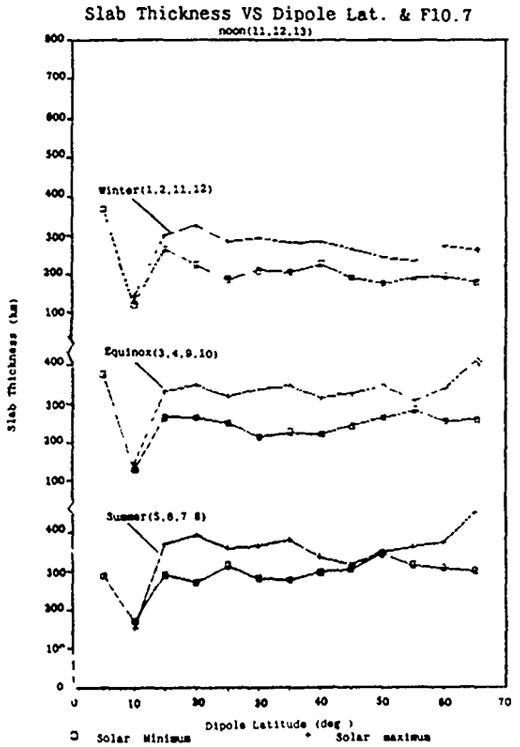


Figure 7. Variation of the noon slab thickness with dipole latitude.

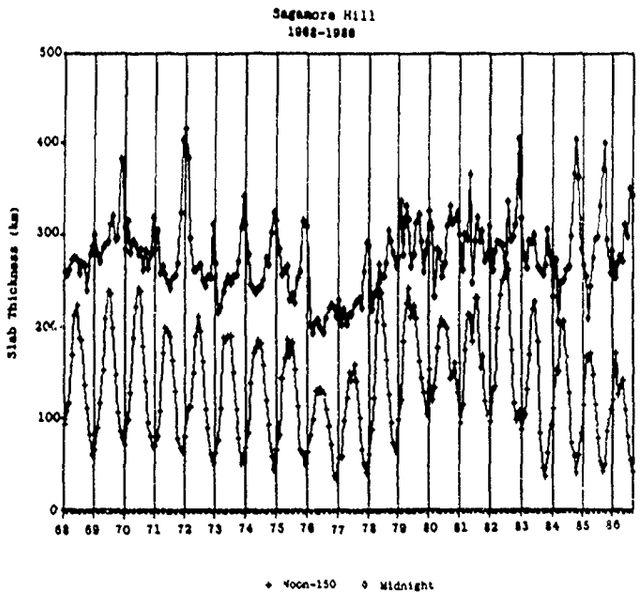


Figure 8. Long-term variations of noon (11, 12, 13 LT) and midnight (12, 00, 01 LT) slab thickness at Sagamore Hill. Note that the seasonal variations at noon and midnight tend to be in antiphase. The noon values are shifted down by 150 to separate the data sets.

$$\overline{\tau}_{12} = 218 + 0.5 * F \quad (3)$$

for $F > 70$, and, for midnight,

$$\overline{\tau}_{0} = 238 + 0.3\ F \quad (4)$$

Hence, for values of F near 100, the smoothed values for noon and midnight are nearly equal.

To examine the long-term variation for different seasons at Sagamore Hill, the noon and midnight data were binned as follows: winter (November, December, January, February), equinox (March, April, September, October), and summer (May, June, July, August). The results are presented in Table 2 for the years 1967–1986. Near noon the slab thickness increases with smoothed solar fluxes during all seasons, whereas near midnight in winter there is a slight decrease of slab thickness with smoothed solar flux.

Somewhat similar data for for 11 stations, for different local times, are summarized in Table 3. Different stations have data for different years, which partly accounts for the spread in the coefficients. Also, the 10.6-cm flux are the averages of the particular months used rather than the 12-month smoothed values.

Table 2. Dependence of noon and midnight slab thickness on 12-month smoothed solar 10.7-cm radio flux (F) and season at Sagamore Hill, 1967 through 1986.

$\tau = A + BF$

	Slab thickness (km)					
Season	Equinox		Summer		Winter	
	A	B	A	B	A	B
Midnight	222	0.38	197	0.52	303	−0.018
Noon	219	0.46	262	0.61	164	0.50

Table 3. Variations of slab thickness (in km) with 10.7-cm solar flux (for the same months) for different local times.

	LT	0000		0400		0800		1200		1600		2000	
Station	Season	a	b	a	b	a	b	a	b	a	b	a	b
	Equinox	100	−0.0050	306	−0 0075	115	−0 0039	129	0 0013	81	0 0012	42	0 0084
Hong Kong	Summer	259	−0 0046	274	−0 0032	161	−0.0019	169	−0 0012	133	−0 0001	144	−0 0015
(69.86)	Winter	158	−0.0068	237	−0 0086	86	−0 0083	124	0 0018	62	0 0098	56	−0 0019
	Equinox	404	−0 0014	315	0 0029	189	0 0011	145	0 0092	280	0 0069	469	−0 0013
Delhi	Summer	177	0 0009	205	0 0017	130	0 0078	153	0 0131	242	0 0056	273	0 0042
(68.74)	Winter	307	0 0255	176	0.0295	105	0 0068	137	0.0108	282	0 0086	487	0.0055
	Equinox	106	0 0092	104	0.0085	173	0 0033	180	0 0047	129	0 0076	100	0 0122
Tokyo	Summer	121	0.0062	135	0.0081	192	0 0048	279	0 0015	223	0.0027	130	0 0069
(77.79)	Winter	70	0 0158	91	0 0100	132	0.0036	106	0 0095	159	0.0049	103	0.0118
	Equinox	182	0.0078	252	0 0040	136	0 0062	138	0 0073	150	0 0031	155	0 0047
Ramey	Summer	265	0 0012	287	−0 0001	249	0 0043	220	0 0023	250	0 0017	218	0 0039
(81.36)	Winter	259	0 0042	224	0 0078	188	0 0019	157	0 0053	143	0 0085	412	0.0003
	Equinox	355	0 0004	403	−0 0015	179	0 0036	143	0 0072	121	0 0093	252	0 0018
Kennedy	Summer	330	−0.0024	381	−0 0025	212	0 0055	242	0 0033	214	0 0035	240	0 0015
(81.86)	Winter	381	0 0008	402	−0 0004	196	0 0012	187	0 0033	201	0 0037	379	−0 0008
	Equinox	292	−0 0024	229	−0 0094	228	0 0013	240	0 0008	201	0 0011	274	−0 0003
Shemya	Summer	246	−0 0045	271	−0 0052	324	−0 0007	304	−0 0013	312	−0 0019	269	−0 0009
(81.86)	Winter	362	0 0018	295	0 0054	153	0 0045	195	0 0020	241	0 0012	369	−0 0003
	Equinox	168	0.0042	195	0 0047	140	0 0060	172	0.0049	160	0 0047	111	0 0116
Graz	Summer	182	0 0021	222	0 0020	199	0.0028	234	0 0025	271	0 0018	221	0 0032
(75.80)	Winter	162	0 0091	168	0 0084	138	0 0052	122	0 0049	150	0.0030	211	0 0045
	Equinox	156	−0 0048	208	−0 0023	218	0 0014	224	0.0028	192	0 0039	214	0.0014
Boulder	Summer	177	−0 0006	181	−0 0006	200	0.0063	199	0 0063	226	0 0038	143	0 0041
(87.89)	Winter	238	−0 0016	234	0 0007	188	0 0008	183	0 0033	164	0 0053	360	−0 0008
	Equinox	315	0 0012	330	0 0024	206	0 0035	266	0 0014	248	0 0016	166	0 0065
Lindau	Summer	210	0 0018	247	0 0023	212	0 0028	263	0 0021	294	0 0014	142	0 0053
(75.80)	Winter	218	0.0082	211	0 0090	178	0 0040	157	0.0025	210	0 0013	280	0 0038
Sagamore	Equinox	202	0 0022	218	0 0015	194	0 0017	234	0 0043	200	0 0032	185	0 0038
Hill	Summer	195	0 0022	253	−0 0005	229	0 0024	262	0 0019	296	0 0013	161	0 0049
(75-79)	Winter	186	0 0030	143	0 0053	141	0 0042	154	0 0039	150	0 0047	243	0 0011
	Equinox	145	0 0121	299	0 0014	125	0 0100	165	0.0071	176	0 0068	188	0 0074
Goose Bay	Summer	180	0 0103	184	0 0089	188	0 0072	209	0 0084	260	0 0043	215	0 0042
(81.86)	Winter	181	0 0065	238	0 0019	160	0 0092	134	0 0054	192	0 0032	238	0 0045

Diurnal curves for the various stations listed in Table 1 have been subjected to Fourier analyses using

$$\tau = A_0 + \Sigma A_n \sin(2\pi nH/24) + \Sigma B_n \cos(2\pi nH/24), \quad (5)$$

where H is local time in hours and n is the harmonic number.

The coefficients A_0, A_n, B_n (for $n = 1$ through 6) show no perceptible variation with geographic or geomagnetic location. Coefficients A_0 for all stations studied depends on the solar flux (see Figure 9) and season:

$$A_0 \text{ (summer)} = 225 + 0.66F, \quad (6a)$$

$$A_0 \text{ (equinox)} = 192 + 0.74F, \quad (6b)$$

$$A_0 \text{ (winter)} = 183 + 0.86F, \quad (6c)$$

for $F > 70$.

The higher harmonics ($n \geq 1$) showed no appreciable dependence on solar flux.

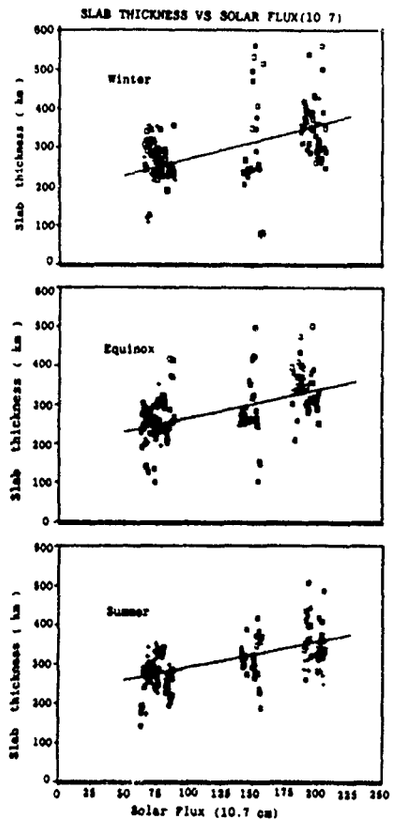


Figure 9. Variation of the zero-order Fourier term A_0 of the seasonally averaged slab thickness with 10.7-cm solar fluxes.

6. Discussion

This study was initiated with a view to mapping the global total electron content. Because of the paucity of TEC data relative to the data available on N_m (i.e., $f_o F_2$) it was decided that the appropriate course was to obtain $f_o F_2$ (i.e., $N_m F_2$) by an existing mapping technique (e.g., IONCAP, see Rush 1989) and to combine these with slab thickness information to obtain the total electron content.

In addition to the problems associated with the measurement of TEC listed in Section 2, another difficulty arises in the determination of slab thickness, namely the geographical separation between the satellite-to-ground ray path in the ionosphere and the location of the ionosonde used to measure $N_m F_2$. Since in the definition of slab thickness $N_m F_2$ appears in the denominator, the effect of any inaccuracy in $N_m F_2$ is magnified at night when $N_m F_2$ is small. This aspect must be borne in mind when attempting a detailed interpretation of slab thickness data.

The more-important conclusions to be drawn from this study are:

(1) It is impossible to discern any geographical (or geomagnetic) structure to within the accuracy (scatter) of the data.

(2) Attention has been concentrated on one of the most reliable middle-latitude stations (Sagamore Hill). The slab thicknesses at Sagamore Hill have dependencies on smoothed solar radio fluxes. The midnight and noon values increase approximately linearly with 12-month smoothed F 10.7 in equinox and summer and also the winter noon values. The winter midnight values are, on the whole, essentially independent of the 12-month smoothed fluxes.

(3) On a global scale the average daily values (as represented by the A_0 term of the Fourier expansion) increase with F 10.7 (see equation 6).

(4) The diurnal variation exhibits a pronounced predawn increase in slab thickness. The magnitude of this increase is larger in low latitudes than in middle latitudes; it is caused by low values of $N_m F_2$ rather than by an increase in total electron content. Thus, at night, the electron density profile appears to erode relatively more near the peak of electron density than at other heights. Rishbeth (1989, page 91) suggests that the predawn decreases of $N_m F_2$ are probably connected with changes of neutral-wind direction.

Acknowledgements

We thank the following for making their data available for this study: J. A. Klobuchar (and his many sources), J. K. Hargreaves, L. Kersley, R. Leitinger; also R. Conkright, for furnishing critical frequencies, and Johnann McKee, for the program for calculating Fourier coefficients. One of the authors (X. M. Liu) was a National Research Council Research Associate during the course of this study.

References

Bhuyan, P. K., and T. R. Tyagi, 1986, An empirical model of ionospheric electron content over Delhi, in Tauriainen, A., ed: *Radio Beacon Contribution to the Study of Ionization and Dynamics of the Ionosphere and to Corrections to Geodesy*, Oulu University, 77.

Davies, K., 1980, Recent progress in satellite radio beacon studies with particular emphasis on the ATS-6 Radio Beacon Experiment, *Space Sci. Rev.* **25**, 357.

Feitchter, E., R. Leitinger, and G. K. Hartmann, 1989, Modell-Parameter für den electroneninhalt der ionosphare in mittleren breiten, *Kleinheubacher Berichte* **32**, 623.

Klobuchar, J. A., 1987, Ionospheric time delay algorithm for single frequency GPS users, *IEEE*, AES-**23**, 325.

Ma Jainmin and Long Qili, 1988, The ionospheric total electron content and the slab thickness at the mid-latitudes in China, in Cao Chong, ed: *Investigation of the Ionosphere by means of Beacon Satellite Measurements*, Int. Acad. Publ., Beijing, 191.

Poulter, E. M., and J. K. Hargreaves, 1981, A harmonic analysis of ATS-6 electron content observations at Lancaster, U.K., during 1975-76, *Ann. Geophys.* **37**, 405.

Rao, P. B., B. Jayachandran, and N. Balan, 1988, Low latitude TEC and N_p variations during the solar cycles of 20 and 21 in Cao Chong, ed: *Investigation of the Ionosphere by Means of Beacon Satellite Measurements*, Int. Acad. Publ., Beijing. 241.

Rishbeth, H., 1989, Basic physics of the ionosphere, in Hall, M. P. M. and L. W. Barclay, eds: *Radiowave Propagation*, P. Peregrinus Ltd., London, U.K., 75.

AD-P006 306

THE VARIATION OF IONOSPHERIC SLAB THICKNESS DURING GEOMAGNETIC STORMS

Matthew W. Fox and Michael Mendillo
Boston University, Center for Space Physics
725 Commonwealth Ave., Boston MA 02215

Paolo Spalla
IROE - CNR
Via Panciatichi 64, 50127 Florence, Italy

Abstract

A first-order estimate of the shape of the electron density profile, N_e(h), is given by the ratio of the profile's total electron content (TEC) to its peak density (Nmax). Applications to transionospheric and hf supported propagation can be aided by models of this quantity, τ, and thus the BU/AFGL slab thickness model has been developed to describe the mean variations of slab thickness.

In this paper, we address the issue of how geomagnetic disturbances affect τ at mid-latitudes, making use of a database of two solar cycles of observations from Hamilton (Ma) and Wallops Island (Va). These results are compared with those deduced from an Italian station pair (Florence and Rome).

INTRODUCTION

An important goal of ionospheric modeling research today involves the specification of the electron density profile. Model profiles are useful operationally (for example, for integrating along a transionospheric raypath to determine ionospheric refraction, or for raytracing between two groundbased stations for location or communications reliability studies), and can be used to assist in the physical interpretation of experimental results. The peak parameters (such as foF2 and h'F2) are generally obtained from numerical representations of observations and can, in principle, be constructed to any desired accuracy over a particular sounding network, according to the particular problem being addressed. The profile shape, however, is a complicated function of the neutral and plasma temperatures, ionic compositions, neutral winds, and therefore of the season, local time, location, as well as the level of solar and geomagnetic activity The equivalent slab thickness is defined as the breadth (in km) of an ionosphere of uniform electron density equal to the peak density, Nmax. Functionally, τ = TEC/Nmax, where TEC = $\int_0^\infty N_e(h)\,dh$, and it is a convenient one-parameter summary of the electron density profile. Indeed, it can be related to a variety of quantities of interest that effect the overall profile shape. Apart from the intuitive relation to the F2 layer semi-thickness, various authors have proposed relations between τ and both neutral temperature and O^+/H^+ transition height (Titheridge, 1973), the scale height of atomic oxygen and drift motion (Furman and Prasad, 1973), and the mean gradient of electron temperature (Amayenc et al., 1971). In addition, τ is operationally a very useful parameter, as it allows one to make a simple conversion between foF2 and TEC The mean behavior of τ at Hamilton-Wallops Island has been summarized in a simple numerical model developed at Boston University (Fox et al., 1990).

The core of this report is a description of the day-to-day variations of τ, concentrating on geomagnetic storm periods when such variations are greatest and when underlying physical processes are qualitatively understood. The departures from quiet (or mean) conditions are presented under a variety of storm conditions in order to demonstrate the fundamental dependencies. These results are then compared and contrasted with those derived from an Italian station pair

91

91-09696

THE DATABASE

The bulk of the analysis contained in this report comes from data obtained in the mid-latitude American sector. Measurements of TEC were obtained from Hamilton, MA, by observing the Faraday rotation of signals from the satellites ATS-3 and ATS-5. The sub-ionospheric point of this raypath lies at nearly the same latitude as an ionosonde located at Wallops Island, VA, but shifted slightly in local time. The sample of data consists of sparse observations from 1965 and 1967, and nearly complete observations from 1968-1986. This gives us access to two solar cycles of TEC, Nmax and τ.

Additional data has been obtained from the Florence-Rome station pair. The TEC measurements were taken at Florence, by observing the Faraday rotation of signals from the satellites ATS-6 and SIRIO (Spalla et al, 1983, Spalla and Capannini, 1988). The sub-ionospheric point of these raypaths lay at a similar latitude as the Rome ionosonde, with again a slight shift in local time. The data sampling is somewhat irregular, running from 1975 to 1982. Much of the data was obtained at very high levels of solar activity.

An important point to bear in mind is that slab thickness data combine the errors of two observational parameters, TEC and Nmax, each of which can carry both error of measurement and interpretive error. A further complication is that there are physical differences in the measurements of foF2 and TEC. The former is deduced from vertical (or nearly vertical) incidence on the ionosphere, while the latter is measured along an oblique path. The physical difference between the exact sub-ionospheric point and the location of the ionosonde will also mean that TEC and Nmax will not necessarily relate to the same profile when horizontal gradients are large, such as in the period immediately following sunrise. As a consequence, the hourly slab thickness database may be quite noisy, and caution should be exercised in the interpretation of results.

We now present the mean behavior of slab thickness, in order to define the zero level for storm-related departures. The seasonal trends in the BU/AFGL slab thickness model (Fox et al., 1990) are shown in Figure 1, at typical low and high solar activity levels. Values are generally higher at higher solar activity and during the summer months. Generally, TEC and Nmax are highly correlated, so that τ is relatively constant compared to the mean diurnal variation of either TEC or Nmax.

The parameter used to measure the variations from 'normal' conditions is the percentage departure of the hourly value from a 31-day running mean hourly value. Running means should be a more stable base than monthly medians, especially around the beginning and end of equinoctial months. Means have the additional advantage of giving rise to departures that average to zero overall. Also, because τ is relatively constant, the percentage variations correspond to roughly the same absolute changes in τ under all conditions. A storm pattern is defined as the time variation of this storm departure parameter over the storm period. Here, the storm period is taken as one week, beginning the day prior to the onset of the storm (in LT), i.e. from SD0 to SD6 inclusive, with SD1 corresponding to the day of storm onset.

The day-to-day variations of τ are, like the mean variations, less than seen in TEC and Nmax. The times of greatest day-to-day variability in τ are the periods of geomagnetic storms. Not only are the physical effects that result in changes to TEC and Nmax the greatest, but the degree of correlation between TEC and Nmax is less and this manifests itself as greater variability in τ.

FEATURES OF THE STORM-RELATED VARIATIONS

We are restricting our attention in this report to average storm patterns. This allows us to highlight the effects that are common to most storms in a particular sample, rather than attempting to explain the complex morphologies of individual storm patterns. This will also reduce the effects of individually spurious hourly τ values that could be expected in our potentially noisy database.

In Figure 2, we show the average storm patterns obtained for τ using all storms (top panel), all summer storms (middle panel) and all winter storms (lower panel). The overall average pattern is worth describing in more detail, as it defines the basis for all the subsequent comparisons of storm patterns. The first noticeable feature is that τ is systematically enhanced during periods of geomegnetic disturbance, both during the positive and negative phases of the TEC and Nmax variations. That the layer thickness should behave so relates to the underlying physical processes of the positive (dynamical uplifting of the layer leading to enhanced electron densities in the topside) and negative (enhanced loss at the layer peak) phases. The enhancements are basically daytime phenomena at Hamilton, peaking at around 20% ($\Delta\tau$ = +60 km) on the day following the storm onset. By SD5, the average pattern shows no discernible departure from the zero level (mean behavior).

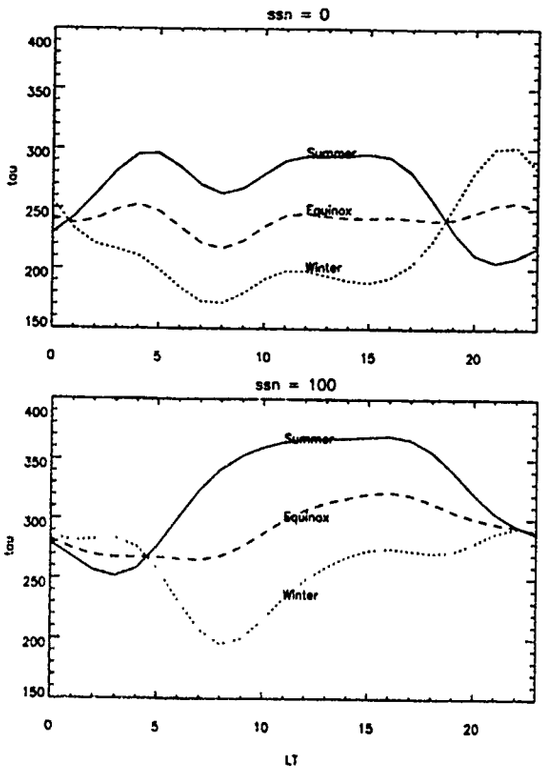


Figure 1. Mean seasonal variations in slab thickness over Hamilton, MA.

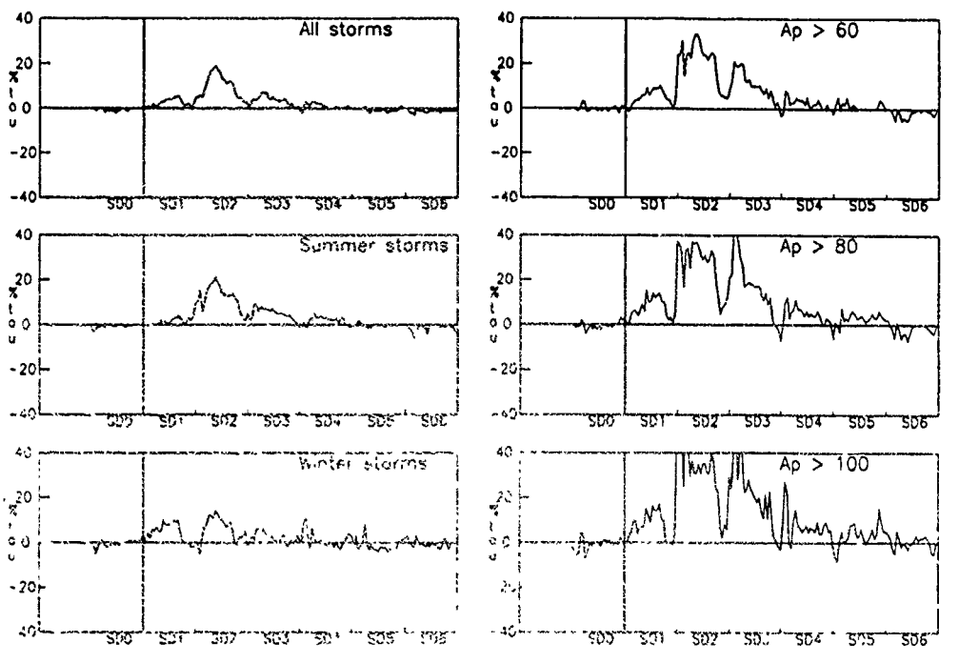


Figure 2 The overall and seasonal average storm pattern of slab thickness over Hamilton

Figure 3 The average storm patterns of slab thickness over Hamilton for

By comparing the storm patterns obtained during summer and winter months, the basic seasonal trends emerge. It is apparent that the enhancement on SD1 is greater in winter, while that observed on SD2 is greater in summer. This reflects an earlier positive phase on SD1 in the winter months, associated with a less severe negative phase on subsequent storm days. The duration of the average storm effect does not appear to be seasonally dependent. An interesting feature in the winter average pattern are midnight enhancements seen at SD3-4 and SD4-5. These are quite significant events at higher latitudes (Mendillo (1978) and Buonsanto et al. (1979)) and are attributed to auroral induced ionization. The effect is greater at Hamilton and less at the higher latitudes in winter because of the equatorward motion of the auroral region.

The panels in Figure 3 show the average storm patterns derived from only those storms corresponding to a peak Ap value of greater than specified cutoffs. Cutoff Ap values in the three panels are 60, 80 and 100 respectively. The relative enhancements in τ are clearly greater, the greater is the cutoff Ap value (this trend is even more apparent when Figure 3 is compared to the overall average storm pattern seen in the upper panel of Figure 2). The enhancements are systematically greater on each day in the storm period and the average storm effects last until SD5 (for Ap > 100 storms). The storm departures on SD3 are particularly enhanced, suggesting a longer-lasting negative phase for these strong geomagnetic storms, with the upwelling of molecular material (that underlies enhanced loss at the layer peak) occurring to a greater degree on later storm days.

The effect of varying the level of solar activity on average storm pattern is shown in Figure 4. The overall average pattern is repeated in the upper panel for ease of comparison, while the middle panel displays the patterns for low (monthly mean sunspot number < 50) solar activity, and the upper panel, high (sunspot number > 125) activity periods. While no dramatic trend is visible, the effect on τ is greater during times of high solar activity. This corresponds to times of greater absolute values of τ, and times when the percentage departures for both TEC and Nmax are significantly lower, and so this effect is worthy of attention.

The final correlation demonstrated here is that with respect to the onset time of the storm. The panels in Figure 5 show the average storm patterns obtained for storms of early, middle and late onset times (all in SD1 by definition). The changes are significant, as expected. When onset occurs early in LT, there is a significant enhancement in τ on SD1, associated with the earlier positive phase of the storm. The effect on SD2 is reduced, relating to a weaker negative phase of the storm. The SD2-3 overnight enhancement corresponds to a negative phase that is significantly stronger in Nmax than TEC- the overnight negative phases are more in line on subsequent storm days. The middle and lower panels show that for later onset times, no effect is seen in the storm patterns on SD1, and that on average the storm goes directly to the negative phase (and the related τ enhancement) on SD2. The enhancement appears to be greater in magnitude and occurs earlier for onset times in the middle of the day, while the effect lasts for more storm days when the storm onset occurs late in the day.

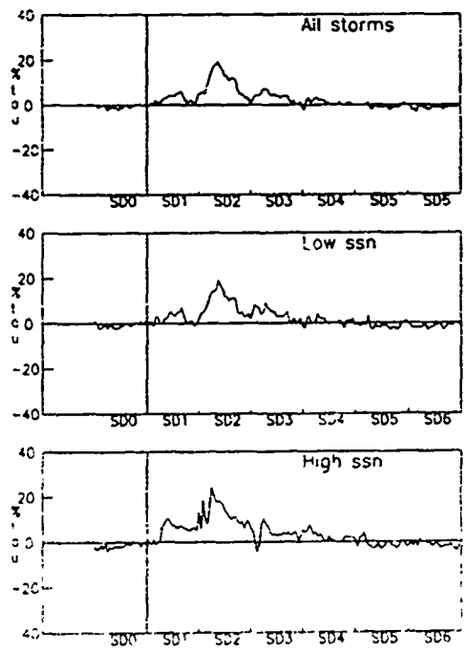


Figure 4. The overall average storm patterns and that of differing levels of solar activity. Hamilton

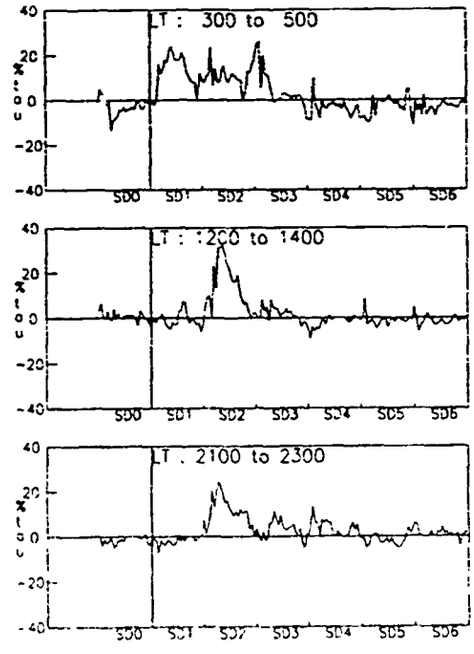


Figure 5. The average storm patterns of storms of differing onset times. Hamilton

Overall, there are clear correlations with storm perturbations in τ with respect to season, peak Ap, solar activity and onset time of the storm, and any numerical description of the average storm patterns should contain dependendencies on each of these observable quantities.

RESULTS OBTAINED FROM THE ITALIAN STATION PAIR

As mentioned earlier, the TEC database obtained from Florence was not as comprehensive as the Hamilton data set, both in the number of years available, and in the continuity within a given period. This meant that individual periods could be less certain with respect to the absolute TEC calibration. Yet, this is an important data set to examine for longitude effects in τ behavior under both quiet and disturbed conditions. Overall, it is not clear that the mean variations of τ are systematically different than seen at Hamilton (or at Aberystwyth, Wales, as described by Hajeb-Hosseinieh and Kersley (1975)). The mean variations of τ over Florence are shown in Figure 6. Differences in τ are apparent, especially in the nighttime values. On the other hand, if one allows for the change in vertical scales, the daytime values deduced at Florence are reasonably similar to those at Hamilton, as in Figure 1.

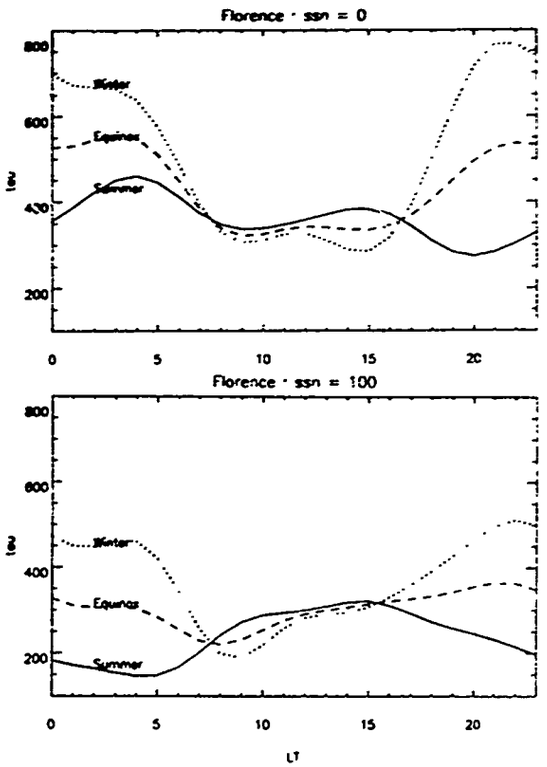


Figure 6. Mean seasonal variations in slab thickness over Florence, Italy.

However, we might expect our storm departure parameter not to be strongly influenced by uncertainty in the absolute TEC calibration. Percentage departures during storm periods with respect to a running mean τ (so that the mean is usually obtained from data in one or a small number, of observing bursts) should be much the same unless the calibration was significantly in error. In addition, we give greater significance to variations in the storm departure parameter during the daylight hours when interpreting the results of average storm patterns over Florence

In Figure 7, we show the average storm patterns derived over Florence, using all storms (upper panel), all summer storms (middle panel) and all winter storms (lower panel) The overall average pattern is quite similar to that obtained at Hamilton, allowing for the noisier pattern (smaller storm sample) Because of the noise, storm departures cannot be discerned beyond SD3, and no clear effect is seen on SD0 However, the SD2 pattern is much the same. An overnight enhancement on SD2-3 appears to occur more significantly in summer than winter, so its origin is not clear A seasonal trend is difficult to discern, as the average winter pattern appears to be little more than noise with a tendency to enhancements But like Hamilton the SD2 daytime enhancement is stronger in summer, with quieter SD0 behavior

The variation of Florence storm pattern with the peak geomagnetic activity is shown in Figure 8. Again the patterns are noisy, but the variability is clearly greater the higher the peak Ap index of the storm. Comparing the upper panels of Figures 7 and 8 (a cutoff Ap of 60 still permits sufficient storms for valid statistics) shows that enhancements are greater the higher the Ap value. In the average patterns for cutoff Ap values of 80 and 100, it is only really clear that the variability is greater.

A comparison of storm patterns obtained at different levels of solar activity is given in Figure 9, in the form of Figure 4. Unlike at Hamilton, departures (enhancements) at low solar activity at Florence appear to be greater although the statistics are weaker. However, like Hamilton, the enhancements in τ at Florence are greater at the level of solar activity when τ itself is greater.

The effect of storm onset time on average storm pattern is indicated in Figure 10, in the form of Figure 5. No clear pattern is seen in the upper panel (the average of early storms). However, there is a clearer tendency towards the anticipated enhancements in the patterns for middle and late onset time storms. The SD2 enhancement appears to be stronger in the middle panel, but unlike over Hamilton, the effect appears to last for longer also. This may simply reflect the smaller storm statistics in the Florence sample.

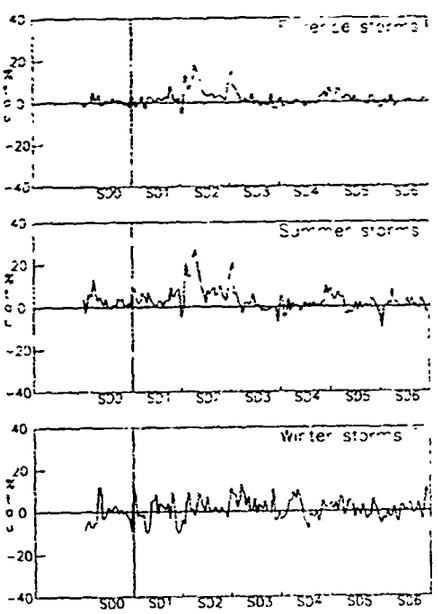


Figure 7 The overall and seasonal average storm patterns of slab thickness over Florence

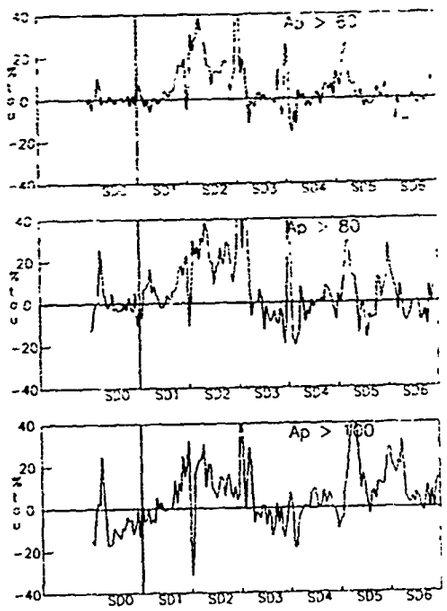


Figure 8 The average storm patterns of slab thickness over Florence for storms of differing strengths

A SIMPLE NUMERICAL DESCRIPTION

As a final stage of this report, we summarize the apparent storm-time changes to τ in simple numerical terms. This description is based on the Hamilton-Wallops station pair as this comprises the larger database, and because the existing slab thickness model of Fox et al. (1990) was also derived there.

We summarise the overall average storm-time perturbations as a cosine term in Local Time to account for the mid-day peak. Specifically, $\Delta\tau$(SD) = Amplitude(SD) cos(π (LT-12)/24) That is, a different amplitude of modulation is used each storm day (SD) In an overall average sense (and therefore, if nothing is known about the storm other than it exists) the amplitudes for SD1-4 are A_1 = 7, A_2 = 20, A_3 = 7, and A_4 = 3%

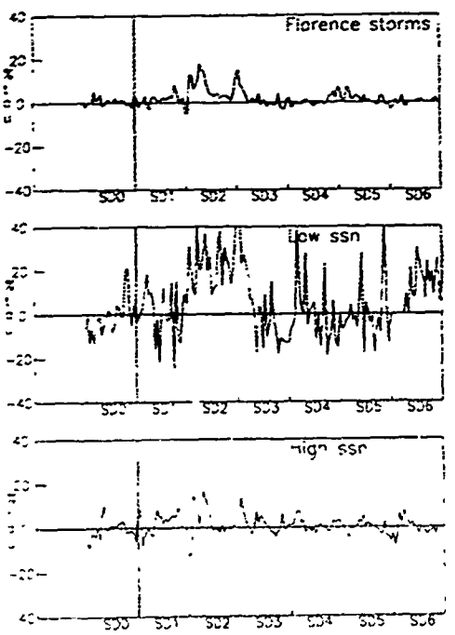


Figure 9 The overall average storm patterns and that of differing levels of solar activity - Florence

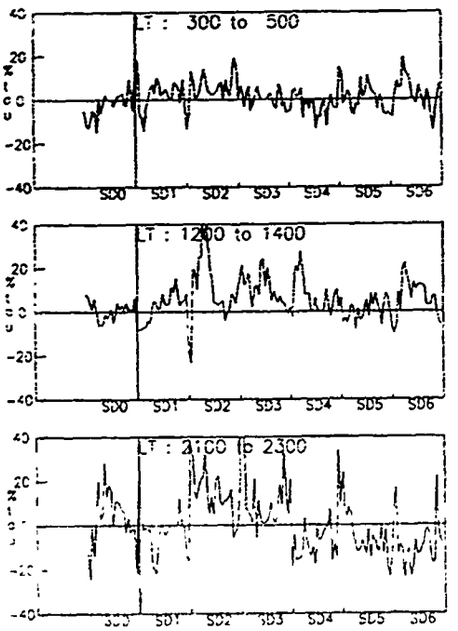


Figure 10 The average storm patterns of storms of differing onset times - Florence

These amplitudes contain dependencies on season, peak Ap index, onset Local Time of the storm, and mean solar activity Each of these are simply expressed here in terms of corrections to these average values, and are all given in percentage units. Also note that any net correction (or rather, net effect) of less than about 2% corresponds to no effect, and that a negative effect should not occur The seasonal adjustment can be written as $B_{SD} = b_{SD} \cos(S\,\pi)$, where S = 0, 1, 2 and 3 correspond to summer, fall, winter and spring respectively Specifically, b_1 = -3%, b_2 = 4%, b_3 = 2%, and b_4 = 2%. Equinoxes patterns were essentially the same as the overall averages

The adjustment due to variation of solar flux is only significant on SD2, and $C_2 = 0.05 \times (\text{ssn} - 50)$ %, with ssn as the monthly mean sunspot number

There is a significant dependence in the average τ pattern on peak Ap index, not only on the magnitude of the adjustment, but also on its diurnal variation. The stronger magnetic storms are more likely to involve nighttime and pre-dawn enhancements, but we do not consider that effect in this simple numerical description The magnetic correction can be summarized as $D_{SD} = d_{SD} \times (\text{Ap} - 40)$ %, with d_1 = 0 20, d_2 = 0 25, d_3 = 0 25, d_4 = 0.05 and d_5 = 0 10

The dependence on onset time of the storm (in LT) is not easily expressed in these simple terms. Physically, the variations are sensible, with early onset storms demonstrating SD1 effects, and later onset storms showing stronger effects on SD2 and beyond. The dominant effects are seen on SD1 and SD2 and these are summarised as follows: E_1 = +18% if onset occurs before dawn, and E_1 = -7% i onset occurs at dawn or after 1200 LT; E_2 = -6% if onset occurs before 1200 LT, E_2 = +15% if onset occurs between 1200 and 1500 LT, and E_2 = +5% if onset occurs after 1500 LT.

Thus, to predict the average amplitude of the storm pattern for any given day, one takes all the terms A, B, C, D and E (whichever are known) and adds them. Thus Amplitude(SD) = $A_{SD} + B_{SD} + C_{SD} + D_{SD} + E_{SD}$, to be used with the overall diurnal cosine modulation defined above

Similarly, we define the modulation of the nighttime peaks with a cosine squared term, but with a narrower width (here, two hours) The overall averages, using the conventions defined above, are NA_1 = $NA_2 = NA_3 = NA_4 = 0$ if ssn < 75 and NA_1 = 10, NA_2 = 7, NA_3 = 4, and NA_4 = 2 otherwise

The seasonal variation is defined similarly to the daytime peak We write $NB_{SD} = Nb_{SD} \cos(S\,\pi)$, with Nb_1 = 5%, Nb_2 = 4%, Nb_3 = -4%, and Nb_4 = -2%.

The dominant term for variation in the nighttime peak magnitude is the dependence on magnetic activity. We derive $ND_{SD} = Nd_{SD} \times$ (Ap-40) with $Nd_1 = 0.5$, $Nd_2 = 0.5$, $Nd_3 = 0.3$ and $Nd_4 = 0.0$. Note the size of the coefficients with respect to the daytime peak values.

This model is the simplest realistic representation of the storm patterns in τ seen over Hamilton, and does at least indicate the degree of sensitivity to each of the variables that are important enough to affect the average storm patterns. However, individual storms demonstrate a significant degree of variation about this average pattern, and we propose to undertake further work on specific cases to investigate the reasons for these additional variations.

SUMMARY AND CONCLUSIONS

Average storm patterns have been derived at two mid-latitude sites and analyzed for a correlation with various observable quantities. There were correlations deduced for the storm pattern with respect to season, solar activity, geomagnetic activity and storm onset time. Some differences were suggested in both the mean behavior of τ and in the storm patterns between the sites (Hamilton, MA and Florence, Italy), but the significance of the differences are diminished both by possible errors in the absolute TEC calibration over Florence, and by the smaller sample of storms covered there. Work is continuing on validating the Florence TEC data, as it constitutes an important independent data set. A more complete numerical description of the average storm patterns (covering TEC, Nmax and τ) at Hamilton is currently being developed, and case studies will be made to further investigate dependencies, to model these where possible and to use this as a guide to the day-to-day variability of the ionosphere.

REFERENCES

Amayenc, P., Bertin, F. and Papet-Lepine, J. An empirical relationship between ionospheric equivalent slab thickness and mean gradient of the electron temperature in the F region, Plan. Space Sci., vol. 19, p. 1313, 1971.

Buonsanto, M.J., Mendillo, M. and Klobuchar, J.A. The ionosphere at L=4: average behavior and the response to geomagnetic storms, Ann. Geophys., vol. 35, p. 15, 1979.

Fox, M.W., Mendillo, M. and Klobuchar, J.A. Ionospheric equivalent slab thickness and its modeling applications, In Preparation, 1990.

Furman, D.R. and Prasad, S.S. Ionospheric slab thickness: Its relation to temperature and dynamics, J.G.R., vol. 78, p. 5837, 1973.

Hajeb-Hosseinieh, H. and Kersley, L. Analysis of satellite recordings. Studies of ionospheric slab thickness, Interim Scientific Report, Air Force Cambridge Research Lab., Bedford, MA, U.S.A., 1975

Mendillo, M. Behavior of the ionospheric F-region during geomagnetic storms, Technical Report AFGL-TR-78-0092(II), Air Force Geophysics Lab., Bedford, MA, U.S.A., 1978.

Spalla, P., Capannini, E., Ciraolo, L. and Ragionieri, R. Use of the VHF signal of the SIRIO for ionospheric investigations, URSI Symposium on Beacon Satellite Studies of the Earth's Environment, National Physical Laboratory, New Delhi, India, 1983.

Spalla, P. and Capannini, E. Total Electron Content (TEC) in the Mediterranean Area: an Atlas, Consiglio Nazionale Delle Ricerche, Rome, Italy, 1988.

Titheridge, J.E. The slab thickness of the mid-latitude ionosphere, Plan. Space Sci., vol. 21, p. 1775, 1973.

AD-P006 307

REDUCTION OF IONOSPHERIC EFFECTS ON POLARIZATION MEASUREMENTS DURING SATELLITE TRACKING

C. Yoo and M. C. Lee
Plasma Fusion Center
Massachusetts Institute of Technology
Cambridge, Massachusetts 02139

ABSTRACT

Since the beginning of 1988, polarization measurements have been made druing the satellite tracking with the L-band (1.295 GHz) radar located at Millstone Hill, Massachusetts. This work aims at reducing the ionospheric effects on the phase differences between the two different circularly polarized signals, both reflected from the tracked satellite. The success of the effort will be conducive to improving the radar capability for space object identification. To predict the phase shifts caused by the trans-ionospheric propagation of the radar signals, we have used an ionospheric model which combines the slab model proposed by Klobuchar and the foF2 model suggested by Chiu. The computed results are compared with the phase shifts derived from the total electron content (TEC) measurements. These data were recorded at Hamilton, Massachusetts with the Faraday rotation measurements of the linearly polarized 137 MHz signal transmitted from GOES-II satellite. It is found that the predictions made with the aforesaid ionospheric model agrees with the measurements for most of the cases. However, during the periods of high solar fluxes or when the ionosperic density reaches its peak values in the day, this model seems to overpredict the TEC.

1. Introduction

The Millstone Hill L-band Satellite Tracking Radar located in Westford, Massachusetts, transmits right-hand-circularly-polarized (RHCP) signals, and the major portion of the reflected energy is contained in the left-hand-circularly-polarized (LHCP) signals. Because of the complexity of the satellite surfaces, however, a significant amount of the reflected signal returns also as RHCP waves. Depending on the shapes of satellites, the energy in the RHCP wave can be greater than 30% of the energy in the LHCP signal. By utilizing the information contained in the LHCP signals, one can greatly improve metrics and space object identification, the two most important aspects in satellite tracking.

The phase shift due to transionospheric propagation between the two polarized waves makes the polarization measurements more difficult. To reduce this ionospheric effect, we calculated the amount of phase shift, using an ionospheric model by Klobuchar and Chiu, which combines the slab model proposed by Klobuchar and the foF2 model suggested by Chiu. The estimated phase shift from this calculation would, then, be subtracted from the measured phase difference between the two polarized signals. This had been incorporated into a FORTRAN code which takes time, metrics, foF2, Hmax, and F10.7cm

91-09697

solar flux as inputs and then calculates the Faraday rotation (FR) angle of 1.295 GHz linearly-polarized (LP) waves due to the ionospheric effect by estimating the total electron content (TEC) along the line of sight from the Millstone Hill to the satellite. It can be easily shown that, given the same anisotropic condition and frequency, the phase shift between two oppositely oriented circularly-polarized (CP) waves is twice the amount of Faraday rotation angle of a LP wave. When the amplitudes of the CP waves are unequal, the direction of the semi-major axis of the polarization ellipse (resulting from the addition of the two CP waves) corresponds exactly to the Faraday rotation angle.

To evaluate the accuracy of the computed FR angles and to improve the model, it became necessary to compare the estimated FR angles with the direct measurements. The Geostationary Operational Environmental Satellite II (GOES-II), transmits 137 MHz linearly-polarized (LP) beacon signal to Hamilton, MA. and the TEC is derived from these FR angles and is recorded by the Geophysics Laboratory (GL) at Hanscom Air Force Base in Massachusetts. After the TEC values are recorded in magnetic storage, the Faraday rotation information is discarded. The equivalent FR angles for 1.295 GHz were obtained through the following equation:

$$\Omega = \frac{2.36 \times 10^{-5} \overline{M} N}{f^2}$$

where

$$\overline{M} = \overline{B \cos\theta \sec\chi}$$

$$N = TEC$$

$$\Omega = FaradayRotation$$

$$B = GeomagneticFieldIntensity$$

$$\theta = PropagationAngle$$

$$\chi = ZenithAngle$$

$$f = 1.295GHz$$

The GL has used the inverse of this equation (with the exception of the difference in frequencies) for the conversion of FR (at 137 MHz) to TEC. Therefore the FR at 1.295 GHz from TEC is almost as accurate as the direct measurements of FR at 137 MHz.

Through the finite element set calculation, the metrics of the GOES-II was obtained, which in turn, was entered into the computer code implementing the adopted ionospheric model. Other parameters such as foF2, Hmax, and F10.7cm flux were entered to generate the set of data from the model. This set of data containing the Faraday rotation angles (i.e. the exact half of the phase difference between the LHCP and the RHCP waves) in the 15 minute interval was compared with the 1.295 GHz Faraday rotation angles calculated from the TEC measurements.

2. Analysis

The two sets of data, FR measured from TEC and FR estimated through calculation, were averaged for each hour of the day from the beginning of October, 1988 to the end of March, 1989. A typical case of diurnal variation in the FR angles, both measured and estimated, is shown in Figure 1. The model produces significantly large amount of hourly fluctuation in the FR angles, whereas the measurements show the changes in the FR angles consistent with the time of the day. Even when averaged for a month, the fluctuations from the model calculations were rather high, especially during the middle of the day (from 1400 to 2100 UT), although the night time averages exhibited reasonably small standard deviations.

The monthly averages at night also show that the model's prediction of the FR angles is acceptable within 10 to 15%. During the daytime, however, the estimation exceeded the the actual measurements by a consistently large bias on average. (See Figure 2.) On certain days, the estimates and the measurements were well within 10° from each other. During these periods, there does not seem to be any significant bias. On the other hand, there were periods when the model overestimated the FR angles by more than 35° in some cases. We have found that these days of overprediction by the model corresponded closely to the days of high solar flux (F10.7cm). For instance, from 14 to 25 December, 1988, the solar fluxes were higher than 200 (i.e. $200 \times 10^{-22} Watt/m^2$), and the estimated FR angles were upto 40° greater than the measured angles. (See Figure 3.) During the month of January, 1989, the solar flux consistently stayed above 200, almost for the whole month, and again, the model overpredicted the the FR angles by 35° on average. These consistent overestimations created during these two months were also observed in February and March, which also had many days of high solar fluxes.

3. Conclusion

The model calculations exhibited large fluctuations in the estimation of the Faraday rotation angles in the small time scale (15 minutes). Out of all the input parameters, only the foF2 and Hmax changed significantly in time as short as 15 minutes, whereas all the others were more or less constant over that time scale. Therefore we suspect that the measurements of foF2 and Hmax by the digisonde were not as nearly accurate as the TEC measurements.

For monthly averages, the model was able to give acceptable predictions at night, with relatively small standard deviations. The eminent failure of the model was the daytime prediction, especially during the hours of the peak electron density on the days of high solar flux (greater than 200). It is suggested that the model should use the more updated weighting factors in incorporating the solar flux to determine the total electron content along the line of propagation.

A newer and more sophisticated ionospheric model using the Chapman profile has been tested, but the more complicated code slowed down the calculation. Within the consideration of the fact that the ultimate purpose of this project is on the real-time basis during satellite tracking, the speed of the estimation is rather essential. As an alternative approach, we recommend that the averaged previous hour's data be used in determining the amount of phase shift, so as to eliminate the effect of large fluctuations.

Acknowledgments. Special thanks to A. J. Coster of M.I.T. Lincoln Laboratory and J. A. Klobuchar of the Geophysics Laboratory for their constructive suggestion and information. This work was supported by the NASA grant NAG 5-1055, Lincoln Laboratory grant (under the Air Force contract to Lincoln Labortory F19628-80-C-0002), and the AFOSR grant AFOSR-88-0217.

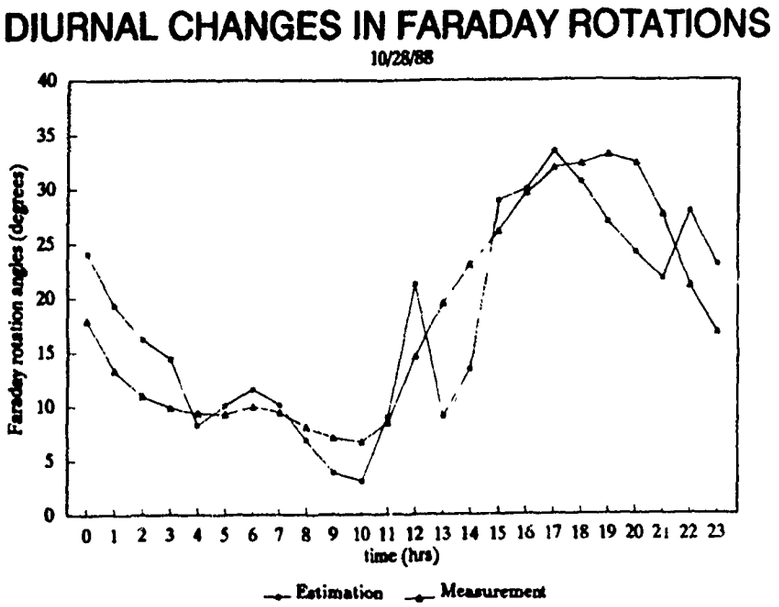


Figure 1. An example of diurnal variations on estimated and measured Faraday rotation angles. The estimation from the model has large fluctuations.

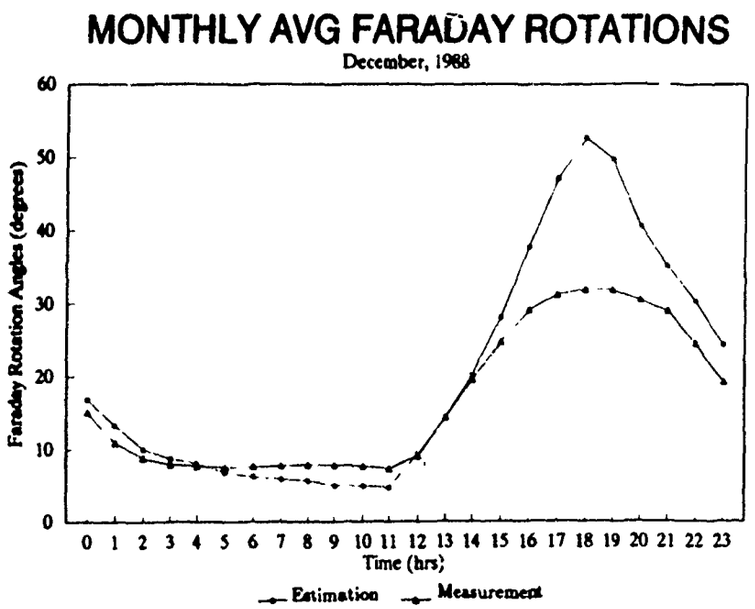


Figure 2 The average Faraday rotation angles in December, 1988 During the daytime, the estimates are much larger than the actual measurements.

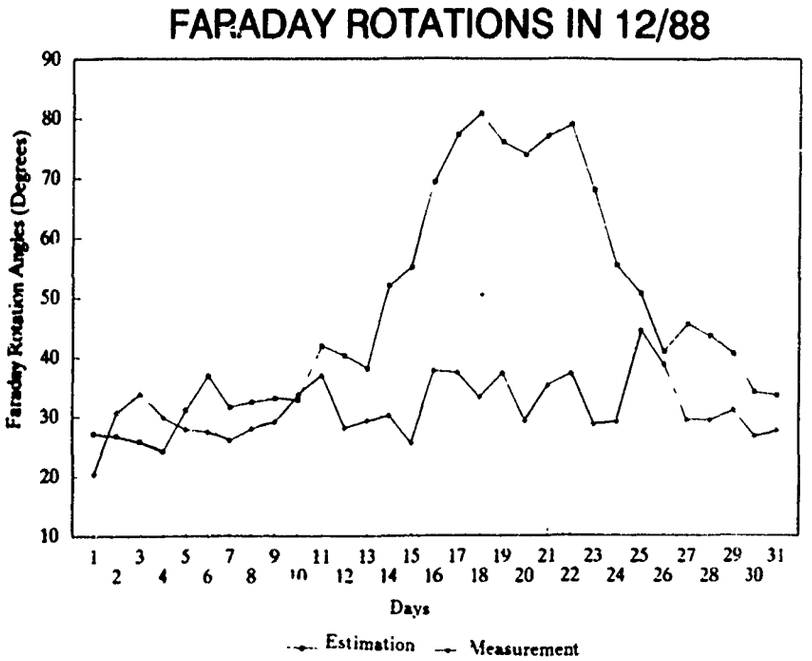


Figure 3 The daily change in the daytime Faraday rotation angles in December, 1988 From 12/14 88 to 12/25/88, in the period of extremely high solar flux, the estimates from the model calculations far exceed the actual values.

AD-P006 308

IONOSPHERIC OBSERVATIONS IN WESTERN-EUROPE

F.J. Gorman and H. Soicher
U.S. Army Communications-Electronics Command
CECOM Center for C³ Systems
Fort Monmouth, NJ 07703

ABSTRACT

Hourly digital ionograms were taken at two Western European locations: Dourbes, Belgium (50.10° N, 4.60° E) and, Roquetas, Tarragona, Spain (40.80° N, 0.30° E) during Winter through Summer of 1988. Ionospheric parameters (e.g. f_0F_2, f_0F_1, h'F, $h'F_2$, f_{min}, f_0E_s, MUF, $f_{min}F$, f_xI, $f_{min}E$, f_0E, h'E, and $h'E_s$) were tabulated and compared at the two stations, which were separated by ∅1000 km, using correlative techniques.

Future plans call for the establishment of additional stations in order to determine correlation distances for the various ionospheric parameters.

INTRODUCTION

Long-range HF radio communication depends on the ability of the ionosphere to return the radio signal incident on it back to earth. Prediction of ionization levels in the various ionospheric regions are derived from models and are used as a basis for planning and frequency management of HF radio systems worldwide. The models permit the calculation of system parameters such as operating frequencies, signal strengths, signal-to-noise ratios, and multipath probability that can be used to describe the performance of HF radio systems. Uncertainties or inaccuracies in the models of the ionosphere have long been known to be one of the major, if not the major cause for inaccuracies in the calculated propagation characteristics. This is particularly true for those applications of ionospheric predictions involving time scales that are less than the monthly average or monthly median.

To reduce average monthly RMS errors in predictions, adaptive techniques that use real-time observations to correct model biases have been devised. Examples of such are sounding the desired communication path prior to information transmittal, or sounding the vertical local ionosphere to determine a real-time model reference point and then to adjust the model for the future - until the next reference sounding. While the first is appropriate for fixed point communication paths it is cumbersome and causes EM interference over wide geographic areas; whereas the second technique, while not causing interference, may not yield values representative of ionospheric conditions at the reflection point of an oblique path. Since the Army is interested in mobile as well as in fixed communication nodes the determination of the correlation of propagation characteristics in the short, medium, and long range oblique modes with those of a local vertical mode, is of particular interest. Thus determined correlation distances will permit the assessment of the reliability of HF tactical communication circuits within an extended area, based on monitoring of the vertical ionosphere at a central location.

To this end the U.S. Army Communications-Electronics Command (CECOM) has initiated the European Communications Analysis Program (EUROCAP). It aims at the study of correlation distances of HF skywave propagation characteristics by

91-09694
91 9 1 136

recording both vertical and oblique ionograms. A regionalized data base is required to develop an accurate HF prediction algorithm, and a cooperative network of ionospheric sounder facilities was established in order to accumulate the propagation data. Preliminary results of the program based on vertical soundings taken by two stations in the EUROCAP network, the Royal Institute of Meteorology, Dourbes, Belgium, and the Ebro Observatory, Roquetas, Tarragona, Spain, separated by 1087 km of ground distance is reported here.

THE DATA

Vertical ionograms using the Digisonde 256 (Reinisch, 1986) were taken at hourly intervals from February 29 to July 3, 1988 at the Spanish location and from February 4 to June 17, 1988 at the Belgium location for a total of ~6200 ionograms. Digital ionograms are scaled automatically using the Automatic Real Time Ionogram Scaler with True Height (ARTIST) (Reinisch and Huang, 1982). Amplitude, phase, incidence angle, polarization and Doppler shift measured by the Digisonde are analyzed to extract the overhead ordinary and extraordinary traces even during disturbed ionospheric conditions. The electron density profile may be calculated from the ordinary trace using profile-fitting methods. Whereas the ARTIST scales 18 parameters: f_oF_2, f_oF_1, h'F, h'F_2, MUF(3000)F_2, M(3000)F_2, f_{min}, $f_{min}F$, f_xI, f_oE, h'E, f_oE_s, h'E_s, $f_{min}E$, range spread F, range spread E, frequency spread E; only a selected sample of these parameters will be reported here. Examples of data will be given for March 1988 when Relative Sunspot number varied from 20 to 120 on a daily basis..

The diurnal variation of f_oF_2 values (ordinary wave critical frequency) at the Belgium and Spain sites were measured at hourly intervals for the month of March 1988 is shown in Figures 1a and 1b, respectively. At Belgium, the mean value varied between 2.9 MHz (0500 UT) and 7.0 MHz (1200 UT) with the standard deviation varying between .5 MHz to 1.0 MHz. In Spain, the minimum and maximum of the diurnal variation occur at the same time as in Belgium with the mean values varying between 3.5 MHz and 9.9 MHz, with the standard deviation varying between .4 MHz and 1.0 MHz. The range of frequencies available to the more southern station is ~60% larger than that of the northern station, mostly at the higher range of the spectrum.

The hourly values of h'F (minimum F layer virtual height) in Belgium and Spain for the month of March 1988 are shown in Figures 2a and 2b, respectively. For Belgium the mean height varies from a minimum of 200 km (1200 UT) to 310 km (2400) with the standard deviation varying from 10 km to 50 km. In Spain the mean height varied from 200 km to 300 km at the same times with the standard deviation varying from 12 km to 33 km.

The hourly values of MUF(3000) F_2 (Maximum Usable Frequency) in Belgium and Spain for the month of March 1988 are shown in Figures 3a and 3b, respectively. For Belgium the mean MUF varied from a minimum of 8.6 MHz (0500 UT) to a maximum of 22.7 MHz (1700 UT) with another maximum of 22.4 MHz (1200 UT). The standard deviation varies from a minimum of 1.7 MHz to a maximum of 4.2 MHz. The MUF at Spain varies from a minimum of 10.2 MHz (0500) to a maximum of 26.3 (1200 UT) and a second maximum of 26.2 (1700 UT). The standard deviation varies from a minimum of 1.2 MHz to a maximum of 3.4 MHz. The range of MUF available to the more southern station is ~11% larger than that of the northern station mostly at the higher range of the spectrum.

The hourly f_{min} values (the lowest frequency at which echo traces are observed) in Belgium and Spain for the month of March 1988 are shown in Figures 4a and 4b, respectively. The mean values for both stations are not highly variable. At both stations the values vary between -1.3 MHz and -1.7 MHz without any trend. The standard deviation varies between .08 MHz and .2 MHz.

As a quantity that does not occur at all time, the variation of f_oF_1 (the critical frequency of the F_1 layer) values for Belgium and Spain are shown in Figures 5a and 5b respectively. At the Belgium location f_oF_1 appeared at 0700 UT and disappeared after 1600 UT, with mean values ranging from 2.7 MHz to 4.5 MHz.

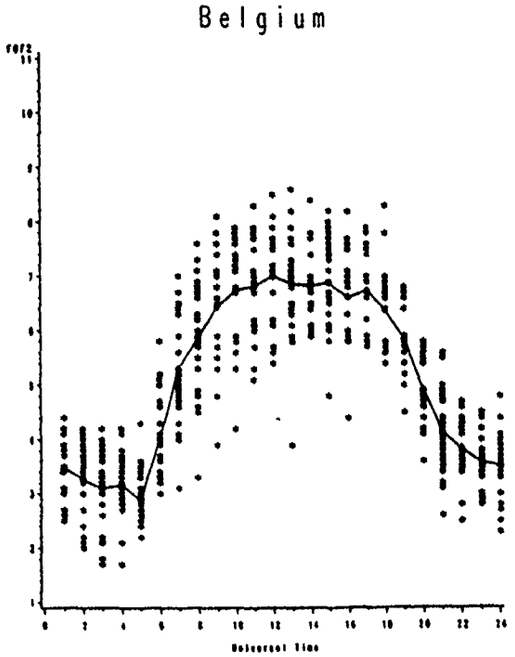


Figure 1a

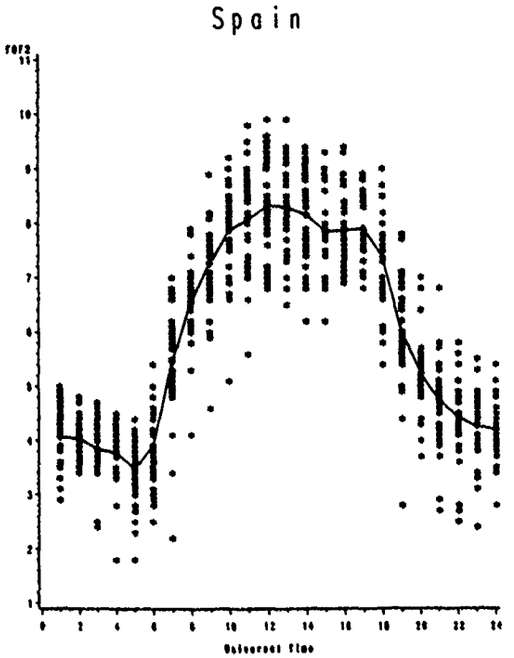


Figure 1b

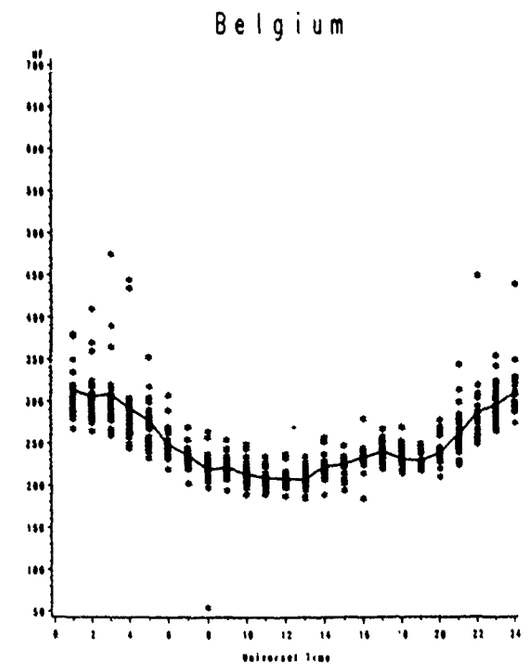


Figure 2a

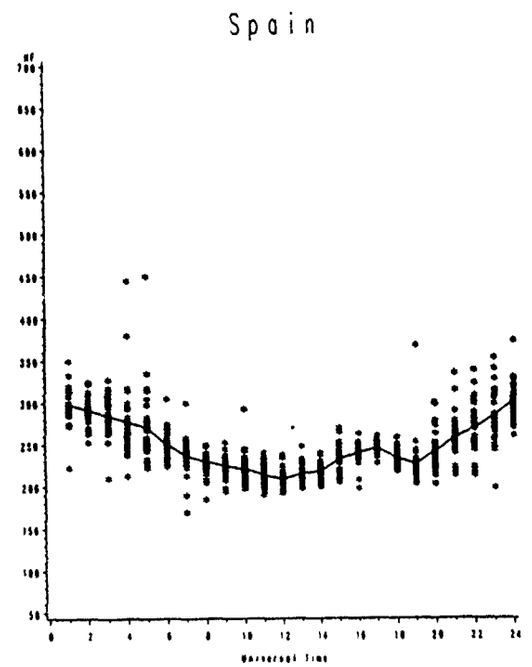


Figure 2b

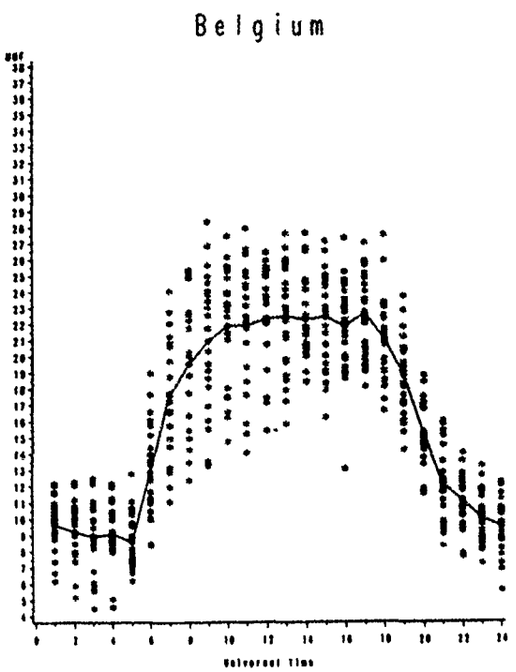
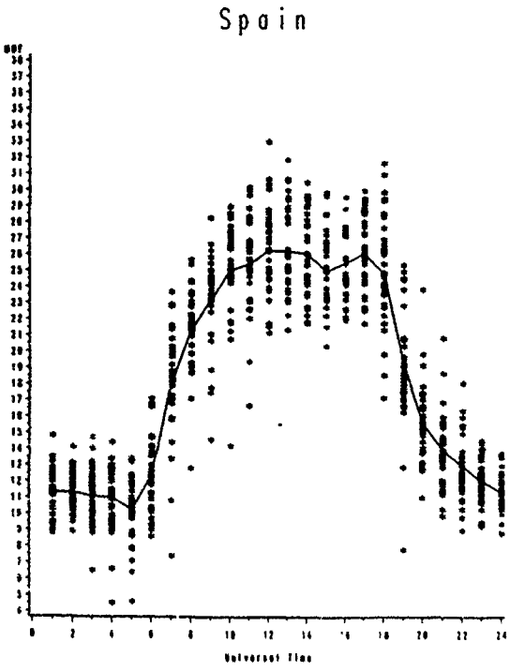


Figure 3a



Fibure 3b

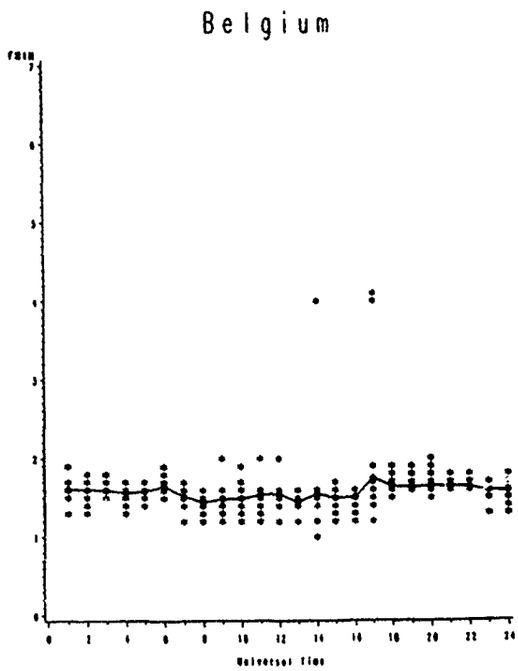


Figure 4a

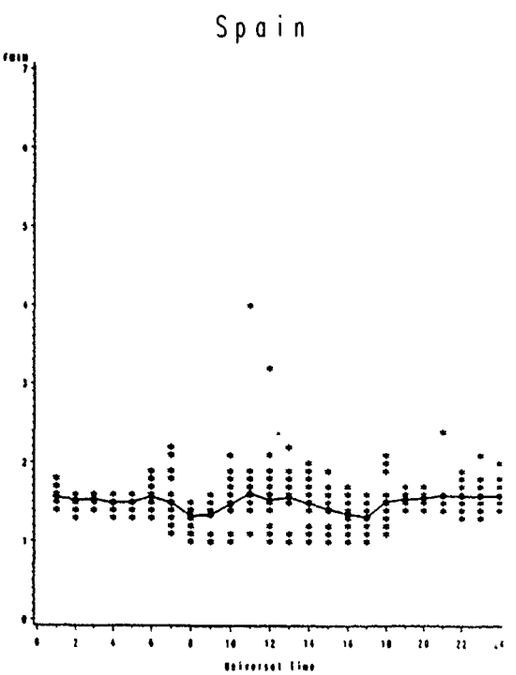


Figure 4b

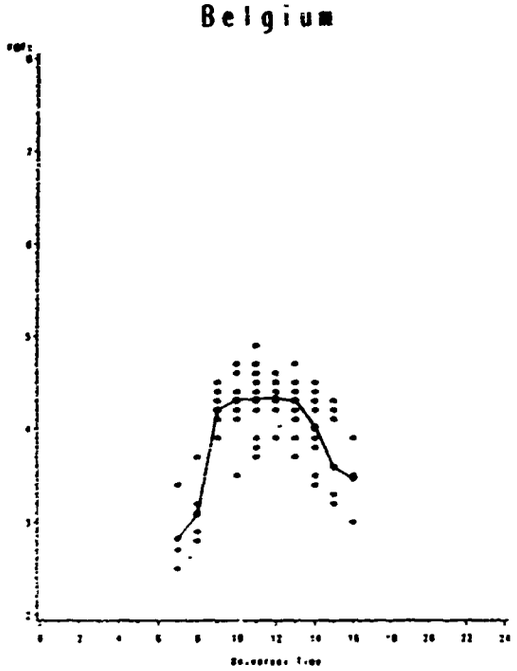


Figure 5a

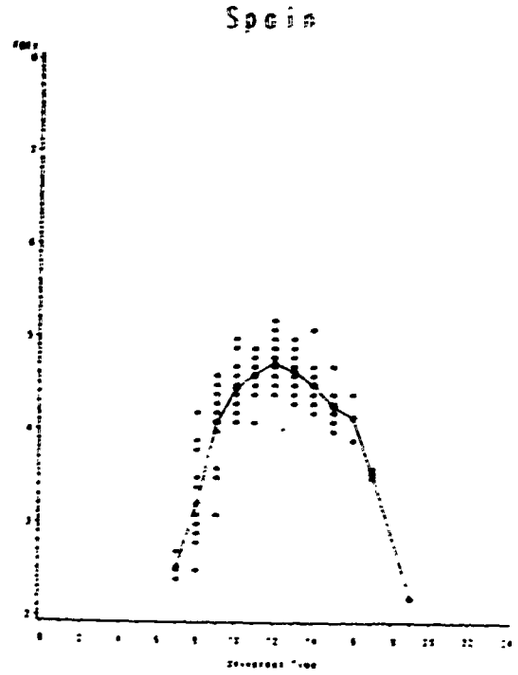


Figure 5b

The standard deviation varied from .17 MHz to .5 MHz. At the Spanish location f_0F_1 appeared at the same time (0700) but disappears after 1900 UT (although from 1600 to 1900 UT it existed only on very few days). The standard deviation varied from .05 MHz to .3 MHz.

Correlation analysis for the selected ionospheric parameters was carried out. A least-squares regression line was fitted to the two data sets with the time parameter eliminated between them. The regression lines for f_0F_2, h'F, MUF(3000)F_2, and f_{min} for March 3, 1988 (day 63) are shown in Figures 6a, b, c, and d, respectively. The coefficient of correlation is defined as

$$r = \pm \left(\frac{\Sigma(Y_{est} - \bar{Y})^2}{\Sigma(Y - \bar{Y})^2} \right)^{1/2}$$

where Y_{est} represents the value of the parameter at Belgium for a given value of the parameter at Spain estimated from the fitted regression line; $\bar{Y}$ is the mean of Y.

The correlation coefficients for the various parameters were calculated in daily intervals, i.e., the data throughout a specific day was correlated at the same time for corresponding paramaters. Since local time differences (~18 minutes) are within data resolution (hourly) the data sets were not shifted in time for adjustment. Correlation coefficients for f_0F_2, h'F, MUF(3000)F_2 and f_{min} for March 1988 are shown in Figures 7a, b, c, and d, respectively.

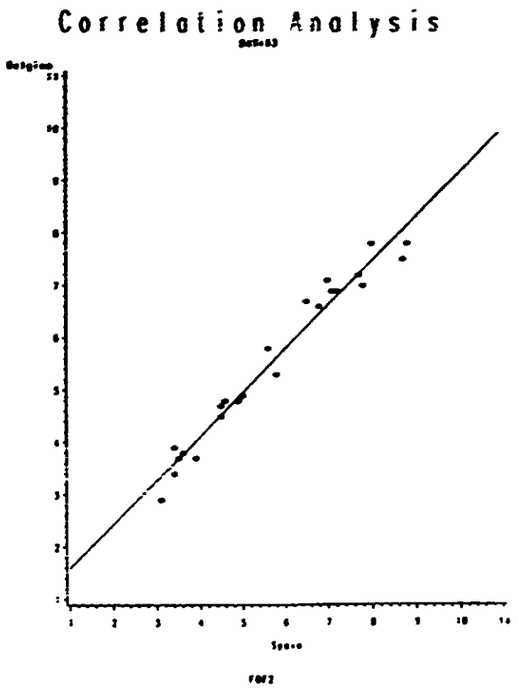


Figure 6a

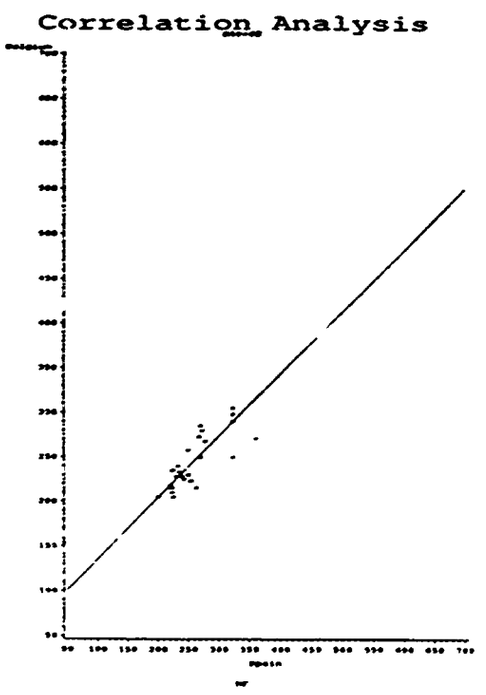


Figure 6b

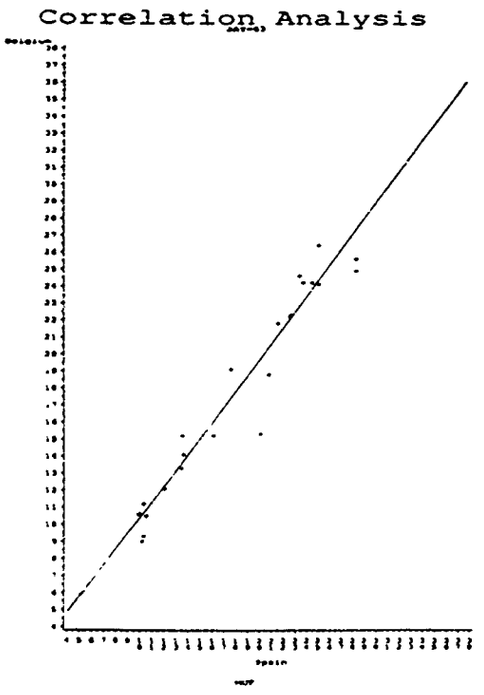


Figure 6c

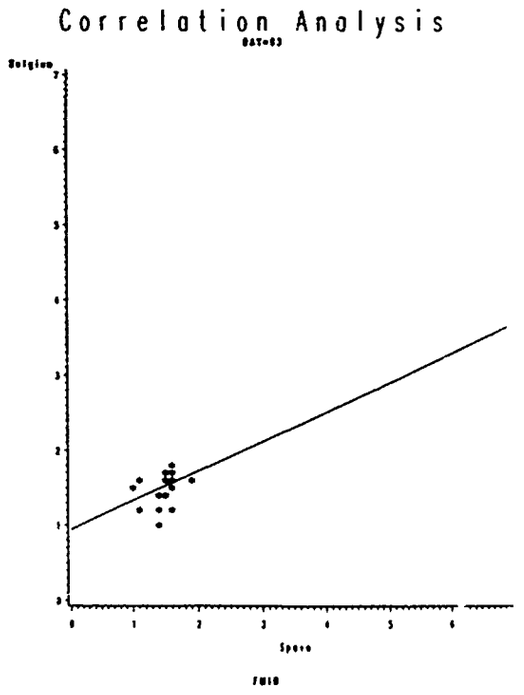


Figure 6d

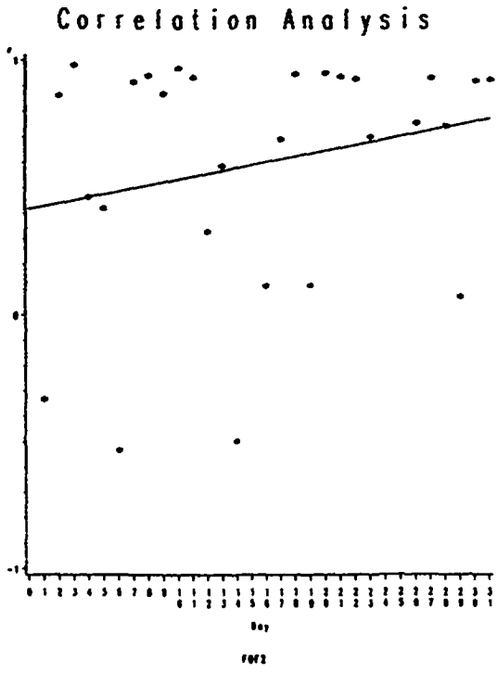


Figure 7a

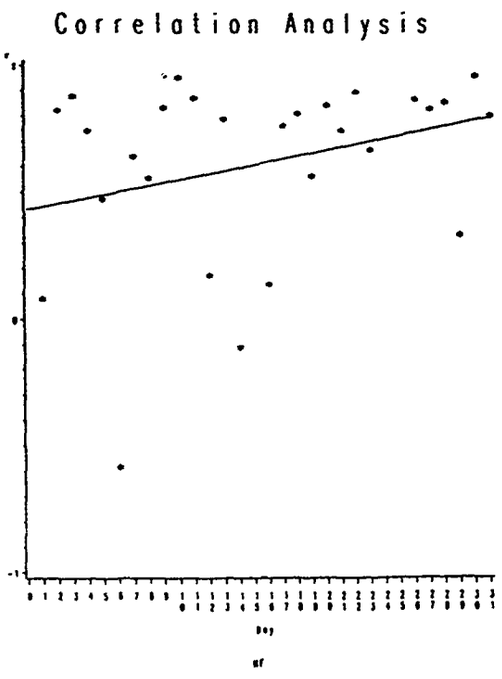


Figure 7b

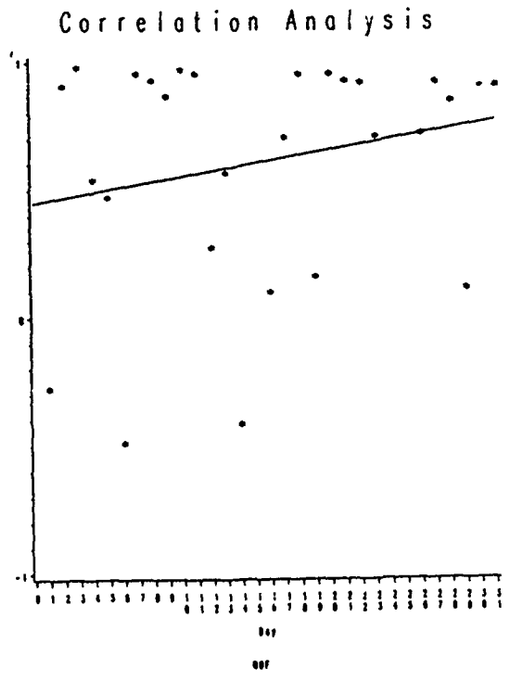


Figure 7c

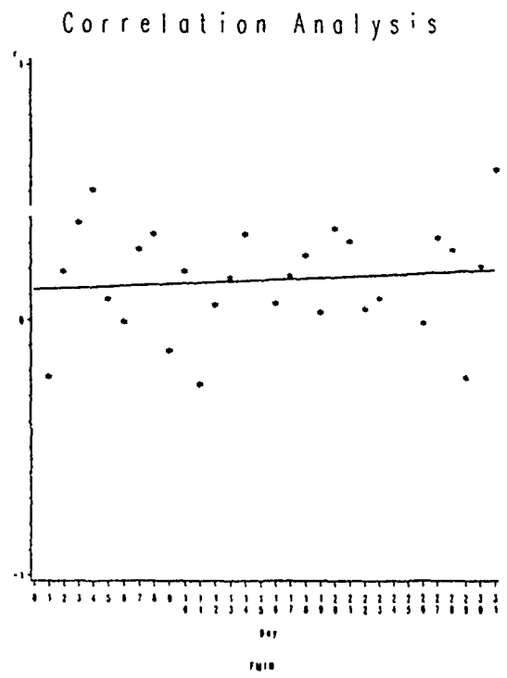


Figure 7d

DISCUSSION

The availability of ionospheric-parameter data at the two European locations at the solstice (March/April) and summer (June) seasons provided a description of seasonal behavior at the individual locations and a means of comparison of such behavior between the two locations. In Belgium the mean values of f_0F_2 varied during solstice between 3.25 MHz and 7.25 MHz (Δf_0F_2 = 4.0 MHz); while during the summer they varied between 5.5 MHz and 7.0 MHz (Δf_0F_2 = 1.5 MHz), allowing for wider range of HF Communication frequencies for use throughout the day even though maximum f_0F_2 is lower. In Spain mean values of f_0F_2 varied during solstice between 3.75 MHz and 8.3 MHz (Δf_0F_2 = 4.55); while they vary during summer between 5.0 MHz and 7.6 MHz (Δf_0F_2 = 2.6 MHz), again allowing for wider range of frequencies for use throughout the day. A comparison of the two locations shows that during solstice and summer, Spain had a wider range of HF Communication frequencies available for use, although during summer Belgium could support propagation of higher frequencies at night. Solar Zenith angle considerations probably account for the higher maximum values at the more southern location.

In Belgium mean values of h'F varied during solstice from about 210 km (1300 UT) to 305 km; while during the summer they varied from 210 km (1300 UT) to 290 km. In Spain mean values of h'F varied during solstice from about 210 km (1200 UT) to 305 km; while in summer they varied from 230 km (1200 UT) to 280 km. Height variability was larger at both locales during solstice than during summer.

In Belgium mean values of $M(3000)F_2$ varied during solstice from about 10 MHz to 22.5 MHz (Δ= 12.5 MHz); while during the summer they varied between 16.5 MHz and 20.0 MHz (Δ= 3.5 MHz), providing a greater range of frequencies for use throughout the day even though maximum MUF is lower. In Spain the mean values of $MUF(3000)F_2$ varied during solstice between 11 MHz and 27 MHz (Δ= 16 MHz); while they varied during summer between 16.5 MHz and 23.5 MHz (Δ= 7.0 MHz), again allowing for greater range of frequencies for use throughout the day. Comparison of the two locations shows that during solstice and summer Spain has a wider range of frequencies available for use throughout the day. Comparison of the two locations showed that during solstice and summer, Spain had a greater selection of HF Communication frequencies availabe for use. This is similar to the discussed variability of f_0F_2, and is attributed to solar Zenith angle, considerations at the more southern station.

In Belgium mean values of f_{min} varied during the solstice between 1.5 MHz and 1.6 MHz; while during the summer they vary between 1.5 MHz and 1.7 MHz The variation does not have a time trend. In Spain mean values of f_{min} varied during solstice between 1.5 MHz and 1.6 MHz; while they varied between 1.8 MHz and 1.9 MHz during the summer. The parameter f_{min} appeared to be little varying at both locations. Summer values at both locations were somewhat higher than corresponding solstice values.

The regression analysis for a specific day (March 3, 1988), shows excellent correlation for f_0F_2, h'F, and MUF_3 and a lesser one for the parameter f_{min}. Correlation coefficients for the entire month of March 1988 indicate values >.6 for most days, with values >.8 for about 50% of all days for f_0F_2 and MUF; while values >.5 for most days for h'F were indicated, with little correlation indicated for f_{min}. It appears that on most days f_0F_2 and MUF may be ascertained at any time at one locale from a known current variation of these parameters at the other locale. No such claim may be made for f_{min}. However, f_{min} is not highly variable quanitity and its usefulness for communications applications is limited.

ACKNOWLEDGEMENTS

The data at Belgium was taken under the guidance of Dr. J.C. Jodogne, while at Spain it was taken under the guidance of Dr. Luis Alberca of the Ebro Observatory.

REFERENCES

REINISCH, B.W. and HUANG, X. [1982a] Automatic calculation of electron density profiles from digital ionograms, 3. Processing of bottomside ionograms. Radio Sci., Vol. 17, 837-844.

REINISCH, B.W. and HUANG, X. [1982b] Automatic calculation of electron density profiles from digital ionograms, 1. Automatic O and X trace identification for topside ionograms. Radio Sci., Vol. 17, 421-434.

REINISCH, B.W. [1986] New techniques in ground-based ionospheric sounding and studies. Radio Sci., Vol. 21, 331-341.

FIGURES

1a,b: Superposed diurnal variation of f_0F_2 at hourly intervals at Belgium (a) and Spain (b), March 1988. Monthly mean values are also indicated.

2a,b: Superposed diurnal variation of h'F at hourly intervals at Belgium (a) and Spain (b), March 1988. Monthly mean values are also indicated.

3a,b: Superposed diurnal variation of M(3000)F_2 at hourly intervals at Belgium (a) and Spain (b), March 1988. Monthly mean values are also indicated.

4a,b: Superposed diurnal variation of f_{min} at hourly intervals at Belgium (a) and Spain (b), March 1988. Monthly mean values are also indicated.

5a,b: Superposed diurnal variation of f_0F_1 at hourly intervals at Belgium (a) and Spain (b), March 1988. Monthly mean values are also indicated.

6a,b,c,d: Regression lines for hourly values of f_0F_2 (a), h'F (b), MUF(3000)F_2 (c), and f_{min} (d) at Belgium (ordinate) and Spain (abscissa) for March 3, 1988 (day 63).

7a,b,c,d: Correlation coefficients versus day of month for hourly values of f_0F_2 (a), f'F (b), MUF(3000)F_2 (c), and f_{min} (d) at Belgium and Spain for March 1988.

AD-P006 309

APPLICATIONS FOR HIGH ACCURACY DIGITAL IONOSONDE DATA

Adolf K. Paul
Ocean and Atmospheric Sciences Division, Code 542
Naval Ocean System Center, San Diego, CA 92118-5000

INTRODUCTION

The new technology used in modern digital ionosondes permits the measurement of traditional (virtual heights and amplitude of echoes) and new (radio phase of echoes) ionospheric data with very high precision. Consequently, higher accuracy for standard ionospheric parameters can be achieved and new types of parameters can be obtained using new processing methods. Details of such data analysis programs may depend on the type of digital ionosonde used, however, the basic physical principles involved are the same. For example, there is no doubt that the change of the radio phase with time is proportional to the Doppler frequency of the echo. In recent years much effort has gone into modeling of the ionosphere. Unfortunately the spatial and the temporal resolution of the most basic parameters of the data base for testing such models is inadequate. For example, it appears that in some areas (e.g., Europe) the spatial resolution of the F-layer maximum electron density may be sufficient, but this is not true for the height of the maximum and the half-thickness of the F-layer, since very few station computed electron density profiles from the recorded ionograms. In the following we will outline a new procedure for computing F-layer profile parameters. The process is simple and its routine application could significantly improve the data base. The accuracy limits of the resulting parameters will be discussed together with some other important ionospheric quantities observable with digital ionosondes.

ELECTRON DENSITY PROFILE PARAMETERS

Assuming that the electron density distribution around the peak of the F-layer is parabolic, it has been shown (Paul and Mackison, 1982) that the inverse of the slope of the virtual height trace h'(f) is approximately linear with frequency in the vicinity of the critical frequency. In a numerical peak fitting process the inverse of the slope is replaced by divided differences, the frequency interval divided by the corresponding virtual height difference. Linear extrapolation of such divided differences to a zero value using a least squares fit yields a very accurate estimate of the critical frequency foF2 or fxF2. It is easy to see that the slope of this linear fit is proportional to the reciprocal of the half-thickness Ym of the model parabola. Multiplying the half-thickness obtained in this way by the proper parabolic coefficient for a given frequency and subtracting this product from the corresponding virtual height then yields the height of the maximum Hm. This is expected to be true for every (frequency, virtual height) pair. In reality such residuals deviate from a constant value in a systematic way due to the difference between the linear and exact behavior of the divided differences mentioned above. In most cases only a very small change of the critical frequency is necessary to minimize the deviations of the residuals from a constant value.

The principle of this concept is illustrated in figure 1. The lower part, for which the left hand scale applies shows the divided differences as a function of frequency for the ordinary (o) and for the extraordinary (x) component. In the upper part of this figure for which the right hand scale applies, estimates of the height of maximum are displayed. Two sets of data are shown for each component. The ones with stronger curvature at the high frequency end were obtained as a result of the straight line least square fit. The other sets with almost constant heights resulted after a correction of the critical frequencies. For this particular ionogram the final critical frequency for the ordinary component was 12.868 MHz after a correction by +31 kHz, the final value for the extraordinary component was

91-09693

91 9 4 135

13.584 MHz after a correction by -55 kHz. The height of maximum derived from the ordinary data was 7.6 km less than the same quantity derived from the extraordinary data, the difference being due to the underlying ionization (F1 and E-region). If the virtual heights for each component are computed using the results of the peak fitting procedure, the agreement with the observed virtual heights is very close over a large portion of the ionogram. This is clearly visible in figure 2, where the dots show the observed and the solid lines the computed virtual heights. We also see here (between 4.0 and 7.2 MHz for the ordinary virtual heights) that the lowest portion of the F-layer differs significantly from our parabolic model. Preliminary results of computing the virtual heights for a correspondingly modified F-layer and comparing them with the observations verifies that the difference of Hm values for the two components is caused by the additional ionization in the lower part of the ionosphere and also shows that this height difference is approximately equal to the absolute error of the height of maximum derived from the ordinary virtual heights. In other words, at least at mid-latitudes, the difference between the Hm values for the two components can directly be used as a first order correction for the underlying ionization. The process described above permits a determination of the critical frequencies with an error of approximately 10 kHz or less, if accurate virtual heights from an undisturbed ionosphere are recorded. Under such conditions temporal changes of 20 kHz or more are significant as can be seen in figure 3. The accuracy is poorer when acoustic gravity waves with small vertical wavelength are passing through the ionosphere causing the electron density profile to differ significantly from a parabolic shape or when small scale irregularities, i.e. spread-F, are present.

Similarly the other two profile parameters Hm and Ym are also very well determined. The absolute values of these quantities, however, may need corrections for underlying ionization which can be done by a few representative model calculations as mentioned above. Note that the difference between the Hm values of the two components are practically constant over the entire period shown. Temporal changes of a few kilometers, less than the difference between the Hm values obtained for o- and x-components, can be significant as figure 4 illustrates. The half-thickness is a measure for the scale height of the ionosphere which increases with height. In figure 5 the Ym value obtained for the ordinary trace is plotted against the ordinary Hm values for 466 ionograms taken in 3 minute intervals covering almost a full day. The height dependence of Ym is clearly visible and a linear fit of the data yields a height dependence factor of 0.2 which means that the half-thickness changes by 20 km for a height increase of 100 km.

The frequency separation of ordinary and extraordinary components is caused by the earth's magnetic field and is related to the electron gyro frequency, fh, which is also height dependent. The accurate determination of ordinary and extraordinary critical frequency also permits the derivation of an accurate value of fh. Most data show a narrow spread of the individual values around the correct mean value, but only on extremely quiet days is the height dependence of fh measurable. This is not due to insufficient accuracy of the determination of the penetration frequencies, but to local deviations from horizontal stratification of the F-layer caused by acoustic gravity waves. In such situations the two modes can experience slightly different electron densities at the penetration points and the penetration frequencies are not only determined by the local plasma frequency, but also by the deviation of the propagation direction from vertical.

FIXED FREQUENCY PARAMETERS

One quantity which can be easily derived from phase measurements of the echoes is the angle of arrival or the apparent reflection point location (if the signal is received at 3 or more receiving antennas). Many recordings show that the apparent position changes with frequency and therefore with height, especially close to penetration. The temporal variability of the medium and its deviation from horizontal stratification mentioned above becomes even more visible with repeated fixed frequency observations. An example is shown in figure 6 of the variation of the apparent reflection point observed in the extraordinary mode. First the echo is seen in the southeast at the approximate coordinates (4.0,-10.0). The apparent reflection point then drifts toward the southeast to a distance almost 30 km away from the overhead point, returns close to the starting point, begins a loop, moves slowly toward west and then for approximately the last second toward the south.

These changes occurred in a continuous fashion within 19.6 seconds. The smallest significant distances over the sampling interval of 0.1 seconds are of the order of a fraction of one kilometer corresponding to 0.1 degree of angular change seen from the ground.

Figure 7 shows an example of Doppler frequency data derived from the consecutive differences of the phase of an echo. The small scatter of the data indicates that the accuracy of individual values is better than 0.01 Hz. The data also show that the Doppler frequency can change quickly and that well defined trends of such changes may only last for tens of seconds. Simultaneously with the phase measurement used for the Doppler data shown in this figure, the phase was recorded on a second frequency 8 kHz higher. This change of phase with frequency can be used to obtain an improved estimate of the group (virtual) height. Figure 8 shows the result of this procedure for the same time interval and the same frequency of the previous figure. Again the noise is very low and errors of the virtual heights average less than one wavelength. Note that the virtual height changed by more than 10 km over the time interval of five minutes.

All three quantities discussed here, including their frequency dependence, can also be obtained during regular ionogram recording. The longer time series examples show however that an individual value may not be representative for a longer period of time and that significant changes can already take place during the recording time of an ionogram in which case the separation of time and frequency dependence may be incomplete.

DISCUSSION AND CONCLUSIONS

The two oscillations of foF2 as seen in the center and the right side of figure 3 are clearly visible, but relatively small. The amplitude amounts to approximately 100 kHz which corresponds to a relative amplitude of 0.7% in critical frequency or 1.4 % in maximum electron density. Most likely these oscillations are coupled with equal relative changes of the neutral density caused by weak acoustic gravity waves. While there are no corresponding variations of Hm or Ym noticeable in figure 4 during the first oscillation of foF2, periodic chances of these quantities are visible during the second one. The maximum of Ym shortly after 1400 hour indicates an expansion of the F-layer coupled with lower electron density, while the minimum of Ym, approximately 20 minutes later, is consistent with a compression of the layer and appears at a time when fof2 shows a relative maximum. There also is a periodic variation of Hm visible in both components during this last period, but no significant variation of Hm or Ym exists during the earlier period. No explanation for the different behavior could be found. This is one example which demonstrates the sensitivity of digital ionosondes. It also shows that the data obtained can be used to study some properties of the near earth environment and its dynamic behavior besides the parameters important for Hf propagation.

The vertical gradient of the thickness parameter Ym (or scale height), if derived routinely from data as shown in figure 5 could be very useful for ionospheric and atmospheric modeling, especially if comparisons with satellite data should show that the linear variation of Ym with height is valid beyond the height of the F-layer maximum. Figure 5 also gives an indication that because of the high variability of the medium, a large number of samples must be recorded and processed in order to establish relations such as the height dependence of the scale height. Hourly data for the 24 hours of one day are, in most cases, not sufficiently representative to study short term variations of the order of days or weeks, e.g., the response of the ionosphere to changing solar output as indicated by the daily sunspot numbers or 10 cm flux.

The seasonal variation of Ym is expected to be a combination of height and temperature changes. Again, if a sufficiently large number of data are recorded and the short term fluctuations can be eliminated it should be possible to separate the seasonal temperature variation from the change of the height of maximum.

Figure 9 shows an example of the complexity of the F-region on a normal day with no magnetic or solar disturbances. The height of the maximum Hm (left hand scale, crosses) is compared with the critical frequency foF2 of the ordinary mode (right hand scale, dots). The critical frequency is approximately constant for the first five hours of the day, while shortly after midnight the height suddenly increases by 40 km. With the beginning of sunrise the height then drops rapidly by

more than 100 km and foF2 increases sharply. The height reaches a minimum approximately one hour after sunrise, starts increasing again and reaches a maximum in the afternoon at about 1500 hours while foF2 reaches its maximum earlier, near noon. In the late afternoon both the height and the critical frequency decrease until sunset when suddenly the height increases again and the critical frequency continues to decreases, but at a distinctly slower rate. The gravity wave activity is clearly visible as strong oscillations superimposed on Hm, while foF2 is less affected. The long and short period variations of Hm clearly demonstrate that the error in Hm due to underlying ionization discussed earlier is relatively minor. At the same time it is obvious that the best possible computed profile may only be representative for a very short period of time and for a relatively small area.

Coordinated multi-station observations are necessary in order to separate global variations (e.g., solar input) from local changes (e.g., acoustic gravity waves). It appears, that the accuracy of digital ionosondes is now sufficient for monitoring of several parameters describing changes in that part of the near earth environment which cannot be routinely observed by satellites or any other method with a sampling rate high enough in time and space for a correct description of the dynamic processes taking place.

References:

Paul, A. K. and D. L. Mackison, Scaling F-layer critical frequency from digital ionograms applied to observations during the solar eclipse on 26 February 1979, J. Atmos. Terr. Phys., 43,221-223,1981.

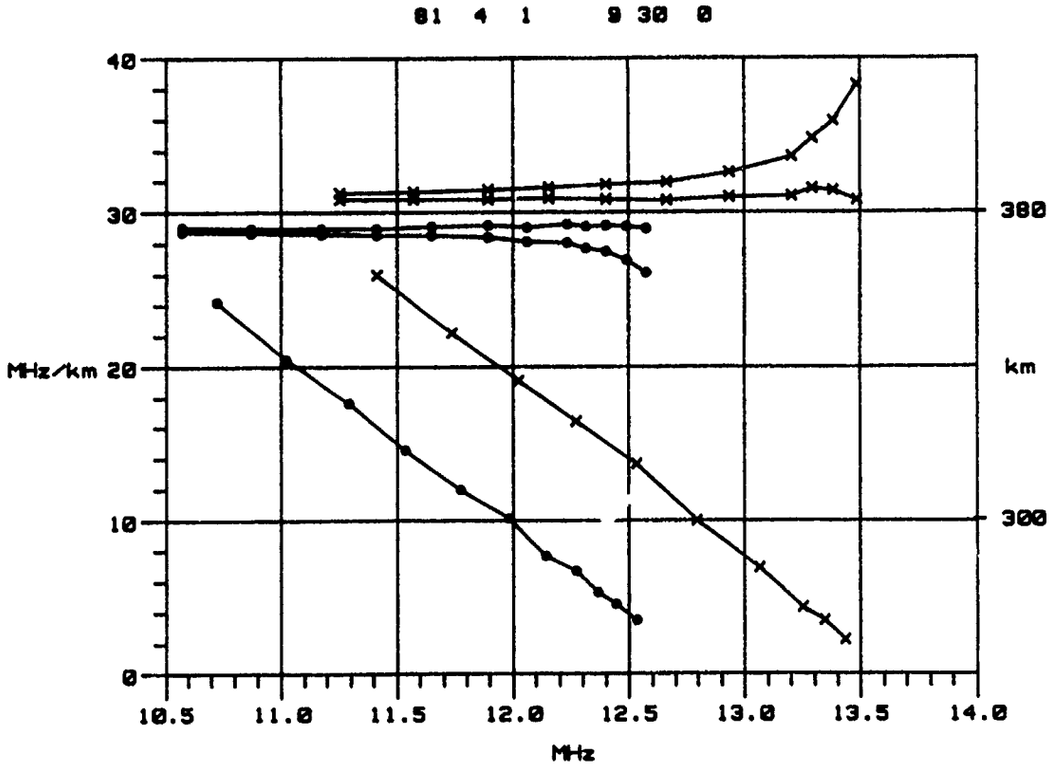


Fig. 1. Divided differences Df/Dh' (left scale) and estimates of the height of the F-layer maximum (right scale) for both components.

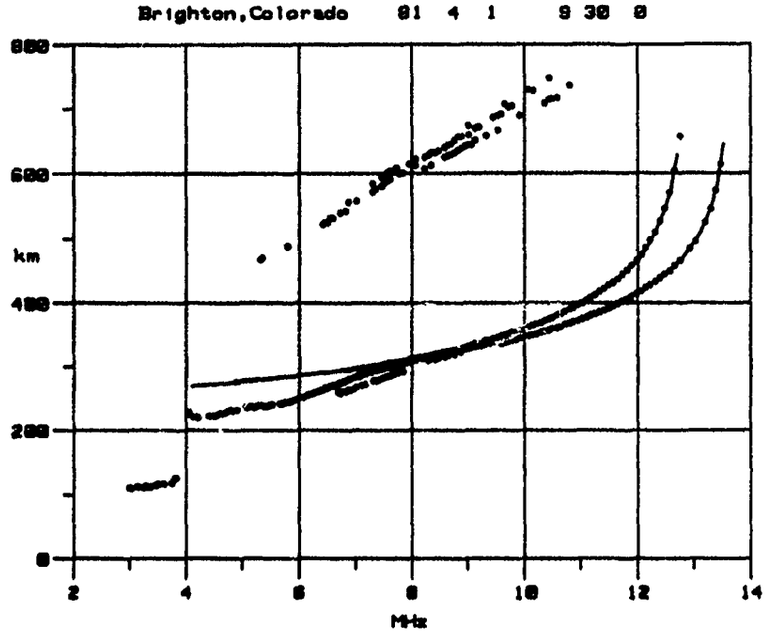


Fig. 2. Comparison of the observed virtual heights (dots) with those computed from the parabolic peak fit (solid lines).

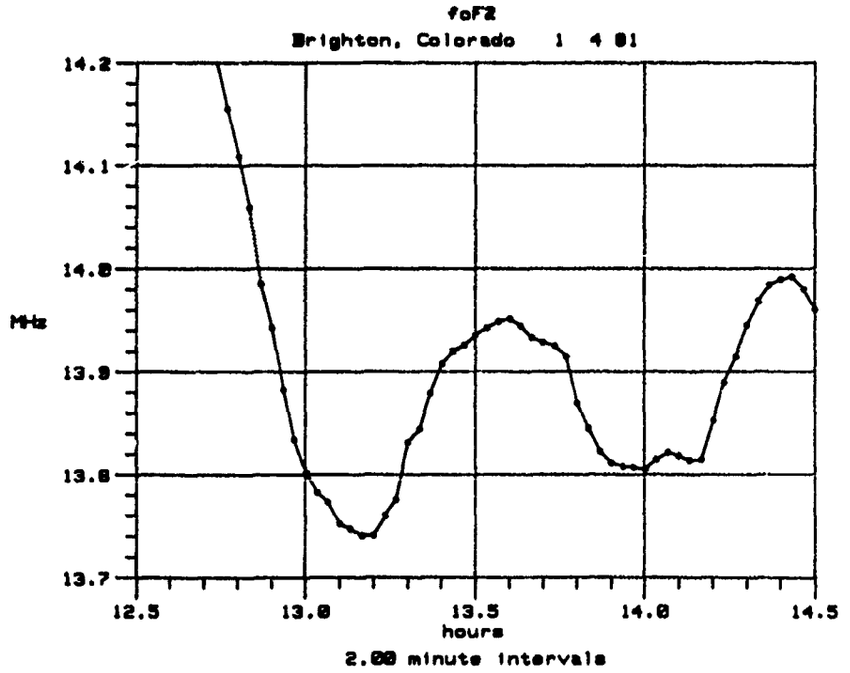


Fig. 3. High resolution (in time and frequency) estimates of the critical frequency foF2 showing small scale periodic variations.

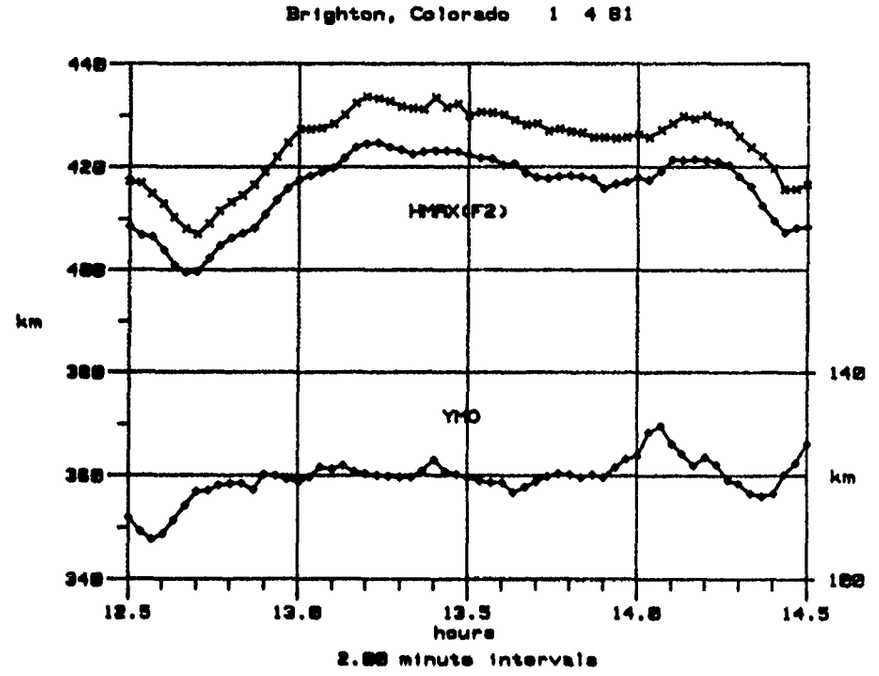


Fig. 4. Estimates of the height of maximum derived from ordinary (o) and extraordinary components (x) (left scale) and of the (ordinary) half-thickness (right scale).

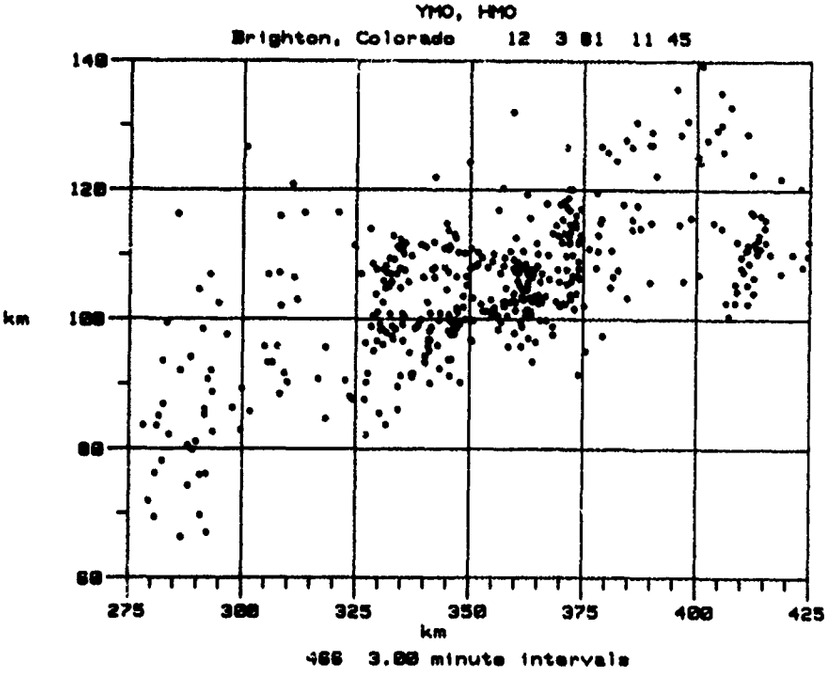


Fig. 5. Scatter plot of half-thickness versus height of maximum.

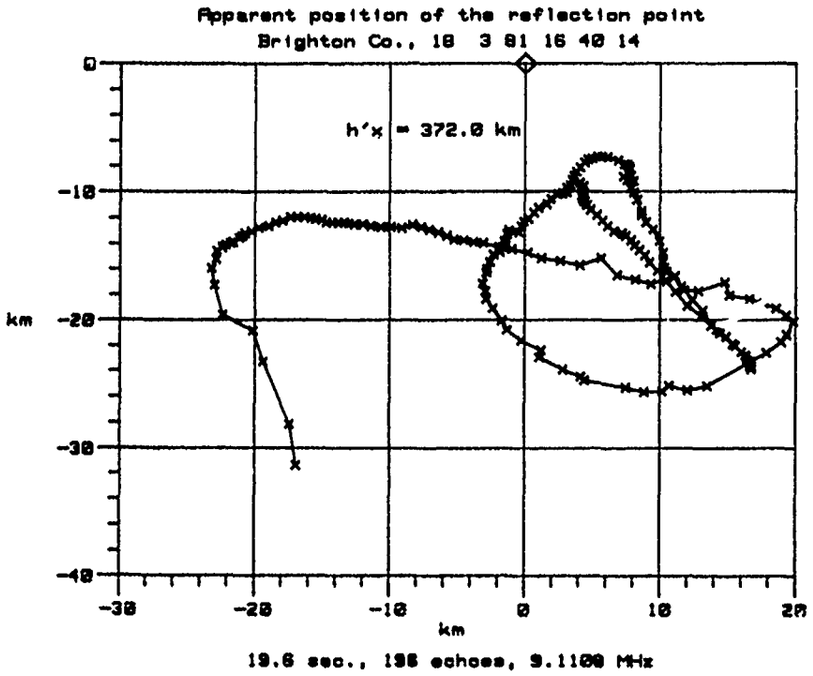


Fig. 6. Variation of the apparent reflection point position with time for a fixed frequency.

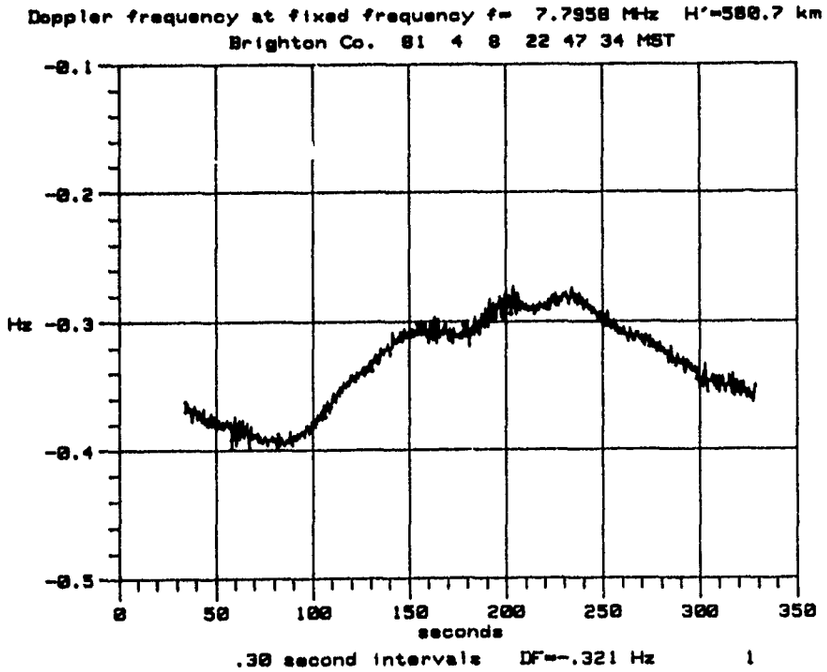


Fig. 7. Doppler frequency derived from phase measurements at a fixed radio frequency.

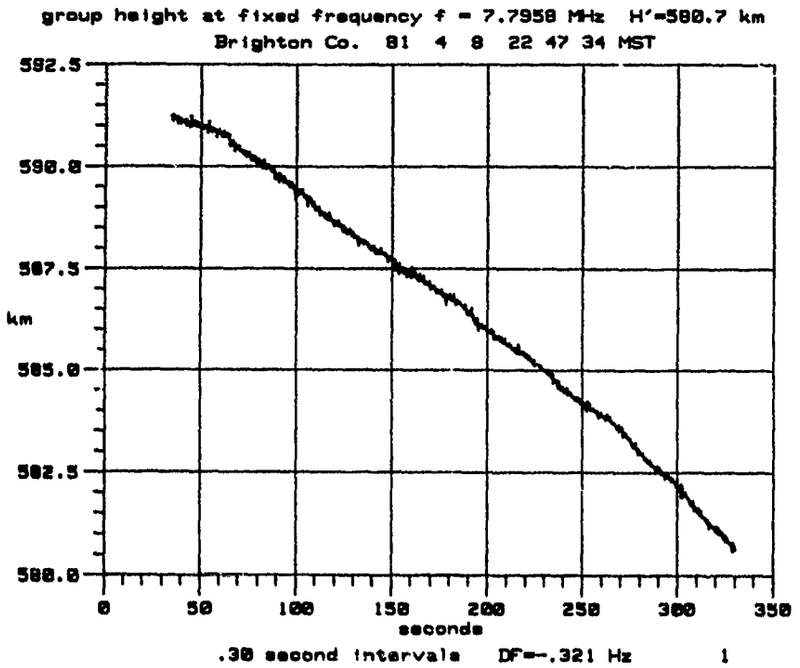


Fig. 8. Virtual height change with time for a fixed frequency.

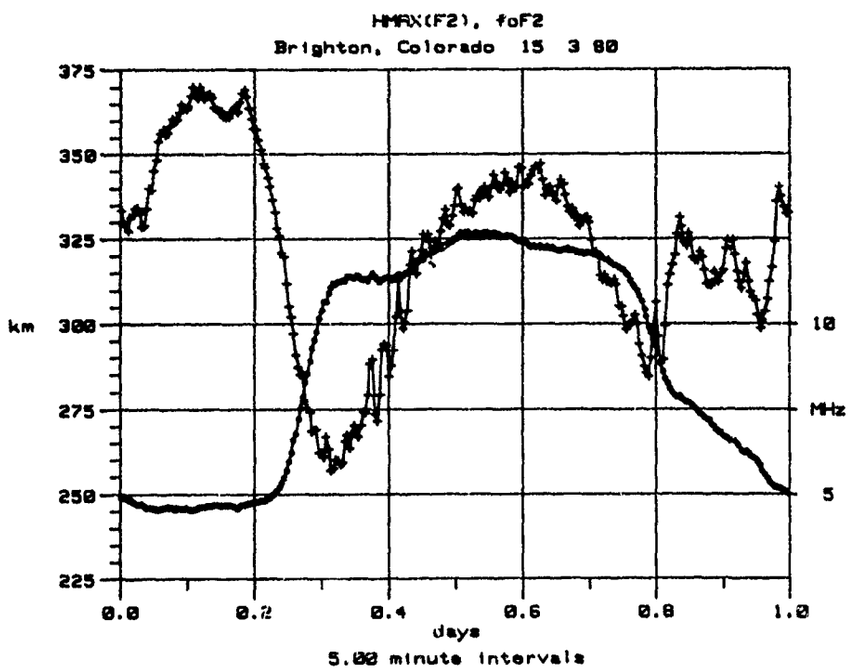


Fig. 9. Diurnal variation of the height of maximum (left scale) and the ordinary critical frequency (right scale).

AD-P006 310

OBLIQUE-INCIDENCE SOUNDER MEASUREMENTS with ABSOLUTE PROPAGATION DELAY TIMING

Mark Daehler
Ionospheric Effects Branch, Space Science Division
E.O. Hulburt Center for Space Research
Naval Research Laboratory, Washington, DC 20375-5000

ABSTRACT

Timing from the Global Position Satellite (GPS) system has been applied to HF oblique-incidence sounder measurements to produce ionograms whose propagation delay time scale is absolutely calibrated. Such a calibration is useful for interpreting ionograms in terms of the electron density true-height profile for the ionosphere responsible for the propagation. Use of the time variations in the shape of the electron density profile, in conjunction with an HF propagation model, is expected to provide better near-term (1-24 hour) HF propagation forecasts than are available from current updating systems, which use only the MUF. Such a capability may provide the basis for HF frequency management techniques which are more efficient than current methods. Absolute timing and other techniques applicable to automatic extraction of the electron-density profile from an ionogram will be discussed.

INTRODUCTION

In this paper I plan to describe some techniques which we are developing to facilitate the automatic interpretation of ionograms from an oblique-incidence sounder. Among these techniques, and one which we think will be an important aid in identifying propagation modes, is the ability to measure absolute propagation delay time. We feel that it will eventually be possible to automatically interpret an ionogram in terms of its mode structure, and then calculate the true-height electron density profile of the ionosphere which is responsible for the propagation.

The reason for wanting this capability is to provide improved frequency management for HF communication systems. Despite the successes of billion-dollar satellite communication systems, relatively simple HF systems using skywave ionospheric propagation have retained considerable importance for both DoD and civilian operations. In fact, HF applications have expanded in the last decade to the point of a severe limitation in available frequencies. Contributing to the dearth of available frequencies is the tendency to provide users with multiple frequencies, the traditional cure for the uncertainties in propagation resulting from a variable and only partially predictable ionosphere. With more crowded HF channels, it is essential to allocate frequencies more efficiently. This means that fewer frequencies will be assigned to a given user, and reassignments will be made more frequently. In return, there must be greater assurance that an assigned frequency will propagate. This more refined type of frequency management virtually requires an automated system.

One of the keys to more efficient frequency management is better forecasting of propagation conditions. Propagation models based on monthly mean ionospheric parameters can serve as a first estimate, and they are in widespread use for frequency management. They are good at representing seasonal and diurnal variations. They cannot, however, take into account those fluctuations, with time of a few hours to a few days, which result from the unpredictability of solar emissions. Oblique-incidence sounders provide an excellent assessment of current HF propagation on a single path. Their usefulness can be extended to "nearby" paths, meaning 1-hop paths whose midpoints are within a few hundred kilometers from the sounded path midpoint. When combined with a propagation model, the sounder results can be used to "adjust" the model forecast, thus permitting forecasts in which the diurnal variations are determined by the model and the unpredictable solar fluctuations are determined by the sounder. Compared with propagation model predictions, this scheme of "updating" provides significantly-improved forecasts of propagation for times of 6-8 hours, or possibly up to 24 hours, in advance (Reilly and Daehler 1986).

91-09692

Current updating schemes typically use the Maximum Observable Frequency (MOF) from the sounder, and calculate an effective sunspot number or effective 10.7-cm. flux which makes the model yield the measured MOF. Even this small amount of information extracted from the sounder ionogram can provide a useful improvement in forecasting. But there is clearly a lot more information about the ionosphere available in the ionograms. In particular, previous work (Reilly 1989) has indicated that it is possible to extract directly from an oblique-incidence ionogram many of the features of the ionospheric electron density true height profile. The electron density profile is a much more fundamental physical property than the MOF, and is the ionospheric property used directly by propagation models. It seems likely to us that it should be possible to use the sounder results to update predictions of the electron density profile, and then use the forecasted electron density profiles to predict propagation at future times. This is a substantially more-complicated procedure, but it also seems reasonable to expect better results from this more fundamental approach.

Our point of attack, then is to attempt to extract from an oblique-incidence the electron density profile, which can be described in a number of ways, but always with a number of parameters (on the order of 6 to 12). The time variation of each of the parameters can then be tracked, and predictions of propagation can be based on forecasts of each of the ionospheric electron-density parameters. If forecasting based on updated electron density profile predictions is to be useful, it is clear that a highly-automated analysis system will be required. The work I am describing here is a collection of techniques which we expect to be useful in an automated version of this electron density profile prediction method.

Automatic interpretation of vertical-incidence sounder data is well-developed (Bibl and Reinisch, 1978; Fox and Blundell, 1989), and an automatic method of calculating the electron density profile has been reported (Reinisch and Xueqin, 1983). Automatic interpetation for oblique-incidence data is not so well developed, but has received attention (Lomax, 1969).

ABSOLUTE PROPAGATION DELAY TIME MEASUREMENTS

The first method we thought would be useful is the addition of absolute propagation delay time calibration to the sounder ionograms. A number of researchers have done this before (Rose, 1988; Iijima et al., 1960; Iijima et al., 1959). However, with the relatively-recent advent of the Global Position Satellite (GPS) system, obtaining absolute time calibration has become much easier. Figure 1 shows the geometry. The oblique-incidence sounder is a part of the Barry Research Corporation AN/TRQ-35 system. The sounder transmitter emits a tone of continuously-increasing frequency; the receiver passband is swept in synchronism with tne transmitter sweep, but with a delay that is just slightly less than the minimum propagation delay time. The procedure is to use a GPS clock at each end of the path, to measure both the moment at which the transmitter sweep starts and the moment at which the receiver sweep starts. The difference between these two times is added to the relative times shown on the sounder display to produce the absolute calibration. The absolute time measurement produced by each GPS clock is accurate to +/- 250 nsec, with the probability that the difference in two such measurements is somewhat smaller. In any case, we expect the accuracy of the propagation time delay measurements be better than +/- 500 nsec. Fig. 2 shows an example of an ionogram with absolute time calibration.

This ionogram shows propagation over the Peterson AFB (Colorado) to Skaggs Island (California) path, a distance of 1541 km. This is one of a series of measurements made every five minutes over a period of about three weeks in May and June 1988. The occasion was a measurement of ionospheric conditions during a test of range and direction-finding equipment. The ionograms shown in this work were all taken on one day, June 5th, of this exercise. The data from the sounder are extracted in analog form, digitized, and recorded on magnetic tape. They are thus available, in near-real-time, in an easily-manipulated form.

EXTRACTION OF ELECTRON DENSITY TRUE HEIGHT PROFILES

One way to determine the electron-density profile which yields a particular ionogram is to assume a likely profile, and then calculate the ionogram which such a density profile would produce. Such a calculation is possible using a technique developed by Reilly. His computer program can accept an ionosphere described in any of several ways, one of them being a sum of parabolic layers. If the calculated ionogram is the same as the measured one, then one can provisionally adopt the assumed electron-density profile as a description of the actual ionosphere. If necessary, the parameters of the assumed profile can be adjusted until the calculated ionogram agrees with the measured ionogram. The examples shown in this work have been produced, manually, by making repeated guesses at which parameter should be adjusted to make the calculated and measured ionograms agree. But we have tried also to keep in mind the possibility of producing an automatic method for doing the fitting.

An approximate electron density profile for starting the fitting process can be obtained from models of mean-value ionospheric parameters. We have used the IONCAP model, which uses the CCIR-recommended ionospheric coefficients, from which a representation of the ionosphere is calculated in the form of a series of parabolic layers.

The ionogram shown in Fig. 2 was chosen because of its simplicity. To start the procedure for determining the electron-density profile of the ionosphere responsible for this ionogram, IONCAP was run with a sunspot number of 120, a plausible value (NOAA, 1988) for the day of the measurement. The ionosphere used by IONCAP for this situation was the sum of two parabolic layers. The ionogram calculated for this ionosphere by Reilly's program is shown, along with the measured ionogram, in Figure 3. In considering the ionogram calculated from the IONCAP ionosphere it must be remembered that no account has been taken of amplitudes, so that it can be expected that some of the features appearing in the calculated ionogram will not appear in the measured data. For example, the measured ionogram shows a large amount of sporadic-E propagation which is not predicted by IONCAP, and the measured ionogram does not show most of the low-frequency E-layer propagation predicted by IONCAP. While there is a clear qualitative correspondance between the 1-hop and 2-hop F modes, the quantitative agreement is not great.

Starting with the ionospheric parameters predicted by IONCAP, adjustments were then made to the ionospheric parameters to try to make the calculated ionogram agree better with the measured ionogram. This is a rather tedious process, but after about twenty attempts the result shown in Figure 4 was obtained. Attempts using only two layers, like IONCAP, weren't very successful at producing ionograms which simultaneously approximate both the high- and low-ray F2 measured traces. Thus a rather weak F1 layer was added, which produced a substantially better result. The presence of an F1 layer is not unreasonable, since the ionogram was made very early in the morning, about 2:15 AM local time, when an F1 layer could start to develop. For this path and time, IONCAP introduces an F1 layer for this path starting approximately 4:00 AM local time, two hours after the time of the measured ionogram. It is clear that the ionogram calculated from the adopted ionosphere leads to a better approximation to the measured ionogram than the one calculated from the IONCAP ionosphere. The adopted ionosphere is compared with the IONCAP ionosphere in Figure 5.

There is of course no guarantee that the adjusted ionosphere is unique. It seems likely that the magnitude of the uniqueness problem will become clearer when the fitting program has been automated and it is possible to follow the changes in the calculated ionosphere over the course of a day.

A somewhat more complicated ionogram, one showing a substantial F2 layer, is shown in Figure 6. Also plotted in this figure is an ionogram calculated from the three-layer ionosphere used by IONCAP. Here, again, there is a clear discrepancy. By adjusting some of the assumed layer parameters, however, a much more reasonable fit can be obtained (Figure 7). A comparison of the two electron density profiles is shown in Figure.

Also shown in Figure 8 is an electron density profile calculated directly from the measured ionogram. This technique, also developed by Reilly (1989), is able to calculate the electron density true height profile by direct inversion of the frequency-time delay ionogram data for the 1-hop E and F layers.

Starting with the minimum path delay, this program calculates the electron density gradients required for propagation at successively-increasing values of path delay. The electron density profile is thereby gradually constructed as a series of connected parabolic and linear segments. The application of this program to arbitrarily-selected ionograms isn't always completely straightforward, particularly because it isn't always possible to find in the ionogram signs of propagation at the very lowest delay times. This can easily result from experimental factors like high absorption, antenna selection, or other conditions which make propagation at lower frequencies impossible. Nevertheless, the fact that the directly-calculated profile in Figure 8 agrees much more closely with the profile determined by adjusting its ionogram to look like the measured ionogram than to the IONCAP profile indicates that this is a potentially useful technique. It seems quite likely that, in an automated system, direct calculation of the electron density profile could be useful in calculating at least some part of the entire profile.

Figure 9 shows an ionogram which looks very much like the one in Figures 6 and 7, except for the presence of an additional mode in between the 1-hop and 2-hop F2 modes. This can be interpreted as a mixed mode, in which one hop is a reflection from the E layer, and the other is from the F2 layer. This interpretation is based on the observation that the time-delay interval between the E-layer reflection and the mixed mode is almost exactly equal to the

interval between the mixed mode and the 2-hop F2 reflection. An aid to this analysis is provided by the side plot in Figure 9, which shows the amplitude of the reflected signals as a function of delay time, for the frequency 7.95 MHz, a kind a ionogram "cross section".

DISCUSSION

This work has shown that oblique-incidence sounders can be operated routinely with absolute propagation-delay timing and that, at least in some cases, ionograms so calibrated can be interpreted in terms of the electron density profile of the ionosphere responsible for propagation. Much work remains to determine the uniqueness of these interpretations, and the extent to which such interpretations agree with measurements by other methods, such as vertical-incidence sounders.

This capability indicates the possibility of improving short-term (1-24 hour) HF propagation forecasting by updating long-term average models (such as IONCAP) with sounder data which directly describes fluctuations in ionospheric morphology. This contrasts with current updating methods in which only the MOF is extracted from the ionogram.

It is clear that these methods can lead to practical applications only if they can be automated. A substantial part of the problem rests in developing usable pattern-recognition routines. Such development will be aided by the facts that the ionosphere changes relatively slowly, compared with the speed at which ionograms can be produced, and that propagation models can provide good starting positions and reasonable parameter limits for the calculated electron-density profiles.

ACKNOWLEDGEMENTS

I am indebted to M. H. Reilly for making his ionogram and electron-density profile inversion programs available, and to both him and J. M. Goodman for many useful discussions. P. Hoover has contributed much of the effort to produce the calculations and illustrations.

REFERENCES

Bibl, K., and B. W. Reinisch (1978), The universal ionosonde, Radio Science,13(3), 519-530

Fox, M. W., and C. Blundell (1989), Automatic scaling of digital ionograms, Radio Science, 24(6), 747-761.

Iijima, S., K. Futsukaichi, and G. Shibutani (1959), Precise measurement of travel times of radio signals between Europe and Japan, Tokyo Astron. Bull., 121, 1409-1495.

Iijima, S., G. Shibutani, and S. Katahara (1960), Precise determination of travel time of radio signals between Hawaii and Tokyo, Tokyo Astron. Bull., 128, 1555-1604.

Lomax, J. (1969), Automated analysis of oblique-incidence sounder data, AGARD Conference Proceedings, #13, 503-524, published by Technivision, Maidenhead, England.

NOAA, Solar-Geophysical Data Prompt Reports (1988), No. 530, Part I, 12.

Reilly, M. H., and M. Daehler (1986), Sounder updates for statistical model predictions of maximum usable frequencies on HF sky wave paths, Radio Science, 21(6), 1001.

Reilly, M. H., and J. D. Kolesar (1989), A method for real height analysis of oblique ionograms, Radio Science, 24(4), 575-583.

Rose, R. B., High-resolution HF time of arrival measurements (1981-1985), Radio Science, 23(3), 257, 1988

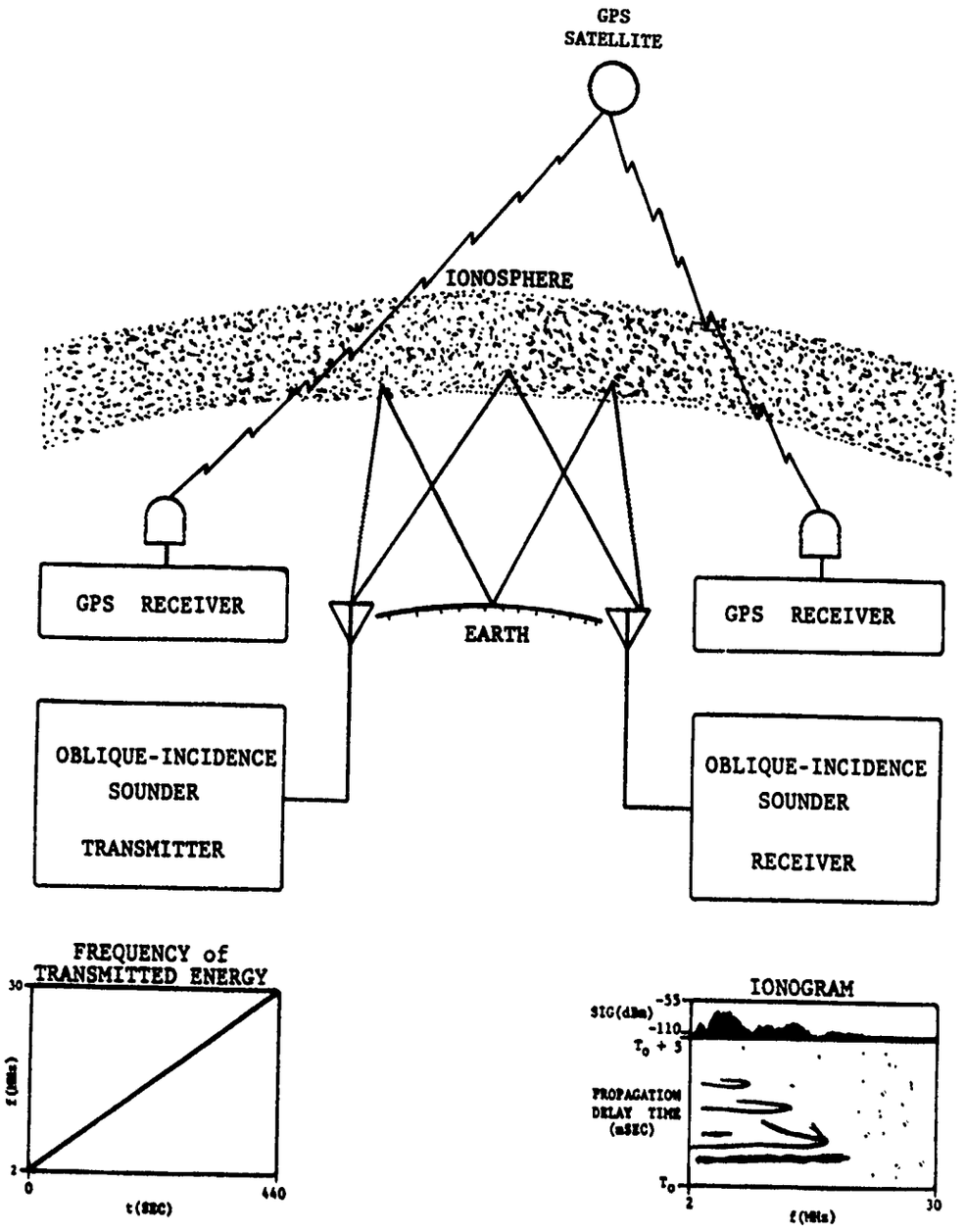


Fig. 1. Geometry of absolute propagation-delay time measurement.

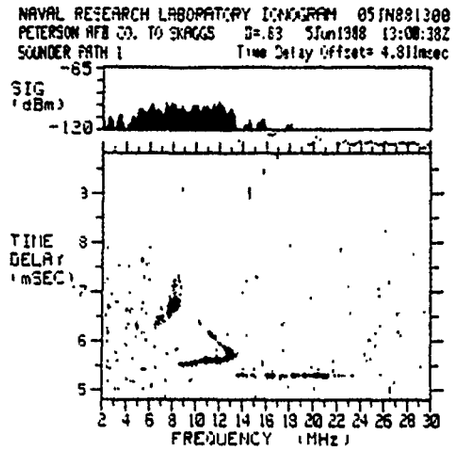


Figure 2. An oblique-incidence ionogram, with absolute time calibration, made on 1988 June 5 1300Z, showing propagation on the Peterson AFB, CO to Skaggs Island, CA path, a distance of 1541 km.

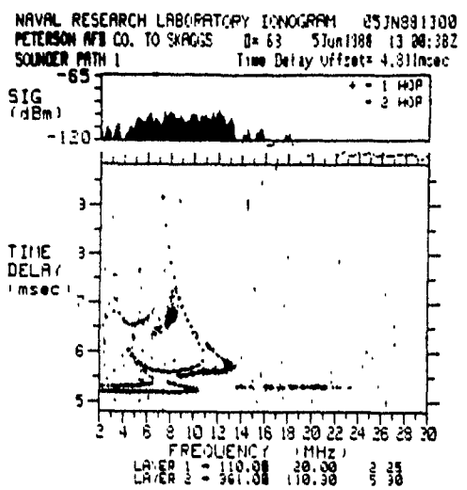


Fig. 3. The measured ionogram of Fig. 2 plotted with the ionogram calculated from the electron density profile used by IONCAP (+ = 1-hop, x = 2-hop). The IONCAP profile is the sum of two parabolic layers whose parameters (height[km] of maximum density, half-width [km] at zero altitude, and maximum density in units of plasma frequency squared [MHz^2]) are listed below the plot.

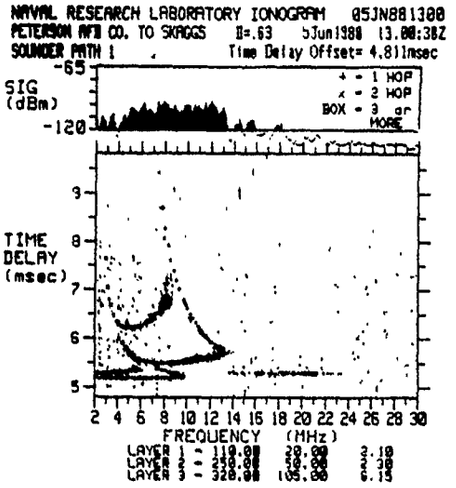


Fig. 4. The measured ionogram of Fig. 2 plotted with the ionogram calculated from an electron density profile whose parameters were adjusted to make the calculated ionogram approximate the measured ionogram. Three parabolic layers (E, F1, and F2) were used; their parameters (height [km] of maximum density, half-width [km] at zero altitude, and maximum density in units of plasma frequency squared [MHz^2]) are listed below the plot.

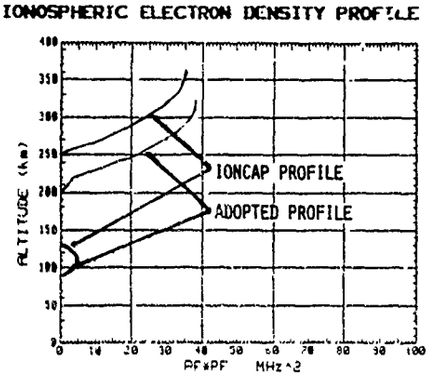


Figure 5. Electron density (in units of plasma frequency squared) as a function of altitude on 1988 June 5 1300Z for the midpoint of the Peterson AFB, CO to Skaggs Island, CA path. Curves show a) the profile used in IONCAP, which results in the calculated ionogram shown in Fig. 3, and b) the adopted profile, which results in the ionogram of Fig. 5, which was calculated to approximate the measured ionogram.

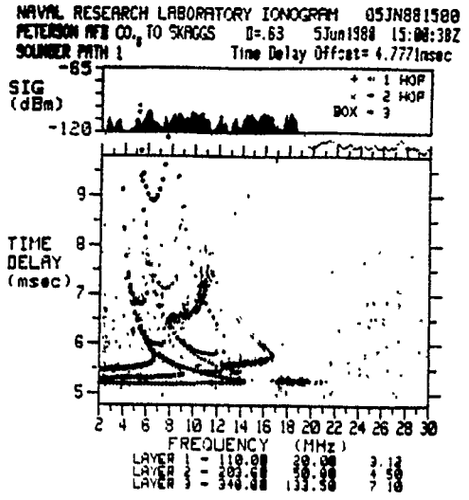


Fig. 6. The ionogram measured on 1988 June 5 1500Z over the Peterson AFB, CO to Skaggs Island, CO path, plotted with an ionogram calculated from the electron density profile used by IONCAP (+ = 1-hop, x = 2-hop, box = 3-hop). The IONCAP profile is the sum of three parabolic layers whose parameters (height [km] of maximum density, half-width [km] at zero altitude, and maximum density in units of plasma frequency [MHz^2]) are listed below the plot.

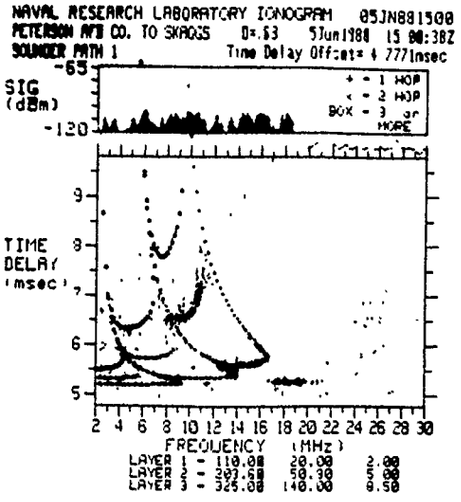


Fig. 7. The ionogram described in Fig. 6 plotted with an ionogram calculated from an electron density profile consisting of three parabolic layers, as in the IONCAP profile, but with layer parameters adjusted to make the ionogram calculated with the adopted profile approximate the measured ionogram.

COEFFICIENT EQUATION: (R+FN)^2=A+R^2+B+R+C

HL(Km) = Hgt at Lower Boundary of Profile Segment,
HU = Hgt at Upper Boundary of Profile Segment,
A(F^2)=A, B(F^2+L)=B, C(F+L)^2=C in Coefficient Equation
05JN88I500

HL(Km)	HU	A(F^2)	B(F^2+L)	C((F+L) 2)
0.0000	90.0000	+0.000000000E+00	+0.000000000E+00	+0.000000000E+00
90.0000	130.1242	-1.015947824E+06	+1.316924825E+10	-6.267626588E+13
130.1242	152.9000	+0.000000000E+00	+0.000000000E+00	+0.000000000E+00
152.8000	233.8494	-3.339638315E+05	+4.391754448E+09	-1.443745357E+13
233.8494	350.0000	-1.226549589E+05	+1.649168131E+09	-5.541178041E+12
LAYER 1 = 110.00	20.00	3.12		
LAYER 2 = 203.60	50.00	4.50		Ioncap Predictions
LAYER 3 = 340.00	133.50	7.10		

IONOSPHERIC ELECTRON DENSITY PROFILE

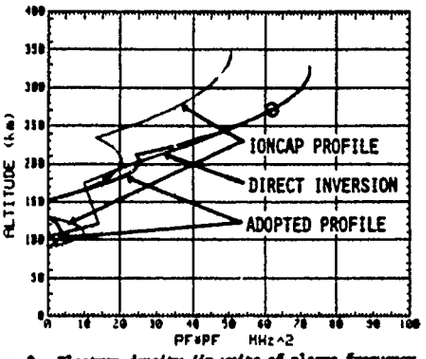


Figure 8. Electron density (in units of plasma frequency squared) as a function of altitude on 1988 June 5 1500Z for the midpoint of the Peterson AFB, CO to Skaggs Island, CA path. Curves show a) the profile used by IONCAP, a sum of three parabolic layers; b) the adopted profile, also a sum of three parabolic layers; and c) a profile calculated by direct inversion of the measured E- and F-layer data from the ionogram. The circle marks the end-point of the direct-inversion data.

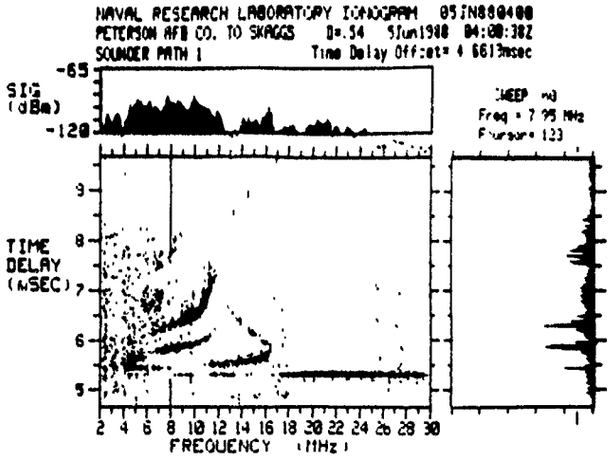


Fig. 9. An ionogram similar in appearance to the ionogram in Figs. 6 and 7, but having an additional mode between the 1-hop and 2-hop F2 modes. (This mode can be described as the propagation roughly following the line 8 MHz, 5.8 msec to 11MHz, 6 msec.) The additional mode is interpreted as a mixed 2-hop mode, with one reflection from the E layer and one from the F2 layer. The side plot records, for the frequency of the cursor position (7.95 MHz), the amplitude of propagation as a function of time delay. It shows, for this frequency, that the time delay interval between the 2-hop E mode (a horizontal line lying about 0.1 msec above the sporadic E line) and the mixed mode is very nearly equal to the time-delay interval between the mixed mode and the 2-hop F2 mode.

AD-P006 311

PROBING THE IONOSPHERE WITH THE UNIVERSITY OF ILLINOIS SOUNDER

K.H. Lin, W.S. Schlegl, K.S. Yang, and K.C. Yeh

Department of Electrical and Computer Engineering, University of Illinois at Urbana-Champaign

ABSTRACT

The University of Illinois has recently deployed an HF oblique sounder, the purpose of which is to probe the ionosphere for wideband response. The sounder is capable of operations in two modes: ranging mode and probing mode. Under the ranging mode, group delay measurements are made using phase modulated PN codes. Such measurements can assist the determination of propagation mode and help the data interpretation. Under the probing mode, a comb of CW signals across the band of interest are sent to probe the ionospheric response. The transmitter is located at Platteville, Colorado and toward the east at approximately one-hop distance is the receiver located at Urbana, Illinois. Within the frequency range of 5 to 15 MHz, the transmitter is designed to synthesize up to six discrete frequencies selectable among a possible eight hundred each separated from its neighbor by 12.5 kHz for probing measurements. Up to a total power of 200 W can be fed to a transmitting antenna which is a log-periodic array consisting of 18 elements above a ground screen. The major lobe at an eastward launch angle of about 22 degrees has a gain of 13.5 dBi. The receiver consists of six independent channels, each coherently heterodynes the received incoming signal to a standard IF and then is I-Q detected and recorded on the magnetic tape for processing. All operations are controlled with PC based software. Remote control and operations of the PCs are carried out via commercial telephone circuits. Local clocks are synchronized using the geostationary satellite GEOS time signals. This system has been used to measure the statistical properties of the ionospheric transfer function which is both frequency selective and time varying. Properties such as symmetry conditions, dispersive bandwidth, coherence bandwidth, and coherence time are experimentally measured for ionospheric conditions.

1. Introduction

Wave propagation through the ionosphere has been written about in many books [e.g., Rydbeck, 1942; Ginzburg, 1964; Davies, 1969; Yeh and Liu, 1972; Budden, 1985]. Recently there has been interest in propagating wideband signals through the ionosphere [Wagner and Goldstein, 1985]. However, being a dispersive medium, different frequency components of a pulse will reach the receiver with different delay and attenuation. This distorts the pulse shape, and more importantly, the pulse width is lengthened which in turn produces intersymbol interference. In addition to the dispersive background, the ionosphere is also embedded with irregularities which can only be described statistically. The scattering of waves from these irregularities further deteriorates the radio signals.

To find a remedy, the characteristics of the channel transfer function have to be determined. Although theoretical work has been done to obtain some numerical results [Wagen and Yeh 1989 a, b and c], they have to be verified from the experimental measurements. The HF sounder put into operation by the University of Illinois is designed for this purpose. Its operations and capabilities are described in Section 2. Some measured results are presented in Section 3. Finally, the paper is concluded in Section 4.

2. Description of the System

The probe consists of separate transmitter and receiver. The transmitter and its supporting electronics are housed in a trailer located in a field site in Platteville, Colorado, while the receiver resides at a farm house south of the University of Illinois campus in Urbana, IL.

The transmitter is a wideband 200 watt power amplifier covering the HF frequencies of interest (5-15 MHz). It is driven with 6 coherently generated frequency synthesizers, each capable of synthesizing any one of 800 evenly spaced discrete frequencies, i.e. every 12.5 kHz, within the 5 to 15 MHz frequency range. The transmitting antenna is a log-period dipole fed with open-wire line coupled to the transmitter with balun transformer, over a wire mesh ground screen. Theoretical gain of 12 dB can be realized at launch angles between 18 and 26 degrees. The dipole elements are supported by two 50 m and two 13 m towers.

Toward east from Platteville about one hop distance (1401 km) away, the receiver with its receiving antennas is capable of receiving up to 6 different frequency channels simultaneously. Due to site space limitation, a relatively small (12×30m) and low (18m) horizontally polarized short dipole array is used for receiving frequencies 10 to 15 MHz. It is phased to produce vertical null planes in the N-S direction to reject signals from the City of Urbana directly north. The receiving antenna for 5 to 10 MHz is 4 dipoles suspended between two 30 m towers.

All the receiver channels are identical and linear from -50 dBm to 100 dBm. The first LO of each channel is coherently generated for 6 channels and translates the input signal to a standard IF of 60 MHz. It is then further translated to its baseband

91-09691

of 50 Hz and I-Q detected. The I-Q outputs are integrated over 8 ms intervals, digitized by a PC and logged onto magnetic tape to be processed off-line.

Two modes of operations are possible. In the probing mode, the complex field strength at the receiver due to a transmitted CW is measured up to a bandwidth of 50 Hz. In the ranging mode, total transmission time delay between the transmitter and the receiver is measured with 10 μsec resolution.

The ranging mode uses maximal linear codes [Dixon, 1976] to phase modulate the CW that is transmitted. These codes have the property that the autocorrelation function is maximum when the code is exactly aligned with itself and has very low side lobes when off aligned as shown in the following example. Each channel uses a different code to avoid interference between channels.

The code that the transmitter and the receiver use is 255 bits long. A bit refers to a 0 or a 1. In Figure 1, a code length of 7 bits is used to illustrate the desirable properties of a maximal linear code. A code length of 255 bits will have these same properties. One of the possible starting patterns for the 7 bit maximal linear code is 1110010. Shifted versions of 1110010 appear below it in the figure. The autocorrelation function Ψ is defined as the number of agreements minus the number of disagreements as the code is compared bit by bit to a shifted version of itself. Ψ equals 7 when there is 0 offset and equals -1 for all the other bit offsets. Figure 2 shows a plot of Ψ vs offset. As the offset goes from -1 bits to 0 bits, Ψ increases linearly from a value of -1 to its maximum value of 7. Ψ decreases linearly back to -1 as the offset goes from 0 bits to 1 bits.

The operation of the receiver using a maximal linear code of length 7 to do ranging is illustrated in Figure 3. Using the actual chip length of 255 to illustrate the ranging operation would be cumbersome.

The diagram of Figure 3(a) describes the notation that will be used. The 7 bit code is drawn in blocks. The length of each block containing a bit is T/7. The length of the entire 7 chip code is then T.

Figure 3(b) illustrates how the code is shifted to produce offset versions of the original starting code. After each T period the current bit of the code remains unchanged an extra T/7. This is shown in the figure by the size of every seventh block being twice as big as the other blocks. The bit in that block remains for 2T/7. Blocks of length T are then broken up by the curly brackets on top of the chips. The first block is the 7 bit code 1110010. The second block is 0111001, which is just the code 1110010 with an offset of 1. The third block is 1011100, which is just the code 1110010 with an offset of 2. This is the idea used to shift the codes at the receiver relative to the incoming code of the transmitted signal.

Figure 3(c) illustrates the operation of the receiver in ranging mode. The code that phase modulates the transmitter signal and an identical code that phase modulates the first local oscillator (LO) of the receiver start at exactly the same time. This starting point is made possible by synchronous 1 second ticks that come from Universal Time (UT) clocks on board the geostationary satellite GOES 2 (1977 048A) parked at 113.74 degrees W longitude. These ticks are received at both the transmitter site and the receiver site. The transmitted signal arrives at the receiver with a delay due to the propagation delay of the channel. This channel delay is the quantity to be measured.

The code that phase modulates the receiver's first LO is shifted relative to the code that is modulating the transmitter signal by the delaying a bit every T seconds as illustrated previously in Figure 3(b). The receiver computes the magnitude of the autocorrelation function $|\Psi|$ over a period of T seconds. This result is the receiver output during the next T second interval. For example, during the interval 2T to 3T the receiver would compute the magnitude of the autocorrelation function between 1011100 from the transmitter signal and 1011100 from the receiver's first LO. The resulting $|\Psi|$ equals 7. This is the receiver output from 3T to 4T. Also during this 3T to 4T interval the receiver will compute the magnitude of the autocorrelation function between 1011100 from the transmitter signal and 0101110 from the receiver's first LO. The resulting $|\Psi|$ equals 1. This is the receiver output from 4T to 5T.

The channel delay is computed by multiplying the interval number that $|\Psi|$ is maximum times the bit length of the code. In Figure 3(b) the maximum value of the receiver output occurs during the third T interval so the maximum value of $|\Psi|$ occurs during the second interval. The delay is 2 times T/7.

The code modulating the receiver's first LO is shifted relative to the code modulating the transmitter signal by 1 bit every T seconds. Since these are 7 bit codes it take 7 such shifts for the code modulating the receiver's first LO to cycle through the code modulating the transmitter signal. Therefore, autocorrelation peaks should occur every 7T seconds. $|\Psi|$ equal 7 again during the 9T to 10T interval which occurs 7T seconds after the 3T to 4T interval.

The code that our system uses to do the ranging experiment is 255 bits long rather than the 7 bit code which has been used to illustrate the receiver operation. The period T of the 255 bit code is 12.75 ms so the bit length is 50 μsec (T/255). In the above illustration the receiver code was offset relative to the transmitter code by an entire bit. However, in the actual system, the receiver code is offset relative to the transmitter code by only a fifth of a bit length, or 10 μsec. 10 μsec will then be the ranging resolution of our system. The two codes will be aligned after 1275 (5 x 255) offsets. The offset occurs in the receiver code every 8 msec, so the two codes will be aligned every 10.2 sec (1275 offsets × 8 msec/offset). Section 3 contains data that was gathered while operating the system in the ranging mode.

3. Preliminary Results

Figure 4 shows ranging data. The ranging data for channel 1 is the top plot and the ranging data for channel 6 is the bottom plot. The horizontal axis of each plot is time in units of seconds. The vertical axis of each plot is the amplitude of the

output signal from the receiver and is in units of volts. The spikes appear in the data when the code that modulates the received signal and the code modulating the receiver's first local oscillator code have 0 offset. The time that a spike occurs is written next to several spikes. For example, in the top plot, a spike occurred at 72.888 seconds. This value is referenced to when the codes started at both the receiver and the transmitter. The period between spikes is shown on the left side of the data along a column. The period between spikes on the top plot is 10.2 seconds which agrees with the expected value.

The relative delays between spikes that occur on different channels using channel 6 as a reference are displayed in a table at the top left corner of the figure. The relative delay between the spikes that occur on channel 6 and channel 5 is .50 ms. However, each of the six channels had a different starting offset of 10 bits from one another. This means that channel 6 started at bit #50 while channel 5 started at bit #40 while channel 4 started at bit #30...etc. This means that the autocorrelation function for channel 6 will be maximum before the autocorrelation functions of the other 5 channels. Channel 6 had a head start so to speak of 10 bits or .50 msec (10 bits × 50 μs per bit). Corrected values of the relative delay are computed which take this offset into account. The true offsets between channels appear in the corrected relative delay column of the table. All the output peaks have 0 relative delay except for channel 3 which occurs .01 ms after the other channels. Since the relative time delay at different frequencies is so small (with an upper bound equal to 0.01 ms) for a frequency separation of 0.0625 MHz (between channel 1 and 4), the dispersive coefficient can be computed to be 0.255×10^{-10} sec^2 or smaller. Using the formulas given by Lin *et al.* [1989], the implied dispersive bandwidth turns out to be 30 kHz or larger for this occasion. This is quite a reasonable value.

The UT clock in Colorado has not been working properly since December of 1989. Therefore we cannot compute the ranging delay from this data since there is not a synchronous tick to start the codes.

The codes can still be used to see whether multiple propagation paths exist. Figure 5 contains ranging data showing that multiple propagation paths exist in channels 1, 2 and 4. Spike #2 of channel 2 took a different path from the transmitter to the receiver than spike #1. The relative time delay in this different path is computed to be .22 ms. After another 10.2 seconds a group of 3 spikes occur which indicate 3 different paths from the transmitter to the receiver.

When the sounder operates in the probing mode, the in-phase and quadrature phase components are measured, and from which amplitude and phase can be computed and plotted as functions of time. One sample of this probing data at 13:43:30 Oct. 2, 1989 is shown in Figure 6. The phase varies nearly linearly in most times, corresponding to a Doppler shift of about 0.44 Hz. It is suspected that this part of the shift comes from the frequency discrepancy between the two oscillators at transmitter and receiver. It is interesting to note that for channel 4 at a time near 4.7 second there is a 180 degree jump in phase. This corresponds to a point of zero value on the amplitude plot. Its time evolution can be clearly demonstrated using a locus plot on a complex plane. The first 10 seconds of this locus plot is shown in Figure 7. In the plot, the signal circulates counterclockwise around the origin along a spiral due to the decreasing phase and amplitude variation with time and at 4.7 second the locus passes through the origin, resulting in a 180 degree phase jump. Apparently at this time the multipath provides a complete destructive interference. Such a phenomenon has been predicted theoretically by an early numerical simulation [Rand and Yeh, 1989]. When this happens, the effect on pulse distortion can be expected to be very severe.

Shown in Figure 8 is another example taken at 16:50:00 Dec. 25, 1989. The signal strength experiences some oscillations. This is probably caused by interference of two signals arriving along two discrete paths. When the signals in two paths are vectorially summed together, a maximum or a minimum will occur depending on whether the two paths are in phase or out of phase. Suppose the two paths also suffer slightly different Doppler shifts. The maximum and minimum will repeat in a time plot as a "beat". The frequency of the beat is about 1.1 Hz for channel 1. The phase should also show a periodic pattern after taking away the linear trend and it is evident in the corresponding phase plot.

For communication applications, the quantity of most interest for this frequency selective fading channel is the complex transfer function $H(\omega,t)$. Based on this, one can go further to find the correlation function of the channel:

$$B_H(\omega,\Omega,t) = \langle H^*(\omega,t_1)H(\omega+\Omega,t_1+t)\rangle \qquad (1)$$

In our experiment $H(\omega,t)$ is measured only at a few discrete frequencies, from which some properties of B_H can be deduced. For example, the autocorrelation for each channel is obtained from (1) for $\Omega = 0$. This autocorrelation can be further normalized. One such example obtained at 12:56:00 Feb. 2, 1990 is shown in Figure 9. From this figure, the coherence time is estimated to be 1.86 seconds, corresponding to a Doppler spread of 0.538 Hz. The cross correlation between channels can be calculated in a similar manner. For a stationary uncorrelated channel, the frequency dependence of the correlation function $B_H(\omega,\Omega,t)$ will be on Ω only. To test this hypothesis, B_H curves have been calculated from experimental data for the same Ω but for different values of ω. Figure 10 shows a few plots, all for the same frequency separation Ω = 37.5 kHz but with different center frequency. Figure 10(a) is quite similar to (b) but not to (c), indicating that $B_H(\omega,\Omega,t)$ depends on Ω only if the center frequency ω is within some range. For the data shown in Figure 10, this range is between 37.5 kHz and 75 kHz. Furthermore, the correlation function shown in Figure 10 does not possess even symmetry with respect to Ω. The frequency separation at which the normalized cross correlation is 0.5 can be defined as the coherence bandwidth. A comparison of Figures 9 and 10a shows that the coherence bandwidth is slightly less than 37.5 kHz. For a larger Ω, the peak cross-correlation drops down from 0.4 to slightly less than 0.3 as shown in Figure 11. In this case, B_H depends not only on Ω but also on ω for a frequency range as small as 37.5 kHz.

4. Conclusion

We have successfully constructed and put into operation an HF sounder capable of CW operation at 800 equally spaced

discrete frequencies from 5 to 15 MHz. Two modes of operations are possible, namely, probing and ranging. The principle of this HF sounder was explained in Section 2.

A few measurements were presented in Section 3. The ranging data possesses periodic peaks which can be used to determine the propagation delay. Indeed, the relative delays have been determined up to 10 μs resolution. Based on this information, the dispersive bandwidth is estimated. In some situation, the existence of multimode can be clearly verified from the ranging data obtained.

Simultaneous measurements of the complex transfer function of up to six different frequencies were shown. They were stored on magnetic tapes for further analysis and a few cases were presented here. The channel characteristics were investigated. For a frequency separation of less than 40 kHz, the uncorrelated channel approximation is applicable, if the center frequency varies in a range between 37.5 and 75 kHz. Coherence time was found, for one example, to be about 2 seconds. The ionospheric conditions exhibit a lot of variations, which depend on many factors. Presented here are only some preliminary results. Further investigation is needed to unveil more details.

Acknowledgements: This research was supported by DAAB-07-87K-A006 from the U.S. Army Center for Command, Control and Communications Systems. R.G. Strauch of NOAA was very helpful in allowing us to use the Platteville site for radio transmissions.

References

Budden, K. G., *Radio waves in the Ionosphere*, Chapters 10 and 11, Cambridge University Press, New York, 1985.
Dixon, R. C., *Spread Spectrum Systems*, Chapter 3, John Wiley & Sons, New York, 1976.
Davies, K., *Ionospheric Radio Waves*, Chapters 4 and 5, Blaisdell, Waltham, Mass., 1969.
Ginzburg, V. L., *The Propagation of Electromagnetic Waves in a Plasma*, Chapter IV, Gordon and Breach, New York, 1964.
Lin, K.H., K.C. Yeh, H. Soicher, Bodo W. Reinisch, and R.R. Gamache, Vertical ionograms and dispersive bandwidth for an oblique path, *Radio Sci.*, 24, 519-526, 1989.
Rydbeck, O.E.H., A theoretical survey of the possibility of determining the distribution of the free electrons in the upper atmosphere, *Chalmers Tek. Hoegsk. Handl.*, 3, 1-74, 1942.
Wagner, L. S., and J. A. Goldstein, High-resolution probing of the HF ionospheric skywave channel: F-2 layer results, *Radio Sci.*, 20, 287-302, 1985.
Wagen, J.-F. and K. C. Yeh, Simulation of HF propagation and angle-of-arrival in a turbulent ionosphere, *Radio Science, 24*, 196-208, 1989a.
Wagen, J.-F. and K. C. Yeh, Numerical investigations of ionospheric effects on HF pulse, Submitted to *IEEE Transactions Antenna Propagation* for publishing, 1989b.
Wagen, J.-F. and K. C. Yeh, Numerical investigations of two-frequency mutual coherence functions of an ionospheric reflection channel, *Radio Science, 24*, 209-223, 1989c.
Yeh, K. C., and C. H. Liu, *Theory of Ionospheric Waves*, Chapter. 4, Pergamon, New York, 1972.

Ψ = number of agreements - number of disagreements

		1	1	1	0	0	1	0	number of agreements	number of disagreements	Ψ
offset	0	1	1	1	0	0	1	0	7	0	7
	1	0	1	1	1	0	0	1	3	4	-1
	2	1	0	1	1	1	0	0	3	4	-1
	3	0	1	0	1	1	1	0	3	4	-1
	4	0	0	1	0	1	1	1	3	4	-1
	5	1	0	0	1	0	1	1	3	4	-1
	6	1	1	0	0	1	0	1	3	4	-1
	7	1	1	1	0	0	1	0	3	4	-1

Figure 1. Autocorrelation function of a 7 bit maximal linear code.

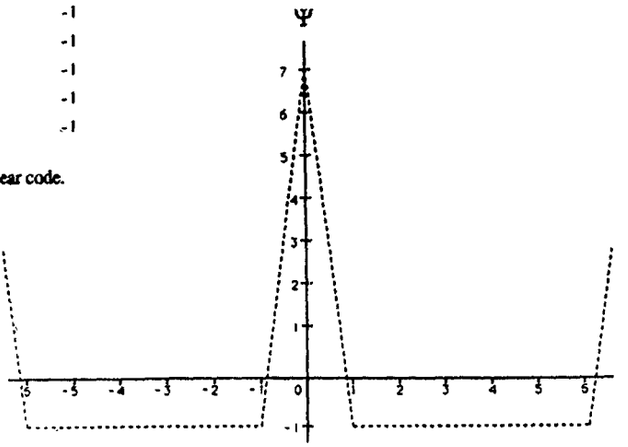
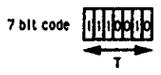
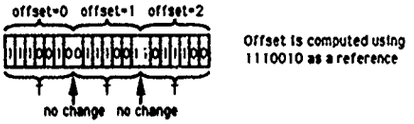


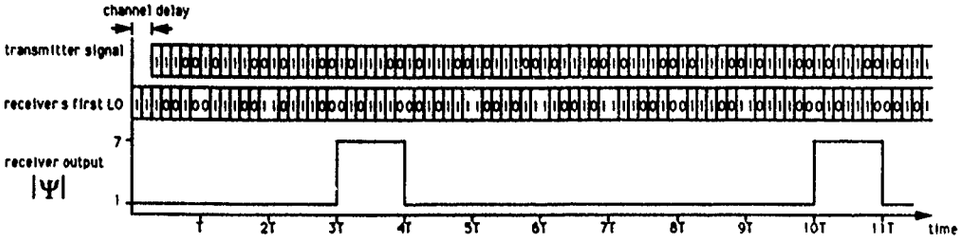
Figure 2. Plot of autocorrelation vs offset for a 7 bit maximal linear code.



(a) Representation of the 7 Bit Code



(b) Shifting Mechanism that the Receiver Code Uses



(c) Code Operation

Figure 3. Receiver operation in the ranging mode.

channel	relative delay (msec)	relative delay (corrected) (msec)
1	2 50	0 00
2	2 00	0.00
3	1.51	0 01
4	1 00	0 00
5	0 50	0.00
6	0 00	0 00

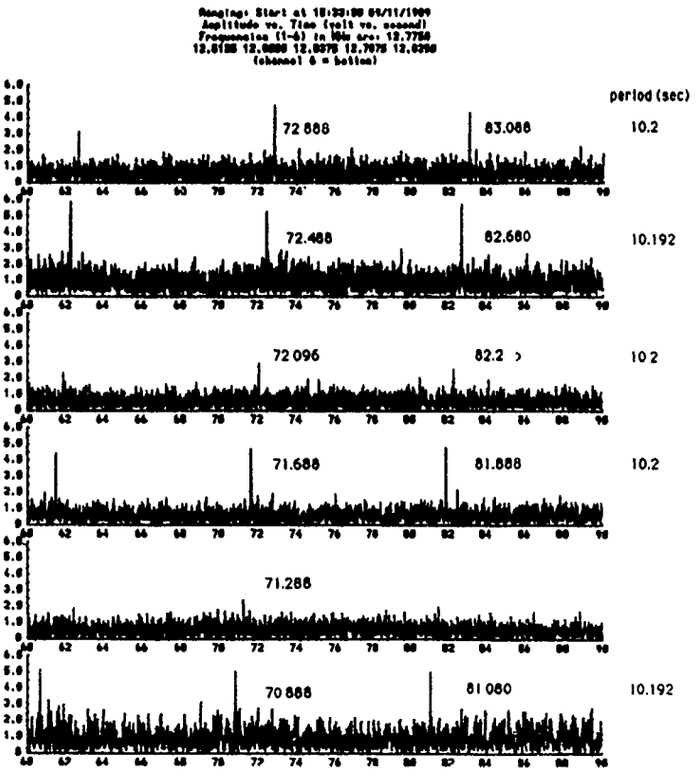


Figure 4. Ranging data showing relative delays.

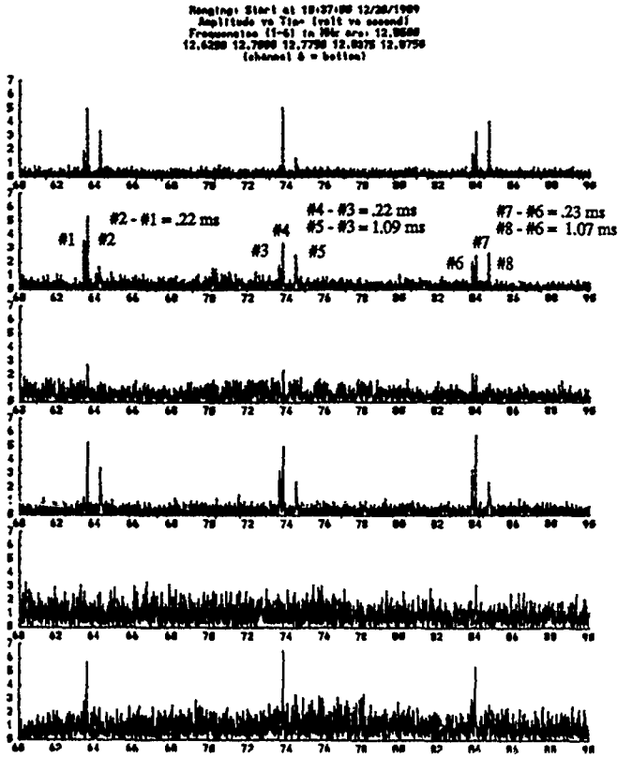


Figure 5. Ranging data showing multiple propagation paths.

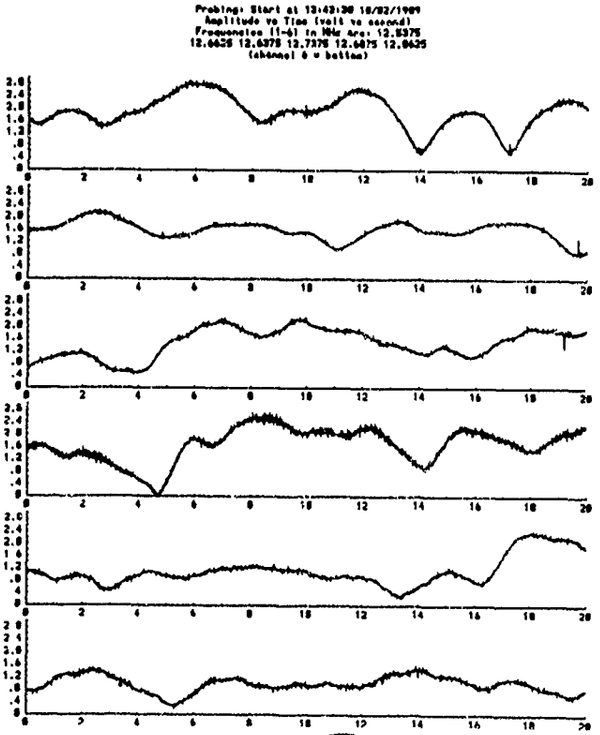


Figure 6a. Six-channel amplitude data taken at 13:43:30, Oct. 2, 1989.

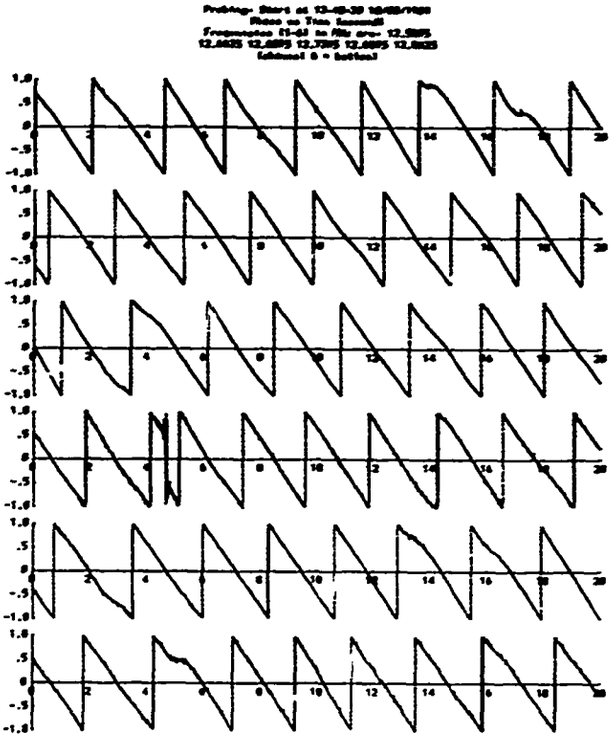


Figure 6b. Six-channel phase data taken at 13:43:30, Oct. 2, 1989.

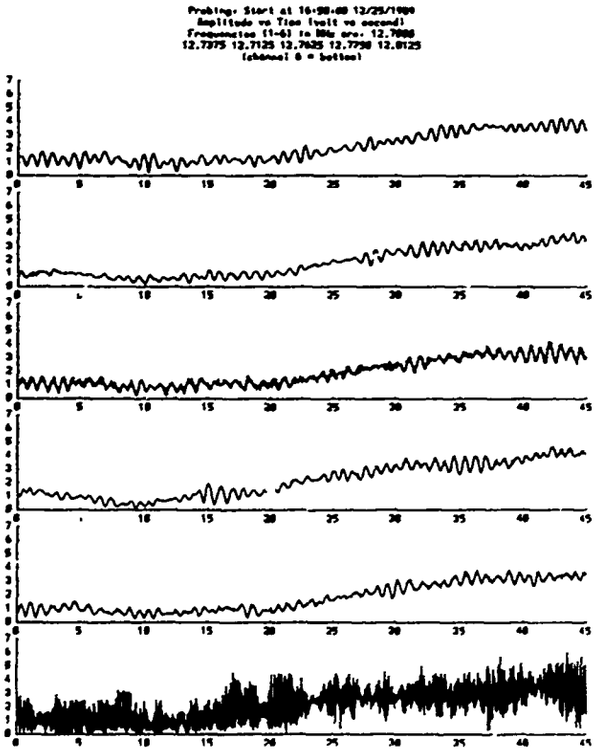


Figure 8a. Six-channel amplitude data taken at 16:50:00, Dec 25, 1989 showing oscillations caused by arrival of signals along two discrete paths

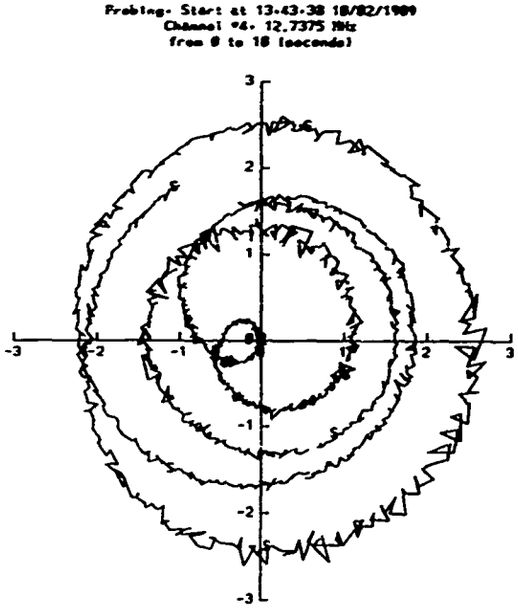


Figure 7. Locus plot of complex amplitude showing the temporal behavior for 10 seconds starting at 13:43:30, Oct. 2, 1989 at a frequency of 12.7375 MHz.

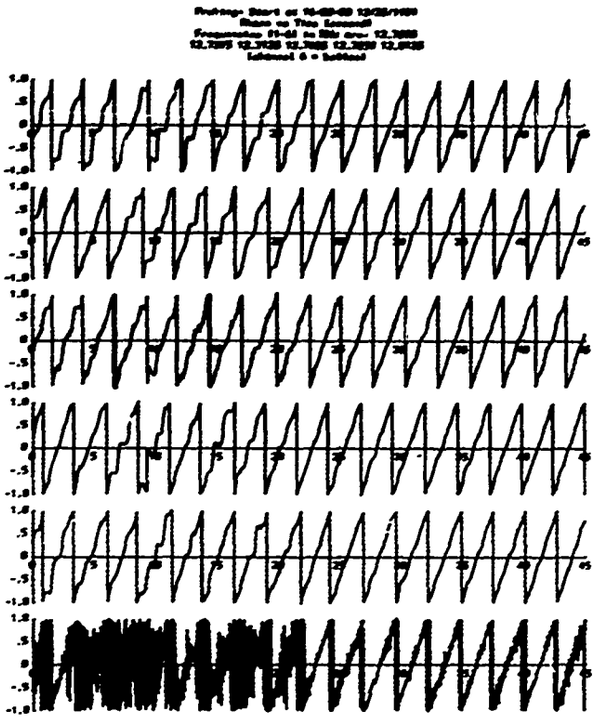


Figure 8b. Six-channel phase data taken at 16:50:00, Dec. 25, 1989 showing oscillations caused by arrival of signals along two discrete paths.

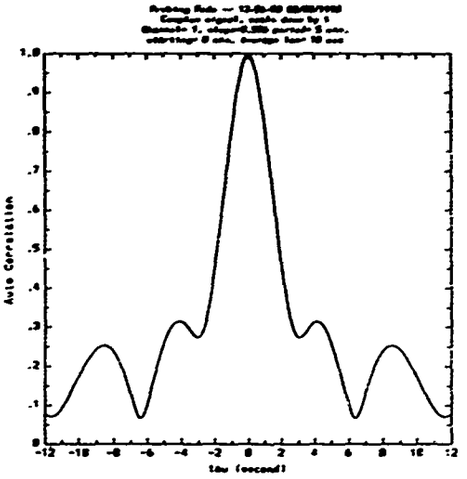


Figure 9. Normalized autocorrelation function at a frequency of 12.2125 MHz obtained at 12:56:00, Feb. 2, 1990.

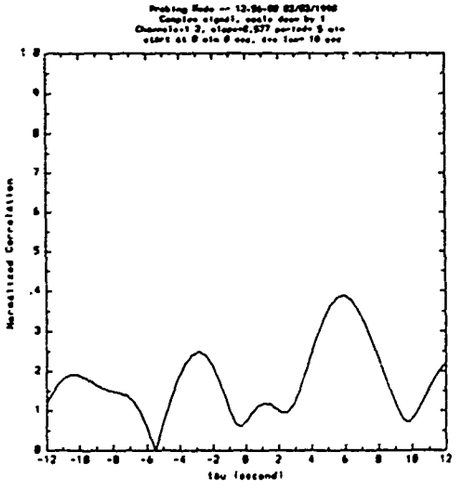


Figure 10a. Cross correlation of the transfer function for a center frequency of 12.2125 MHz and a frequency separation of 37.5 kHz.

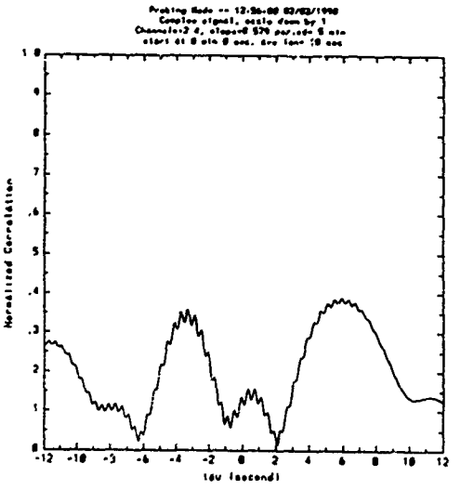


Figure 10b. Cross correlation of the transfer function for a center frequency of 12.25 MHz and a frequency separation of 37.5 kHz.

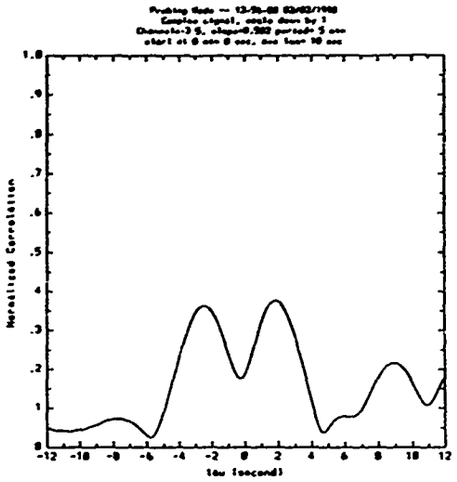


Figure 10c. Cross correlation of the transfer function for a center frequency of 12.325 MHz and a frequency separation of 37.5 kHz.

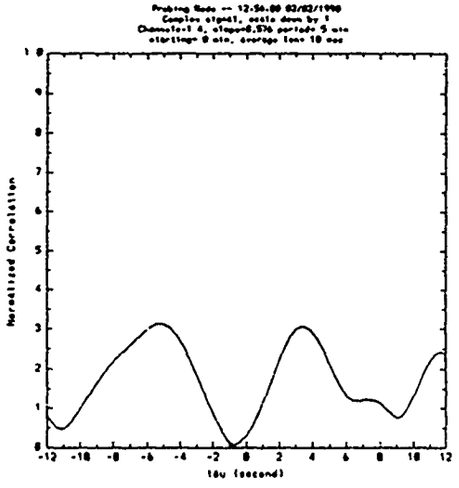


Figure 11a. Cross correlation of the transfer function for a center frequency of 12.2125 MHz and a frequency separation of 75 kHz.

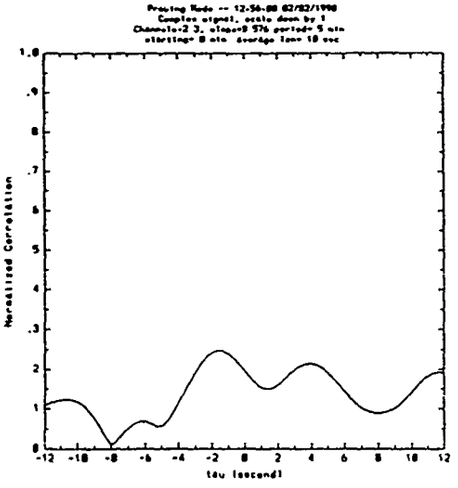


Figure 11b. Cross correlation of the transfer function for a center frequency of 12.25 MHz and a frequency separation of 75 kHz.

AD-P006 312

DETERMINATION OF DAYTIME MIDLATITUDE ELECTRON DENSITY PROFILES FROM LIMITED REAL TIME MEASUREMENTS

Dwight T. Decker
Institute for Space Research, Boston College
Chestnut Hill, Massachusetts 02167

John M. Retterer, John R. Jasperse, David N. Anderson
Ionospheric Physics Division, Geophysics Laboratory
Hanscom Air Force Base, Massachusetts 01731

Frederick J. Rich
Space Physics Division, Geophysics Laboratory
Hanscom Air Force Base, Massachusetts 01731

John C. Foster
Haystack Observatory, Massachusetts Institute of Technology
Westford, Massachusetts 01886

ABSTRACT

This paper addresses the problem of determining accurate, real time ionospheric electron density profiles (EDP) using a limited set of satellite or ground-based measurements. This is done by using real-time data to constrain the geophysical parameters that appear in an ab initio theoretical daytime midlatitude ionospheric model, creating what we call a data constrained-EDP model. The algorithm developed to perform this fitting is quite general and can be used with a variety of real-time data. With the appropriate data, the method can determine not only EDP but various other geophysical parameters (e.g. neutral composition, neutral winds). In earlier work, we have tested this approach using ionosonde data, radar data and UV airglow data. In this presentation we consider the use of satellite in-situ plasma measurements. In February to October 1989, a series of coincident measurements were made by the Millstone Hill incoherent scatter radar and the thermal plasma monitors (SSIES) on the F8 and F9 DMSP satellites. These observations of electron density profiles (radar measurements) and in-situ total ion density (SSIES measurements) give us an excellent opportunity to further test the constrained model and the potential of using SSIES measurements to help in specifying the EDP. We discuss the analysis of this data set as well as the use of other types of real-time data.

INTRODUCTION

Over the years, a number of empirical and semi-empirical models of the ionosphere's global and temporal morphology have been developed. While these statistical models have been useful in studying ionospheric climatology, it has become clear that they are not adequate for describing ionospheric "weather", the day-to-day variation of the conditions in the ionosphere. As a result, there has been increased interest in the idea of using real-time data to drive an ionospheric model in order to specify the actual conditions of the ionosphere at a given time In particular, work in progress at GL is aimed at assessing the utility of satellite in-situ plasma measurements in deducing the ionospheric electron density profile (EDP). In this paper, we consider only the daytime midlatitude ionosphere. Initially, we were interested in using optical emissions to deduce the EDP and chose to focus on UV emissions observed from a satellite viewing in the nadir direction. A difficulty that we encountered was that there is no known emission feature that in the daytime midlatitudes gives a direct signature of the EDP. As a result, we considered an indirect approach for obtaining the EDP from UV measurements. In this approach, measurements of the N_2 Lyman-Birge-Hopfield (LBH) band system and the atomic oxygen line at 1356 A are used to infer, via an airglow model, information about the solar flux and the densities of neutral atmosphere constituents. This information can then be used in a numerical calculation of the EDP using a theoretical ionospheric model based on first principles.

91-09690

One problem with this approach is that a model for the EDP depends on more geophysical input parameters than can be determined by the UV measurements. For example, the neutral winds, plasma temperatures and electric fields are important for determining the EDP but have little or no effect on the airglow. This means while the UV measurements may help to determine the solar flux and neutral densities, a unique EDP is not defined without some method to determine the other geophysical inputs. This led us to consider other measurements besides those made of optical features. In particular, we studied a four measurement concept, whereby a pair of two in-situ measurements made in the topside ionosphere (the plasma density and temperature) combined with two satellite UV emission measurements are used to provide the information necessary to specify the EDP.

To study this concept, we developed an ab initio airglow/EDP model for the daytime midlatitude ionosphere. We also developed a generalized-inverse, least-squares (GILS) algorithm that allows the use of real-time data to constrain the geophysical parameters that appear in the ab initio model. The ab initio model coupled with the GILS algorithm make up what we referred to as a data constrained-EDP model. In developing this model, we came to realize that we had an excellent tool for studying not just the use of UV and in-situ measurements together but a variety of scenarios involving different combinations of measurements with which to deduce the EDP. In this paper, we explore the case of using just satellite in-situ plasma measurements to deduce the EDP. An example of the type of in-situ instrument we are considering is the thermal plasma monitor (SSIES) presently flown on the DMSP satellites.

THE DATA CONSTRAINED EDP MODEL

Our ab initio midlatitude model includes a local steady-state photochemistry model for the molecular ions in the E and F1 regions [Jasperse, 1977]. A calculation of the photoelectron flux is included from which the daytime airglow can be determined when wanted. For the O^+ ions, the full time-dependent ion continuity equation including transport is solved [Anderson, 1973]. An estimate of the H^+ ions is made by assuming chemical equilibrium at low altitudes and diffusive equilibrium at high altitudes. The geophysical parameters that are needed by the model are as follows:

(1) the solar extreme ultraviolet (EUV) flux
(2) the neutral densities of O, N_2, O_2, NO, and H
(3) the neutral temperature (T_n)
(4) the neutral wind
(5) the temperatures T_e and T_i of the electrons and ions
(6) the electric field

The purpose of using "real time" data in conjunction with the EDP model is to constrain the geophysical parameters to "reasonable" values and thus determine a "reasonable" EDP. A fundamental problem with this concept is the fact that the number of independent quantities that can be measured remotely is smaller than the number of geophysical parameters in the theoretical model for the EDP. Using a set of values of the remote measurements as constraints on the model, we are unable to determine a unique set of values of the parameters of the theoretical model, and the procedure fails to produce a unique EDP. A generalized-inverse, least-squares (GILS) algorithm was developed [Retterer, 1987] to use statistical information on the natural variations of the parameters to choose the most likely set of parameter values that, at the same time, satisfies the constraints of the remote sensing measurements. The algorithm also gives us estimates of the likely variations of the parameters about the most probable value, allowing us to estimate the resulting uncertainty in the EDP.

Based on a linearization of the relationships between the parameters of the EDP model and the values of the quantities to be measured, the first step in the calculation is to find the hypersurface in parameter space on which the model gives the desired values of the measured quantities. For example, with a model dependent on six parameters subject to constraint by four measured quantities, this hypersurface is six - four = two dimensional; the only variations in parameters that we can allow are those that lie on this surface. The next step is determining the point on this surface that is most probable; this is done by minimizing the distance from the a priori expected values of the parameters, weighted by the inverse of the standard deviations of the parameters about the expected values. From the projection of the standard deviations of the parameters onto the constrained surface, as well the uncertainties in the measured values of the constraints, the expected uncertainties in the constrained parameters can be calculated; this gives us the expected uncertainties in the resulting EDP.

SENSITIVITY STUDY

In this study, the GILS algorithm was used to investigate the sensitivity of the EDP to variations in the parameters of the model and to investigate the effect of imposing either no observational constraints, constraint by one in-situ measurement (the plasma density), or constraint by two in-situ measurements (plasma density and electron temperature). From specified uncertainties in the geophysical parameters, we calculated the resulting uncertainty in the EDP, with and without the observational constraints. For simplicity, we neglected the electric field and effectively included the T_n variation through the neutral density variations. The NO density was assumed to be 10^{-3} times the oxygen density. The effect of variations in T_i were included when varying T_e. For this study, we fixed the spectral shape of the solar flux using the algorithms of Hinteregger et al. [1981], and we characterized

the flux by a single parameter: its energy content. The effects of varying the three neutral densities as well as the neutral winds and electron temperature were treated through the use of multiplicative scale factors applied to empirical models of these parameters. Also, the calculation of the H^+ density was not included in this study.

Case 1: No measurements

Here we examine the effect on the EDP of variations in the six parameters without constraint by measurements. We assumed a 30% standard deviation for the variation of each of the parameters and calculated the expected EDP (solid curve) and its resultant uncertainty (dashed curves) which are shown in Figure 1. In the unconstrained case, the uncertainty in the electron density at altitude z is given by the simple formula

$$\delta_n^2(z) = \sum_{i=1}^{N} \left(\frac{\partial n}{\partial \alpha_i}\right) \delta_{\alpha_i}^2$$

where $\{\alpha_i, i = 1 \text{ to } N\}$ represents the set of parameters of the model, assumed to be uncorrelated with each other in this case [Bevington, 1969]. We can see the relative uncertainty of the EDP is altitude dependent and is of the same order as the across-the-board 30 percent variation that was used for the parameters. A 20 percent marker is shown on the figure for reference purpose.

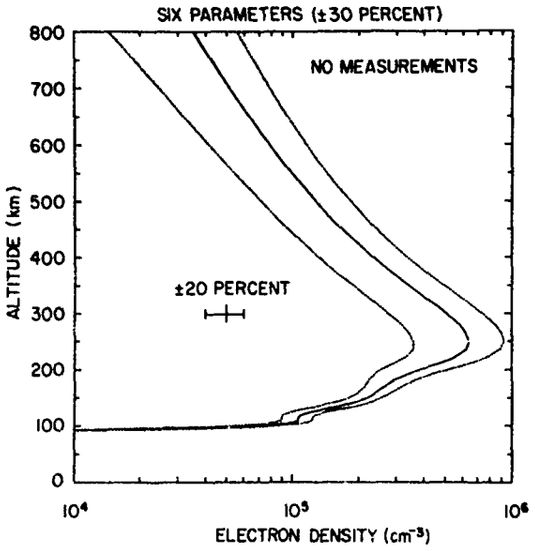


Figure 1. Unconstrained EDP

Case 2: One in-situ measurement

We considered a measurement of the electron density or the total ion density at an altitude of 800 km. Obviously, with only one measurement as a constraint we cannot uniquely determine the geophysical parameters and we expect that the allowed variations in the six parameters will still cause uncertainty in the EDP. But since the constraint is a measurement of the quantity of interest, the electron density, it should be effective in defining the topside electron density. For this case we again took 30% standard deviation in the unconstrained parameters and for the measurement we assumed a standard deviation of 10%. In Figure 2, we again present the expected EDP (solid curve) and its uncertainty (dashed curves). We see the relative uncertainty in the EDP is greatly improved at high altitudes but is only modestly reduced around the F2 peak (still around 30%). Again, we show a 20 percent marker for reference.

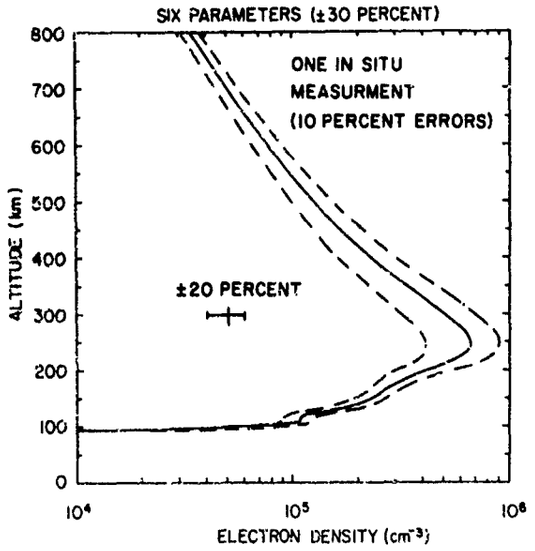


Figure 2. Constrained EDP

Case 3: Two in-situ measurements

We now consider two topside in-situ measurements. The two types of in-situ measurements are: electron temperature (Te) and ion density. We again assume a 30% standard deviation for the unconstrained parameters and a 10% standard deviation in the measurements. The results given in Figure 3 show that at all altitudes the relative uncertainty in the calculated EDP has improved as compared to the no measurement case. At 800 km, the remaining error reflects the assumed error of 10% in the density measurement. We see that having the two constraints gives a 10% uncertainty for most of the topside EDP, though at the F2 peak and below the uncertainty still runs on the order of 20%. Again, the 20 percent marker is shown.

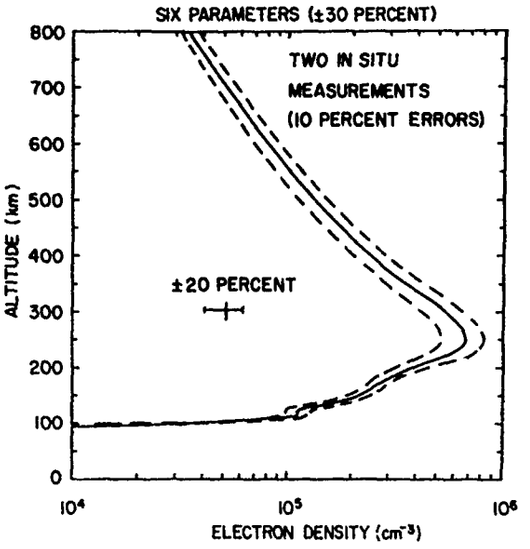


Figure 3. Constrained EDP

COMPARISONS BETWEEN THEORY AND EXPERIMENT

While a sensitivity study can demonstrate theoretically the degree of certainty possible in a determination of the EDP, questions remain concerning the accuracy of our particular EDP model. Are all the important physical processes accurately described? Are all the parameters of the microscopic physics, such as cross sections, reaction rates and diffusion coefficients, accurately known? To answer these questions requires that comparisons be made between the model and actual EDP data.

In order to make such comparisons, we arranged for a series of experiments involving observations by the Millstone Hill Incoherent Radar facility coincident with daytime overflights by a DMSP satellite. The purpose of these experiments was twofold: (1) to obtain a database of coincident in-situ satellite measurements of the total ion density in the topside ionosphere, along with profiles of the total ion density measured by the radar, and (2) to use radar measurements of the ion density at the DMSP satellite altitude (840 km) to calibrate the DMSP ion density sensor (SSIES). The database is being used to test our modeling when in-situ plasma measurements are available as constraints on the model. The comparison between satellite and radar measurements are used to give us some idea as to the accuracy of an in-situ instrument such as is flown on DMSP satellites. Knowing this accuracy is important in assessing the usefulness of the instrument because as we have seen from the sensitivity study any error in measurement will translate into at least a comparable error in the predicted EDP. Observations were made from February to October 1989 involving the F9 satellite (0930 local time) and the F8 satellite (1800 local time). The early experiments consisted of approximately four hours of radar elevation scans. This produced a time history of the electron density profile (EDP) before, during, and after a satellite overflight. In later experiments, by considering only very close overflights, it was appropriate to use the zenith antenna and analyze its data using the high resolution OASIS (Optimal Analysis of Signals from Incoherent Scatter) processing. The steerable antenna was used to obtain vector drift measurements.

To date, there have been eleven experiments, ten of which have been analyzed to some extent. In eight of the experiments, comparisons have been made between the radar measured ion density and the satellite measured density. Four of the experiments involved the F9 satellite, and on average the satellite and radar measurements agreed to within 30 percent. The best case was around 10 percent, and the worst around 60 percent. The worst case clearly suffered from poorer radar data. The other four cases involved the F8 satellite, and in those situations the average agreement between satellite and radar was 13 percent. The best case was around 5 percent and the worst case around 20 percent.

For eight experiments, we have performed our preliminary analysis using the data constrained-EDP model. The analysis thus far has consisted of running the EDP model in its unconstrained or ab initio mode and in its one constraint mode. For the one constraint mode, we used the measured ion density at the satellite altitude as the indirect constraint on the geophysical parameters of the EDP model. We then compared the EDP determined from the one constraint modeling with the ab initio result and the radar profile, as well as the EDPs predicted by the

International Reference Ionosphere (IRI) and the Ionospheric Conductivity and Electron Density (ICED) model. In general, we found that our EDP model in either of these two simple modes performed better than IRI and comparable to ICED.

In Fig. 4, we give as an example of our analysis the results for the March 24, 1989 experiment. The solid curve is the radar profile. The long dashed curve is the EDP predicted by IRI (curve 1). The short dashed curve is the ab initio mode result (curve 2) from the EDP model. The dot dashed curve is the one constraint result (curve 3) from the model. This is an excellent example of how the real time in-situ data can bring the predicted EDP into agreement with the observed profile. If the ab initio result is characteristic of the predictions of ionospheric climatology, then the differences between the ab initio profile and either the radar or constrained-model profiles suggest the significance of ionospheric weather.

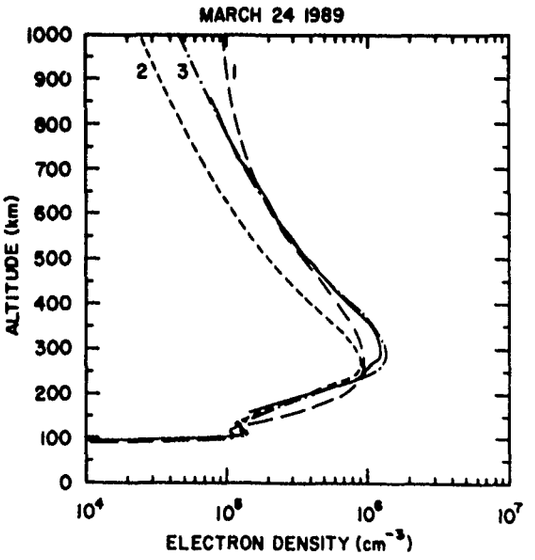


Figure 4. EDP Comparisons

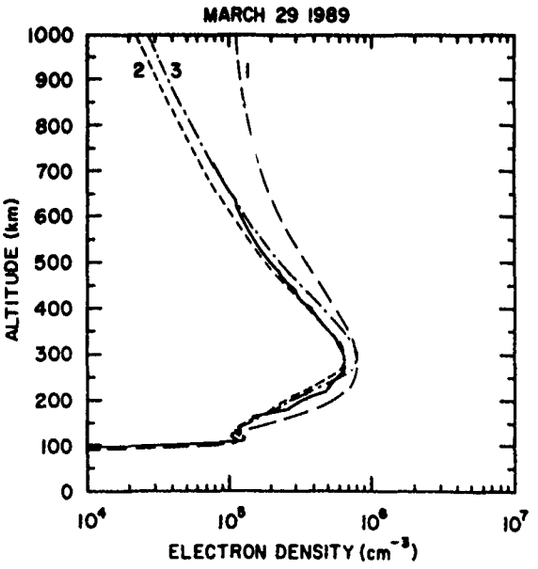


Figure 5. EDP Comparisons

In Fig. 5, we give a similar plot from our analysis of the March 29, 1989 experiment. The same sequence of results are plotted, but this time we see that the in-situ measurement actually degraded our agreement at the F2 peak. This is an example of how just one constraint cannot insure a perfect fit to the measured profile. We note that both the ab initio and one constraint calculations still perform better than IRI.

DISCUSSION

We have shown that in-situ plasma measurements have potential for helping to determine the electron density profile (EDP). From both a sensitivity study and a preliminary data analysis, we see that a topside in-situ measurement of the total ion density improves the prediction of the topside profile. However, to obtain an accurate determination of the F2 peak, it appears that more information besides the in-situ density is needed. We have also seen that the in-situ ion density measurements being made by the thermal plasma monitor (SSIES) on the F9 and F8 DMSP satellites agree with radar measurements to within 30 percent and in some cases within 10 percent.

The next step in our analysis will involve a 2 constraint fit (ion density and T_e) to the data, followed by a 3 constraint fit (ion density, T_e, and T_i). Another scenario we wish to study involves a different set of 3 constraints: an in-situ ion density, the F2 peak density, and the F2 peak altitude. This combination of data would be available when a satellite would overfly a ground-based digisonde. These three constraints practically assure a good profile. The interesting question would be after fitting the data at the location of the digisonde, how well can you model the EDP along the rest of the track of the satellite. Looking at this question will take us from considering a strictly local use of the real-time data towards a more global perspective.

ACKNOWLEDGEMENT

The research reported in this paper was supported by the DMSP Program Office, Space Systems Division, Air Force Systems Command.

REFERENCES

Anderson, D. N., A theoretical study of the ionospheric F region equatorial anomaly, 1, Theory, Planet. Space Sci., 21, 409, 1973.

Bevington, P. R., Data reduction and error analysis for the physical sciences, 366 pp., McGraw-Hill, New York, 1969.

Hinteregger, H. E., K. Fukui, and B. R. Gilson, Observational, reference and model data on solar EUV, from measurements on AE-E, Geophys. Res. Lett., 8, 1147, 1981.

Jasperse, J. R., Electron distribution function and ion concentrations in the earth's lower ionosphere from Boltzmann-Fokker-Planck theory, Planet. Space Sci., 25, 743, 1977.

Retterer, J.M., Optimal model fitting with fewer constraints than parameters, Boston College SDAL Technical Report BC-SDAL-87-3, 1987.

AD-P006 313

USING IONOSONDES FOR GLOBAL MAPPING OF THERMOSPHERIC MERIDIONAL WINDS

K. L. Miller
Center for Atmospheric and Space Sciences
Utah State University
Logan, UT 84322-4405

P. G. Richards and D. G. Torr
Center for Space Plasma and Aeronomic Research
The University of Alabama in Huntsville
Huntsville, AL 35899

ABSTRACT

This paper describes a recently developed method of deriving the meridional component of the horizontal neutral wind in the thermosphere. The method is based on the approximately linear relationship between the height of the ionospheric *F2*-layer peak and the strength of the meridional neutral wind. An ionospheric photochemical model is used to compute the parameters that describe this relationship. The neutral wind speed is derived from a measurement of the height of the *F2*-layer (h_{max}). The measurement of h_{max} can be made by incoherent scatter radar and ionogram inversion techniques, but is also derivable from parameters that are routinely scaled from ionograms and archived in world data centers. The wind speeds derived from h_{max} are compared with winds from Fabry-Perot interferometry and incoherent scatter radar, and with results of the NCAR Thermospheric General Circulation Model.

1. INTRODUCTION

A measurement of the neutral air motions is valuable in any study of ionosphere-thermosphere coupling. The neutral wind affects many of the observable quantities and physical processes of the ionosphere, including the density profile of the ionospheric *F* region, and the generation and maintenance of electric fields. Several models have been developed to describe the photochemistry and dynamics of the thermosphere and ionosphere [Torr et al, 1988; Sojka, 1989]. The models are becoming more and more complex, and are to a great extent limited by the accuracies of input parameters such as reaction rates, collision cross-sections, and solar flux. One of the major limitations, especially to the electron density profile, is the lack of knowledge of the meridional neutral wind.

This paper describes a method to derive the meridional component of the neutral wind in the thermosphere from existing ionosonde measurements. The method has been shown to derive winds with comparable accuracy to other techniques. It has the advantage of being able to derive winds at any local time, and at any of the midlatitude ionosonde sites. Winds can be derived retroactively, since data from many ionosondes are routinely archived at the world data centers.

2. DEPENDENCE OF THE *F2*-LAYER HEIGHT ON MERIDIONAL WIND SPEED

Because of the magnetic control of the processes that determine the *F2*-layer height, the most useful coordinate system is one that reflects the direction of the geomagnetic field. Hence the use of the words "north" and "northward" in this paper always refer to magnetic north, and "meridional" always refers to the

91-09689

magnetic meridian. The declination of the magnetic field is generally less than 20° throughout the mid-latitude region. The largest declinations at mid-latitudes occur in the South Indian Ocean.

The component of the neutral wind that is parallel to the geomagnetic field affects the profile shape and thus the concentration of the ionization at *F*-region heights through ion-neutral collisions [Rishbeth, 1972]. Since vertical neutral drifts are normally small, the dominant collisional forcing is from the magnetically north-south component of the horizontal wind. A poleward component of the horizontal neutral wind will have the effect of forcing ionization downward parallel to the magnetic field, while an equator-ward component will have the opposite effect.

The motion of the ions maintained by a meridional neutral wind is dependent on the dip angle, I, of the geomagnetic field. Under quasi-steady-state conditions, the contribution to the northward parallel ion velocity (V_i) from the magnetically-northward neutral wind component (U_N) is

$$V_i = U_N \cos(I) \tag{1}$$

The vertical component of the ion velocity (V_z) that is due to this effect is

$$V_z = -U_N \sin(I) \cos(I) \tag{2}$$

If there is no externally applied forcing of vertical motion, the shape of the *F2* layer is determined by production, loss, and diffusion. The motion of the ions that is induced through the motion of the neutral wind either adds to or subtracts from the diffusion velocity of the ions. The height of the maximum ion density, or balance height, is determined by the altitude where the influences of loss and diffusion are approximately equal. If a force with a component parallel to the magnetic field is applied by an external source such as neutral wind, the peak electron density is displaced from the balance height. The effectiveness of the neutral wind to affect the height of the *F2* layer decreases significantly at high and low latitudes. It is most effective near 25° from the geomagnetic dip equator where the dip angle is 45°. It is, however, a significant factor in determining the vertical ion motions in the *F* region throughout the midlatitude region.

3. DEVELOPMENT OF THE "SERVO" MODEL

Rishbeth [1966] and Hanson and Patterson [1964] showed that if the external forcing from the neutral atmosphere is not too great, there is a linear relationship between the height of the layer (h_{max}) and the neutral wind speed. Rishbeth [1967] compared the behavior of the F2 region under the influence of a meridional neutral wind to a servo system, and developed the equations describing the rate of vertical movement of the layer and the equilibrium position of the peak. According to the servo model, an equatorward wind will force the ionization to higher altitudes. The raising of the layer increases the downward diffusion, which in turn opposes further upward motion of the layer. Similarly, a poleward wind will force the ionization to lower altitudes where the increased recombination rate opposes a further lowering of the layer maximum. Except for extreme wind speeds, and if electric fields are small, a linear relationship holds approximately between the neutral flow (U_N) and the resulting change in the height of the *F*-layer peak [Rishbeth and Barron, 1960; Hanson and Patterson, 1964; Rishbeth, 1966; Buonsanto, et al. 1989].

By assuming an *F2* region composed primarily of O^+ with an ionization peak well above the production region and a loss process dominated by the reaction of O^+ with N_2, Rishbeth et al. [1978] derived a relationship between an externally induced vertical drift and the rate of change of the layer height. Buonsanto et al. [1989] showed that, for most conditions in the *F* region, the effects of both the time rate of change of the layer height and the non-linearity are small. Under these conditions, the vertical drift V_z becomes

$$V_z = \frac{(k+1)\, D_m \sin^2(I)}{2H^2} (h_m - h_o) \tag{3}$$

where

D_m = diffusion coefficient for O^+

I = magnetic dip angle

h_m = height of the *F2*-layer peak (h_{max})

h_o = balance height

H = scale height of the neutral ionizable gas

kH = scale height of O^+, which controls diffusion

If the vertical drift of the neutral atmosphere is assumed to be small, the vertical drift of the ionization can be assumed to be the result of a combination of ion collisions with the meridional component of the horizontal neutral wind and ion drift induced by the east-west electric field. Combining the E x B drift with Equation 2, the vertical drift can be written in terms of the northward neutral wind (U_N) and the eastward electric field (E_E) as

$$V_Z = -U_N \sin(I) \cos(I) + \frac{E_E \cos(I)}{B} \tag{4}$$

For conditions existing most of the time at mid-latitudes, the "servo" equation yields meridional neutral wind speeds from measurements of the *F2*-layer height and a neutral atmosphere model using a simple relationship from Equations 3 and 4:

$$U_N = \frac{(h_o - h_m)}{\alpha} + \frac{E_E}{B \sin(I)} \tag{5}$$

where

$$\alpha = \frac{2H^2 \cos(I)}{(k+1)D_m \sin(I)}$$

If the east-west electric field is small, or if the effects of the meridional neutral wind and electric field are to be considered together, Equation 5 becomes a simple proportionality between wind and layer height:

$$U_N = \frac{(h_o - h_m)}{\alpha} \tag{6}$$

It would be convenient to be able to use a servo-type model to derive the meridional wind from h_{max}. The approach is complicated, however, by the requirement to accurately locate h_o and calculate D_m. Ionospheric photochemical models contain the complex photochemistry that defines how h_{max} depends on the meridional wind speed. In most applications, the models use this information in computing species concentrations. They are dependent on winds and electric fields to define the profile shape and density of the *F* region. Miller et al. [1986] took advantage of the capability of the FLIP model [Richards and Torr, 1985, 1988; Young, et al., 1980] to compute h_{max} for a specified wind speed and developed a method to derive an approximate constant of proportionality for the relationship between *F2*-layer height and meridional wind speed.

In a recent study, Buonsanto et al. [1989] derived winds from an extensive Millstone Hill Radar data set using both the servo model and the FLIP model. Although the derived winds are similar, servo-model winds are consistently more negative (stronger southward) than the winds derived using the FLIP model. The ratio of the two models is about 1.2 at night, and drops to near 0.5 in the daytime.

4. DEVELOPMENT OF THE "WIND FROM H_{MAX}" MODEL

4.1. Derivation of α and h_o

A plot of h_{max} vs. meridional neutral wind speed has the characteristic "S"-shape shown in Figure 1. Figure 1 also shows the day to night difference in the response of h_{max}. The greater electron density in the daytime, combined with larger production and recombination rates result in a much stronger response of the layer height to small wind speeds, and then a much more rapid end to the linear portion of the curve. On

the other hand, the nighttime change in layer height is smaller, with the linear region extending to greater equatorward wind speeds. The day-to-night difference in the response of the layer height to the neutral wind is important to the accuracy of wind speeds derived using this technique. The wind during the daytime is normally poleward and relatively small, while the nighttime wind is equatorward and often near 200 m s^{-1}.

The values of the parameters h_o and α to be used in Equation 6 can be found by deriving layer heights at several wind speeds, including h_o at $U_N = 0$, and then determining the slope (α) of the linear region. In practice, little accuracy is lost by finding h_{max} at two wind speeds that span the expected wind speeds but are still within the linear region, and then determining α from the slope of the line connecting the two points and h_o from the height where the line crosses $U_N = 0$.

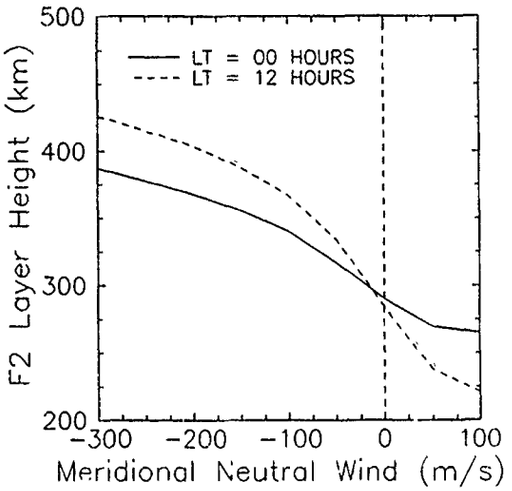


Figure 1. Height of the peak of the *F2* layer as a function of meridional wind speed at noon and midnight for equinox conditions.

4.2. Dependence of h_o and α on Geographic and Solar Parameters

The response of the height of the *F2* layer to changes in the meridional neutral wind speed depends strongly on the inclination of the geomagnetic field and of the local atmospheric conditions. As discussed above, the main determining factors of the layer height are the component of the neutral wind that is parallel to the magnetic field, diffusion of the ions along the field, and recombination.

Magnetic latitude is the primary geographic parameter in determining α and h_o. Even though the ionospheric density is affected by the neutral atmosphere, the effect of geographic latitude on the derivation of winds is much less important than the effect of magnetic latitude. Figure 2 shows the local time variation of α and h_o at four different magnetic latitudes, as computed by the FLIP model. The calculations were done for magnetically quiet conditions ($A_p = 5$) and moderate solar flux (F10.7 = 130) at a magnetic longitude of 0°.

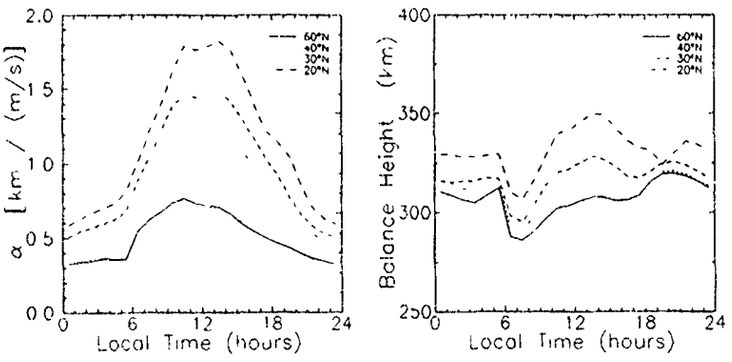


Figure 2. Local time variation of α and h_o at four geomagnetic latitudes.

Variation of α and h_o with respect to magnetic activity is primarily in the balance height, although there is a small change in α. The insensitivity of these parameters to magnetic activity has important consequences in the study of storm effects, since they occur quickly and it is difficult to model the ionosphere with adequate time resolution. A larger effect is found in the variation of α and h_o with changes in solar flux. There are major differences in both α and h_o over the range of F10.7 values during a sunspot cycle, mainly the result of the increased production rate of ionospheric O^+ and the larger scale height of the warmer neutral atmosphere at solar maximum. Although α and h_o are both strongly dependent on F10.7, it is such a slowly-varying function that it does not introduce a large uncertainty into the technique for deriving winds. While it will change significantly from month to month, there is very little day-to-day variation, even during magnetic storms.

4.3. Measurement of h_{max}

One of the advantages of having the capability to derive meridional wind speeds from h_{max} is the availability of data on a global scale from which h_{max} can be derived. The more traditional optical and incoherent scatter radar techniques of measuring neutral winds are valuable and provide checks on this method, but are limited in their coverage by the weather and hours of darkness in the case of the optical measurements, and by the limited number and locations of the radars.

Ionosondes provide the data base from which global arrays of h_{max} can be derived. In simple terms, an ionosonde measures the time of flight of a radio pulse from the transmitter to the reflection point in the ionosphere and back to the ionosonde receiver. As the frequency increases, a virtual-height profile of the plasma frequency of the ionosphere, or ionogram, is generated. To determine h_{max} the virtual height profile must be converted to a true height profile, or at least the true height of the reflection at the peak plasma frequency of the *F2* region *(foF2)* must be determined. In the best situation, the ionogram is either recorded digitally or hand scaled, and a true-height analysis converts the vertical heights of reflection to true heights [Titheridge, 1985]. It is tedious work to hand-scale a complete ionogram, and is not done routinely. Complete electron density profiles are being recorded, however, by a new generation of digital HF sounders.

At the present, any global study that relies on ionosonde data must derive its parameters from data that are routinely scaled from ionograms and archived in the World Data Centers. The parameters most relevant to this study are the transmission factor *M(3000)F2,* the maximum plasma frequency of the *F2* region, *foF2* and the maximum frequency of the *E* region, *foE*. Dudeney [1983] presented an empirical model, based on work by Shimazaki [1955] and by Bradley and Dudeney [1973], that determines the layer height primarily from the approximately inverse relationship between the height of the peak of the *F2* layer and *M(3000)F2*. In this model the ratio of *foF2* to *foE* is included as an indicator of the underlying ionization that retards the pulse and increases the virtual height.

4.4. Accuracy of Derived Winds.

Accuracies of h_o and α are based on the accuracy of the FLIP model. The height of the *F2*-layer peak is determined by a parabolic fit to the three points that define the peak ion density in the model. The step size can be adjusted, but is nominally about 8 km in the region of interest. A larger source of uncertainty is the degree to which the model represents the real ionosphere.

The accuracy of h_{max} is dependent on the type of measurement, the method of derivation, and also the time of day. h_{max} measured by incoherent scatter radar has the potential of being the most accurate, although the accuracy depends on various radar parameters such as the pulse length, signal-to-noise ratio, the conversion between scattered power and electron density, and the method of interpolation used to determine the maximum in the ionization density. There have been few studies published of the accuracy of the *M(3000)F2* technique. Dudeney [1983] suggests that the accuracy in the derived h_{max} is about 5%.

Figure 3 shows the derived meridional wind at Tokyo on 18 September 1984, during the Equinox Transition Study. The error bars in Figure 3 are statistical error bars from the RMS error of Equation 6. The error bars are small compared with the diurnal variation of the derived winds, and are of comparable size to the hourly variability on this particular day. The winds are derived from data taken on a magnetically quiet day, and would normally be much more variable.

The greater sources of uncertainty are not statistical, however, but are systematic and derive from the lack of knowledge of many of the physical processes involved. Although the FLIP model generates the ionization density from photochemical reactions, it uses the MSIS empirical model for neutral densities and neutral temperature. The model performs well in tests against measurements, but is unable to predict unusual or transient features. Recently Burnside et al. [1987] has shown evidence that the O-O$^+$ collision cross section should be increased by a factor of approximately 1.7 over the previously used value. An increase in collision cross section in the FLIP model has the effect of significantly reducing the magnitude of the derived winds. Unless otherwise stated, the model results illustrated in this paper were derived using the larger collision cross section.

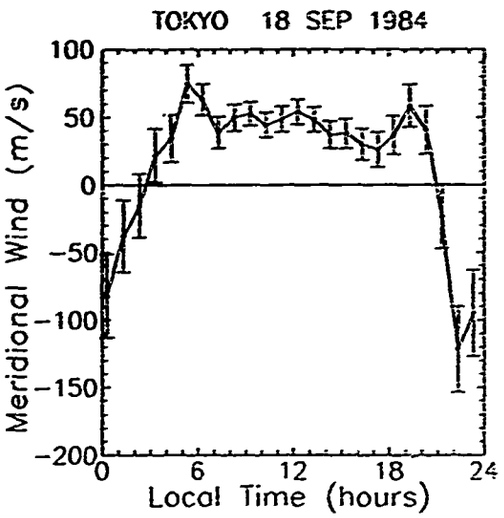


Figure 3. Meridional wind speed at Tokyo (36° N, 140° E) on 18 September 1984.

The magnitude of the east-west electric field also introduces a systematic error into the derived winds (Equation 5). There is currently no electric field model that has been shown to be accurate on a global scale, although some have successfully reproduced local measurements. This is, however, not a serious limitation if the results of the wind derivation is used as input to an ionospheric model, since the sum of the forces due to the meridional neutral wind and the electric field is the quantity needed by most models.

5. CURRENT RESULTS

5.1 Comparisons With Other Wind Measurements

Several comparisons have been made of winds derived from h_{max} with meridional winds derived from other methods. These have been primarily with incoherent scatter radar, although Miller et al. [1986] also compared results with winds measured by Fabry-Perot interferometry. In each of these comparisons the wind speeds derived from h_{max} are consistent with the results of the other technique, both in the magnitude and in the pattern of variability of the winds.

Plots of the comparisons from Miller et al. [1986] are shown in Figure 4. The winds were measured at Arecibo using Fabry-Perot interferometry (F.P.I.) and derived from incoherent scatter ion drift measurements (v_i), [Burnside et al., 1983]. The curves labeled "Model" are derived from h_{max} without the 1.7 factor in the collision cross section.

Miller et al. [1987] made a detailed comparison with the Millstone Hill incoherent scatter radar. This study had two objectives: First, to compare the difference between the results with the statistical uncertainty of the two derivations of wind speed. Second, to include measured electric fields in the calculation of winds from h_{max} to see if this would account for the difference.

The comparison was made for the Global Thermospheric Mapping Study (GTMS) campaigns of 26-18 June 1984 and 15-17 January 1985. Meridional neutral winds were derived from ion drift measurements for both campaigns. The factor of 1.7 was not used in the collision cross section in either wind derivation of this study. The east-west electric field was derived from ion velocity measurements made at Millstone Hill during the June campaign. The measured field was small throughout most of the time interval. The largest electric fields, reaching 4 mV/m in a westward direction, were measured on June 27.

Using Equation 5, we can determine the relative magnitude of the error introduced by neglecting the electric field. Figure 5 illustrates the results of including this measured field in the wind calculation. In

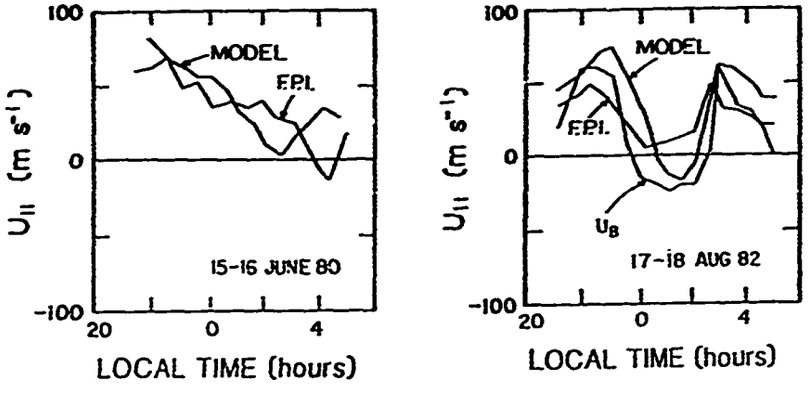


Figure 4. Comparisons of winds from h_{max} with winds measured by F.P.I. and by incoherent scatter ion drift measurements (U_B) for Arecibo.

Figure 5, the dashed curve is the wind from the incoherent scatter measurement. The solid curve was derived from h_{max}, and the dotted curve is the wind from h_{max} corrected by the electric field measurement, as described by Equation 5. On June 27, when the strongest electric fields were measured, the difference between the curves in the morning is accounted for by the 4 mV/m field. However, the measured electric field is not enough to reconcile the two wind curves in the afternoon.

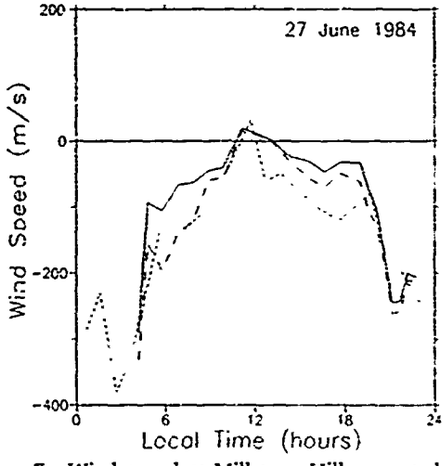


Figure 5. Wind speed at Millstone Hill corrected for the effects of electric fields.

Winds derived from ionosonde measurements cannot be independently verified. However, ionosondes near incoherent scatter radars should give similar results to the radars. An example is shown in Figure 6, where the meridional wind derived from ion drift measurements by the Millstone Hill radar are compared with the winds from h_{max} measurements by the Ottawa ionosonde for 26 June 1984. The two stations are separated by about 500 km in distance. Millstone Hill is at 42.5° N, 71.5° W, and Ottawa is at 45.1° N, 76.1° W. The three-parameter formulation developed by Dudeney [1983] was used to find the ionosonde values of h_{max}. There is agreement in both the wind speed and in the times of occurrence of the major variations at the two locations. The major feature is the abatement in the southward wind that occurs near sunset. This abatement was also seen in the wind values for the January GTMS period at Millstone Hill.

A comparison of winds derived from h_{max} with results of the NCAR TGCM [Crowley, et al., 1989] is shown in Figure 7. This example is from the Equinox Transition Study of September 1984. Magnetic conditions were quiet on the first day, but the second day includes an equatorward surge at the onset of a magnetic storm. This was a "blind" comparison; no attempt was made to adjust either model to make it agree with the other.

5.3 Development of a Global Wind Model.

As the reliability of the derivation of meridional neutral winds from ionosondes becomes established for quiet times, it will be possible to develop an empirical neutral wind model based entirely on ionosonde measurements. The size of the data base is already large enough to provide adequate diurnal and solar cycle coverage, although allowance must be made for the lack of coverage in some parts of the globe.

The possibility of a global wind model based on ionosonde measurements is illustrated by a study by Miller et al. [1990] to compare the winds derived from median ionosonde data for a 15-day period with winds derived from the International Reference Ionosphere (IRI) values of h_{max}. Figure 8 shows agreement between ionosonde values and IRI values of the meridional wind speed at a set of ionosonde stations at approximately constant longitude. The stars show wind speed derived from 15-day median ionosonde measurements of h_{max}. The solid line is the wind speed derived from IRI values of h_{max}. The dashed line that is included in Figure 8 is the wind speed from the HWM87 empirical wind model [Hedin et al, 1988]. The HWM87 model is based on a vector spherical harmonic fit to satellite wind measurements.

Comparing winds derived from ionosonde measurements with winds derived from the IRI is equivalent to comparing the IRI predictions of h_{max} with ionosonde measurement. However, it suggests the possibility of developing a global wind model based on the IRI. The reliability of such an empirical meridional wind model would be critically dependent on the reliability of the IRI values of h_{max}. The IRI uses CCIR maps of ionosonde data to generate h_{max} values. These maps are, of course, most reliable in areas where the ionosonde measurements are most extensive, such as Europe and the Japanese-Australian meridian.

5.4 Support of Measurement Campaigns

Efforts to study ionosphere-thermosphere coupling are being coordinated through CEDAR on a national level, and through WITS and an international level. Both programs are interested in global wind values during campaigns. To date, the GTMS campaigns of June, 1984, and January, 1985; the ETS campaign of September, 1984; each of the GITCAD campaigns, and all but the first SUNDIAL campaign are supported by global-scale derivation of meridional wind speeds. Since the data required for the derivation of h_{max} are collected

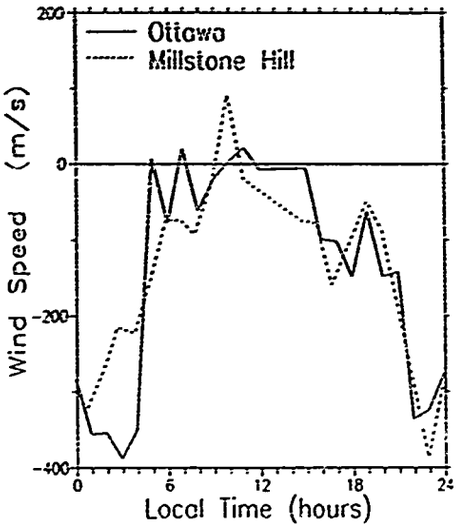


Figure 6. Meridional wind derived from ion drift at Millstone Hill compared with the winds from h_{max} at Ottawa.

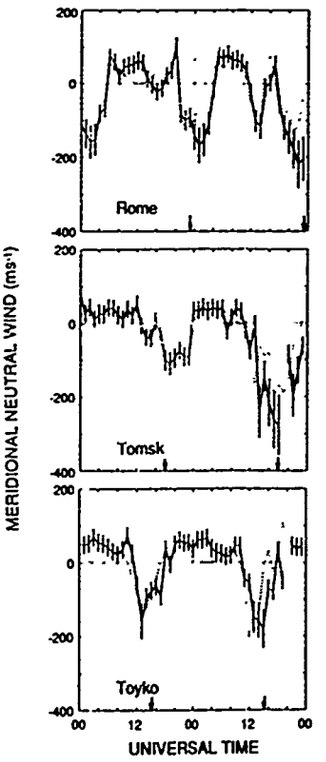


Figure 7. Meridional neutral winds measured (solid) and predicted by the NCAR TGCM (broken). Negative values indicate equatorward winds; units are m s^{-1}. Vertical arrows indicate local midnight [Crowley, et al., 1989].

and archived routinely, it is possible to derive global scale meridional winds for any campaign.

Acknowledgements: This work was supported by NSF Grant ATM-87-15367 to Utah State University and by NSF Grant ATM-87-16036 to the University of Alabama in Huntsville.

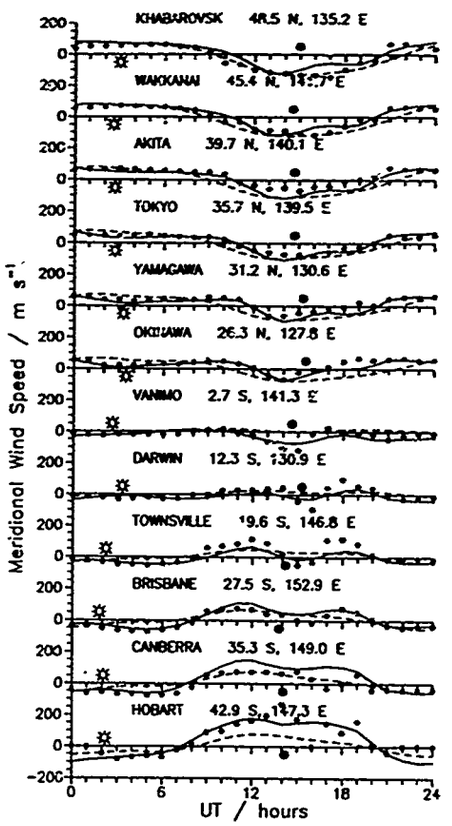


Figure 8. Meridional neutral winds (m s^{-1}) at a meridional chain of ionosonde stations. Positive winds are northward.

REFERENCES

Bradley, P. A., and J. R. Dudeney, A simple model of the vertical distribution of electron concentration in the ionosphere, J. Atmos. Terr. Phys., 35, 2131-2146, 1973.

Buonsanto, M. J., J. E. Salah, K. L. Miller, W. L. Oliver, R.G. Burnside, and P. G. Richards, Observations of neutral circulation at mid-latitudes during the equinox transition study, J. Geophys. Res., 94, 16,987-16,997, 1989.

Burnside, R. G., R. A. Behnke, and J. C. G. Walker, Meridional neutral winds in the thermosphere at Arecibo: Simultaneous incoherent scatter and airglow observations, J. Geophys. Res., 88, 3181-3189, 1983.

Burnside, R. G., C. A. Tepley, and V. B. Wickwar, The O^+-O collision cross-section: Can it be inferred from aeronomical measurements?, Annales Geophysicae, 5A, 343-350, 1987.

Crowley, G., B. A. Emery, R. G. Roble,, H. C. Carlson, Jr., J. E. Salah, V. B. Wickwar, K. L. Miller, W. L. Oliver, R. H. Burnside, and F. A. Marcos, Thermospheric dynamics during September 18-19, 1984. 1. Model simulations, J. Geophys. Res., 94, 16,925-16,944, December 1989.

Dudeney, J. R., The accuracy of simple methods for determining the height of the maximum electron concentration of the *F2*-layer from scaled ionospheric characteristics, J. Atmos. Terr. Phys., 45, 629-640, 1983.

Hanson, W. B., and T, N. Patterson, The maintenance of the night-time *F*-layer, Planet. Space Sci., 12, 979-997, 1964.

Hedin, A. E., N. W. Spencer, and T. L. Kileen, Empirical global model of upper thermosphere winds based on Atmosphere and Dynamics Explorer satellite data, J. Geophys. Res., 93, 9959-9978, 1988.

Miller, K. L., D. G. Torr, and P. G. Richards, Meridional winds in the thermosphere derived from measurement of F2 layer height, J. Geophys. Res., 91, 4531-4535, 1986.

Miller, K. L., J. E. Salah, and D. G. Torr, The effect of electric fields on measurements of meridional neutral winds in the thermosphere, Annales Geophysicae, 5A, 337-342, 1987.

Miller, K. L., A. E. Hedin, P. J. Wilkinson, D. G. Torr, and P. G. Richards, Neutral winds derived from IRI parameters and from the HWM87 wind model for the SUNDIAL campaign of September, 1986, Adv. Space Res. 10, (8)99-(8)102, 1990.

Richards P. G. and D. G. Torr, Seasonal, diurnal, and solar cycle variations of the limiting H^+ flux in the earth's topside ionosphere, J. Geophys. Res., 90, 5261-5268, 1985.

Richards, P. G. and D. G. Torr, Ratio of photoelectron to EUV ionization rates for aeronomic studies, J. Geophys. Res., 93 4060, 1988.

Rishbeth, H., and D. W. Barron, Equilibrium electron distributions in the ionospheric F2-layer, J. Atmos. Terr. Phys., 18, 234-252, 1960.

Rishbeth, H., *F2*-layer rates at sunspot minimum, J. Atmos. Terr. Phys., 28, 911-918, 1966.

Rishbeth, H., The effects of winds on the ionospheric *F2*-peak, J. Atmos. Terr. Phys., 29, 225-238, 1967.

Rishbeth, H., Thermospheric winds and the *F*-region: A review, J. Atmos. Terr. Phys., 34, 1-47, 1972.

Rishbeth, H., S. Ganguly, and J. C. G. Walker, Field-aligned and field-perpendicular velocities in the ionospheric *F2*-layer, J. Atmos. Terr. Phys., 40, 767-784, 1978.

Shimazaki, (1955), J. Radio Research Labs, Japan, 2, 85, 1955.

Sojka, J. J., Global scale, physical models of the *F* region ionosphere, Rev. Geophys, 27, 371-403, 1989.

Titheridge, J. E., Ionogram analysis with the generalized program POLAN, Report UAG-93, World Data Center A for Solar-Terrestrial Physics, NOAA, E/GC2, Boulder, CO, 80303, 1985.

Torr, D. G., P. G. Richards, and M. R. Torr, Photochemistry of the ionosphere, WITS Handbook, 1, 1-38, 1988.

Young E. R., D. G. Torr, P. Richards and A. F. Nagy, A computer simulation of the midlatitude plasmasphere and ionosphere, Planet. Space Sci., 28, 881, 1980.

AD-P006 314

A SURVEY OF POLAR CAP F-REGION ELECTRON DENSITIES MEASURED BY THE SONDRESTROM RADAR

R. M. Robinson
Space Sciences Laboratory
Lockheed Palo Alto Research Laboratories
Palo Alto, California 94304

91-09688

Abstract

The Sondrestrom incoherent scatter radar has been used to study F-region electron densities in the polar cap. The data were obtained while the radar antenna was scanned in the plane of the magnetic meridian. In this mode, electron densities are measured as a function of altitude and latitude so that ionospheric boundaries can be easily identified. One boundary that is present in much of the data is a vertical wall of electron density between 80 and 200 km altitude that separates the region of aurorally-produced electron density from the polar regions where auroral precipitation is absent. Based on this criterion we identified more than 500 radar scans during which the radar measured F-region electron densities in the polar cap. The data were sorted according to magnetic activity, solar flux, invariant latitude, and magnetic local time. Besides calculating average electron density profiles, we also examined the altitude and latitude variations of the F-region peak density. The peak density in the polar cap F region varies by a factor of four between solar zenith angles of 85° and 120°. The relative amplitude of large scale irregularities is approximately twice as large when the F-region is not sunlit. The only evidence for transport effects in the measured electron density is a factor of three decrease in peak density in the dawn local time sector for the same solar zenith angle.

INTRODUCTION

The Sondrestrom incoherent scatter radar has been operating in Greenland since March 1982. At an invariant latitude of 74°, it is ideally situated for observing the ionosphere in the dayside auroral zone and polar cap. Thus far, most of the studies involving Sondrestrom radar data have dealt with the ionospheric effects of precipitating particles in the auroral zone and in polar cap aurora (see, for example, De la Beaujardiere et al., 1983, and review by Vondrak and Robinson, 1989). In this study, we characterize the ionization in the polar cap when energetic particle precipitation is not present. In the absence of precipitation, ionospheric electron density is either produced by sunlight or is transported from other regions. The transport effects, which are important above 200 km altitude, make the high latitude ionosphere particularly complex. In this study, we take a statistical approach to characterizing the polar cap electron density distribution. We examine the average F-region electron densities at different local times and for different magnetic and solar illumination conditions. Because of the importance of ionospheric irregularities, we also examine the latitudinal variations in the F-region electron density. The results show the relative importance of magnetic activity, solar illumination and transport in determining polar cap F region electron densities.

DESCRIPTION OF DATA

The data used for this study were obtained while the Sondrestrom radar was scanning in the magnetic meridian plane. The limits of the scan were about 25° elevation angle in the magnetic north and south directions. Each scan was completed in five minutes. The power returned to the radar was analyzed to determine the electron density between 60 and 600 km altitude. The radar transmits two pulses of length 60 μsec and 320 μsec. The returns from both are sampled at times corresponding to 4.5 km in range, but the actual resolution is 9 km for the short pulse and 48 km for the long pulse. Below 150 km we use short pulse densities; above 250 km we use long pulse densities, between 150 and 250 km we use a weighted average of the two. All measurements are Debye-length and temperature corrected. The off-line integration time used in processing the density data was 20 sec. The resulting horizontal resolution is a function of altitude and elevation angle, but is generally between 10 and 30 km. In processing the data we interpolate the measurements to a latitudinal grid of 0.1° or 10 km and an altitude grid of 5 km below 200 km and 20 km above that altitude.

In selecting data for the study we examined a total of 830 scans. These data were obtained during 18 experiments from November 1985 to October 1986. The F10.7 solar flux was generally in the range 70 to 80 so the data presented here is typical of solar minimum conditions. From the 830 scans we selected data obtained during times when the radar measured F-region electron densities poleward of the auroral oval. This selection of data distinguishes this study from previous studies of ionization at very high latitudes. The ability to identify measurements obtained within the polar cap is possible because in the meridian scan mode the radar observes a cross section of the ionosphere over more than 5° of latitude at an altitude of 200 km. At the latitude of Sondre Stromfjord several types of ionization patterns are observed. At times, kev electron precipitation occurs that produces a distinct E-region peak over a broad latitudinal range. This can occur on the dayside during all levels of magnetic activity and on the nightside only during magnetically active conditions. During quiet times on the nightside, auroral precipitation is typically equatorward of Sondre Stromfjord and the distinct E-region is either absent or can be seen near the southern horizon. Under these conditions, the latitudes void of E-region ionization can be assumed to be within the polar cap. During extremely quiet conditions, the polar regions may be filled with weak electron precipitation that may or may not produce a distinct F layer. In this case, the polar cap boundary is poorly defined and its identification solely on the basis of radar data is not possible. In choosing the data for this study, we selected data from two types of observations. In one type, no E-region ionization was present over the entire latitudinal range sampled by the radar. In such situations the polar cap boundary must be to the south out of the radar field of view. In the second type, the auroral zone/polar cap boundary is to the south within the radar field of view and there is a region at least 3° of latitude wide to the north where E region ionization is largely absent. An example of ionization data obtained under these conditions is shown in Figure 1. These data were taken at a magnetic local time of 2200. The ionization below 200 km in the south is the poleward edge of the auroral oval and the region north of that is in the polar cap. The electron density measurements obtained during the 830 radar scans were plotted in the same format as that shown in Figure 1 and each plot was examined to determine which invariant latitudes were poleward of the polar cap boundary. Height profiles of ionization from these latitudes were extracted to create a data base containing only polar cap F-region ionization.

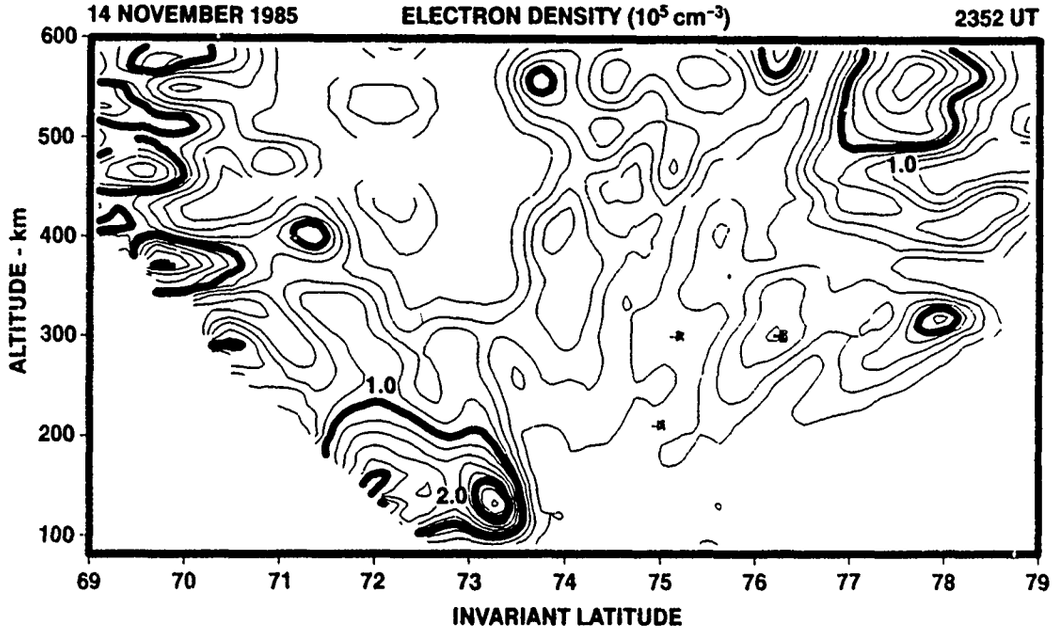


Figure 1. Contour plot of electron densities in the magnetic meridian plane measured by the Sondrestrom radar during an experiment in November, 1985. The data have been ordered according to invariant latitude so that magnetic field lines are vertical in this display.

Two additional points should be made concerning the selection process described above. First, it is possible that there can be significant electron precipitation at energies below 100 eV at latitudes poleward of the boundary defined above. This precipitation may be part of the auroral oval precipitation or may be associated with polar cap aurora. In either case, such events will invalidate our selection criteria which attempts to isolate ionization not being produced in situ by local particle precipitation. Fortunately, situations in which the electron flux is only enhanced at energies below 100 eV are rare. Furthermore, even if these electrons are present, their fluxes are probably small so their effects on the F-region will be minimal. The second point concerns the effects of solar illumination. Much of the data in this study were obtained during times when a measurable E region produced by solar extreme ultraviolet fluxes was present. These data were included in the data base if the E-region did not exceed that expected given the solar zenith angle at the time of observation. The expected E region density for a given solar zenith angle and solar flux condition was determined using the results presented by Robinson and Vondrak (1985). Any particle precipitation present would produce a localized enhancement in E region electron density that would disqualify the data from inclusion in the data base.

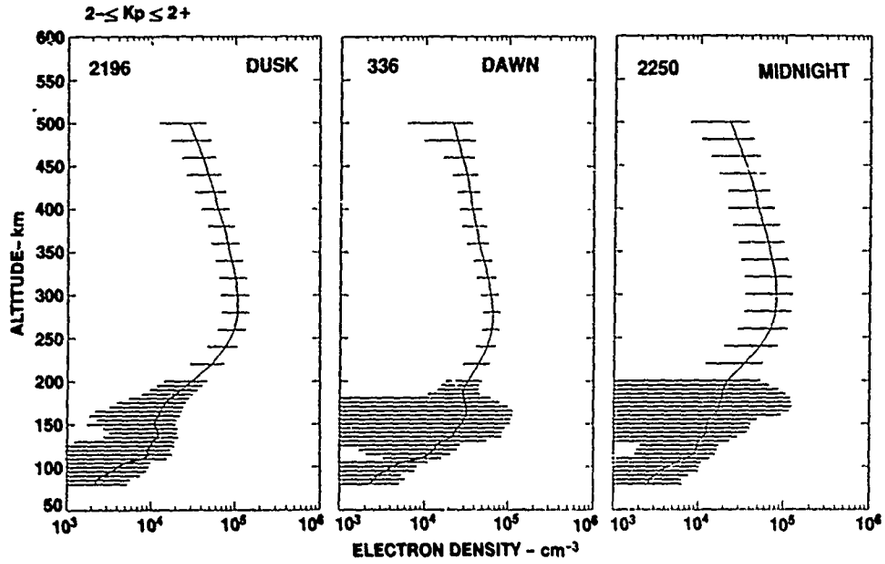


Figure 2. Examples of average electron density profiles for Kp values between 2- and 2+. The number of profiles used in the averaging is given in the upper left in each panel. The horizontal lines show one standard deviation from the mean.

Applying the above selection criteria to the 830 scans resulted in a data base constructed from measurements made during 527 scans. Because electron density profiles were obtained at 0.1°intervals between 73°and 78°latitude, as many as 50 individual profiles were obtained per scan. Because not all of these profiles were within the polar cap, the total number of profiles used for this study was 14,239. These profiles were divided into three magnetic local time quadrants: 09-15 (noon), 15-21 (dusk), 21-03 (midnight), and 03-09 (dawn). The number of profiles in each sector is 593, 4833, 7875, and 938, respectively. Most of the measurements obtained in the noon local time sector were rejected because Sondre Stromfjord is usually within the auroral zone at this local time irrespective of magnetic activity. The dawn sector was poorly sampled in the data because of the relatively fewer number of experiments performed at that local time. In this paper, we show results from the dusk, midnight and dawn local time sectors. These data were binned according to solar zenith angle and magnetic activity. The solar zenith angle bins were as follows: 80°to 90°, 90°to 100°, 100°to 110°, 110°to 120°, and >120°. The magnetic activity was categorized according to Kp as follows: 0 to 1+, 2- to 2+, 3- to 3+, and > 4-. The data from each local time sector was sorted according to these parameters and the electron density profiles averaged by averaging the densities measured at each altitude. An example of the average profiles computed in this manner is shown in Figure 2. Here the average profiles are shown for the three local time sectors for Kp between 2- and 2+. The horizontal bars show one standard deviation from the average. The number of profiles averaged for each is shown in the upper left. Note the large standard deviations in the results for altitudes below 200 km. The large spread in densities shows the influence of solar illumination which can produce E region electron densities up to $10^5 cm^{-3}$.

RESULTS

Solar zenith angle dependence

The electron density profiles in the three local time sectors were sorted according to the solar zenith angle at the time of observation. The average profiles are shown in Figure 3. Not all solar zenith angle conditions were sampled in the different sectors. Figure 3 shows that the solar zenith angle dependence is most pronounced in the midnight local time sector. In the dusk sector the solar zenith angle dependencc is weaker, while in the dawn sector no significant variation with solar zenith angle is observed.

The solar zenith angle dependence of the F region peak is shown in Figure 4 for the three local time sectors. The upper panel shows the variation in the F region peak density and the lower panel shows the variation in the altitude of the peak. The peak density varies by about a factor of four between solar zenith angles of 80°and 120°. The altitude of the F region peak varies between 280 and 320 km and exhibits no systematic change with solar zenith angle. It should be noted that at an altitude of 300 km, the F region peak is sunlit for solar zenith angles less than 107°.

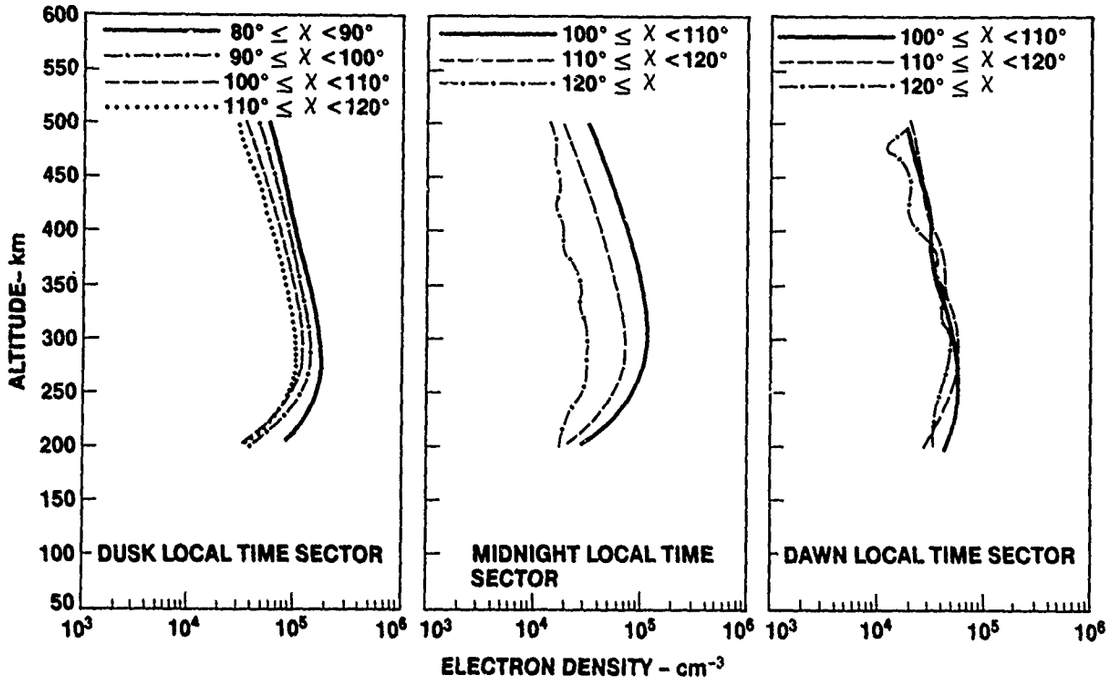


Figure 3. Average profiles in three local time sectors when sorted according to solar zenith angle χ.

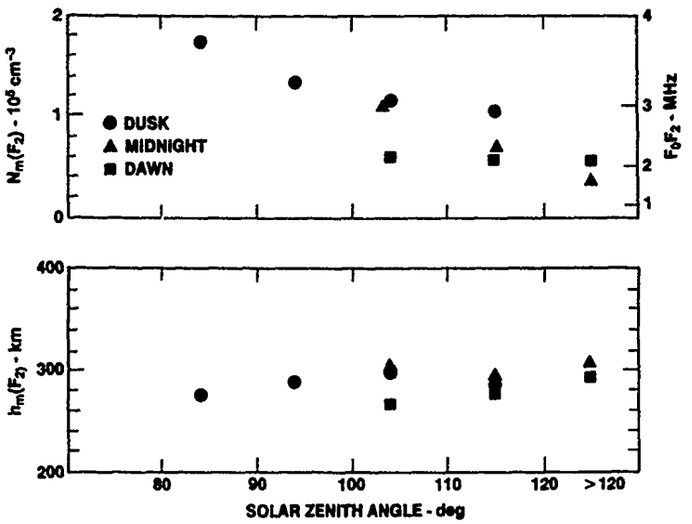


Figure 4. Variation in the F region peak density $[N_m(F_2)]$ and the height of the F region peak $[h_m(F_2)]$ as a function of solar zenith angle. F_0F_2 values in Mhz are given on the right side of the upper panel.

Kp dependence

Figure 5 shows the average profiles for the three local time sectors when sorted according to the magnetic activity index Kp. Although radar measurements were made during times when Kp was greater than 4+, these were eliminated from the data base because of the presence of particle precipitation. It is fairly clear from these results that there is no systematic variation in the F region profiles with magnetic activity. This is also shown in Figure 6 where we display the variations in F region peak density and height as a function of Kp. For the F region peak, note the systematic differences in the densities in the three local time sectors. This reflects the solar illumination effects as shown in the upper panel of Figure 4.

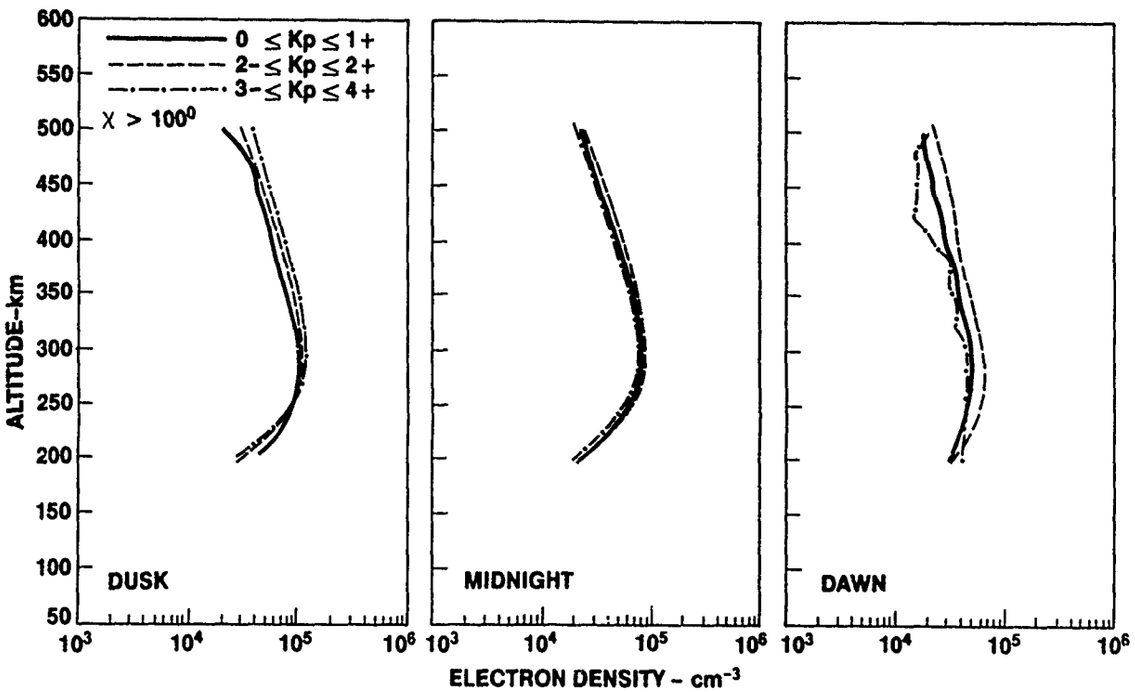


Figure 5. Average profiles in three local time sectors when sorted according to the magnetic activity index Kp.

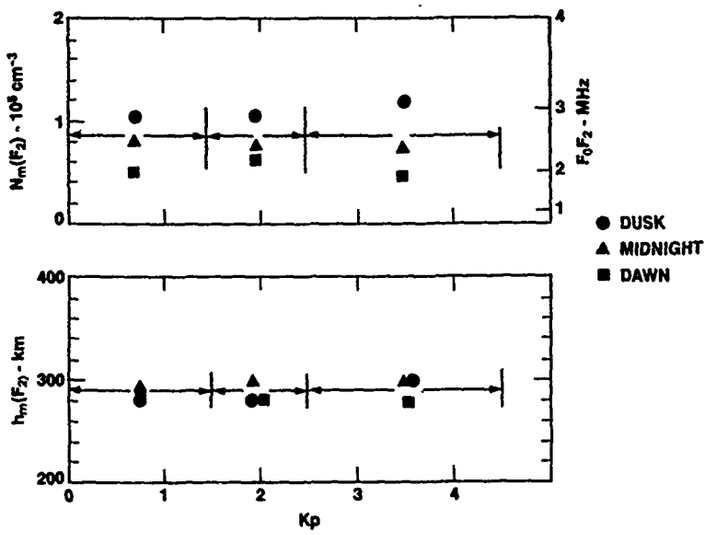


Figure 6. Variation in the F region peak density $[N_m(F_2)]$ and the height of the F region peak $[h_m(F_2)]$ as a function of Kp. F_0F_2 values in Mhz are given on the right side of the upper panel.

Latitudinal variations

The meridian scan data used in this study allow investigation of the latitudinal variations in F region density. The polar cap F region is known to be structured at a variety of scale sizes. The source of this structure is of great interest as it reflects dynamic processes and is intimately tied to scintillation-producing irregularities (see review by Tsunoda, 1989). An example of the way in which the polar cap F region is structured is shown in Figure 7. Each trace gives the latitudinal variation in F region peak density measured during one meridian scan. The Kp and solar zenith angles at the time of the scans are shown in the figure. The vertical scale is percent variation from the mean value N_0. These examples show that the F region peak density can vary by as much as 30 per cent over distances on the order of 50 km. A good deal of smaller scale structure must also be present, but cannot be resolved because of the way the radar data were analyzed for this study.

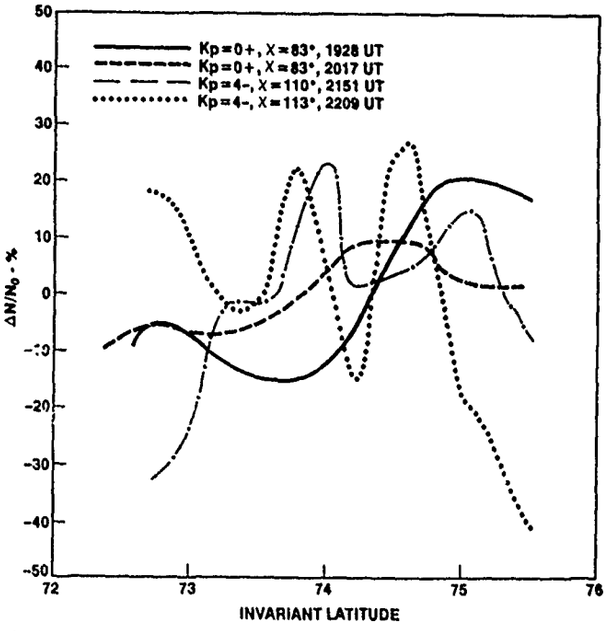


Figure 7. Examples of the latitudinal variations in the F region peak density. The vertical scale is per cent deviation from the mean value N_0 of the F region peak densities measured during each radar scan.

To study these latitudinal variations, we calculated the mean value of $\Delta N/N_0$ for each scan, sorting the scans by solar zenith angle and Kp as above. The average values are plotted as a function of solar zenith angle and Kp in Figure 8. The upper panel shows that the amplitude of the irregularities increases with solar zenith angle. The lower panel shows no systematic variation with magnetic activity. The level of irregularities also shows no obvious change with respect to local time.

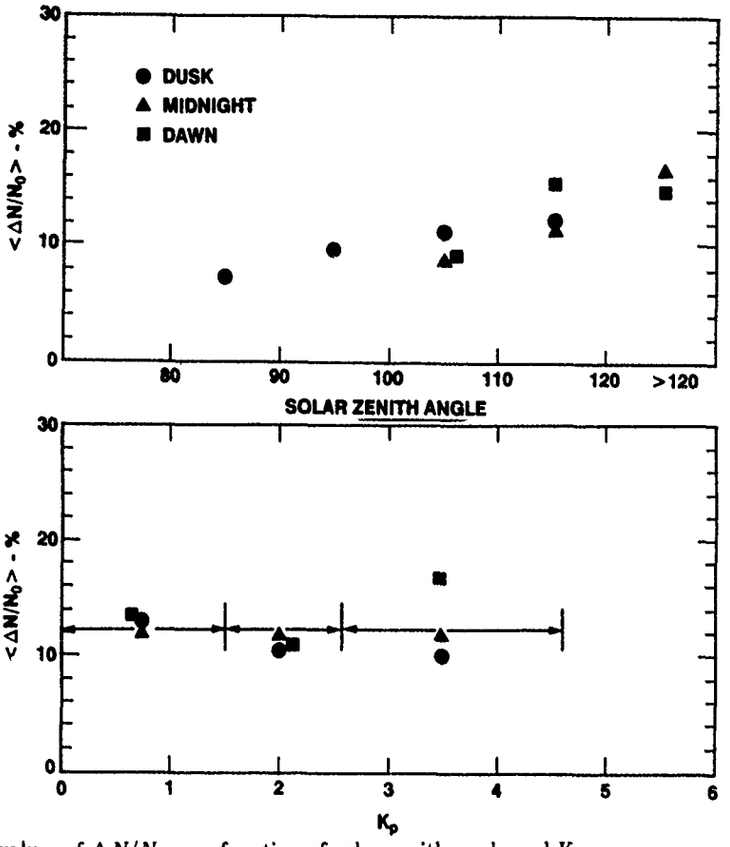


Figure 8. Average values of $\Delta N/N_0$ as a function of solar zenith angle and Kp.

DISCUSSION

The results of this study demonstrate the importance of solar illumination in determining the properties of the polar cap F region. However, more subtle effects can also be seen that may be due to chemical or transport processes. The solar zenith angle variation is summarized in Figure 4. Measurements were made in all three local time sectors for solar zenith angles in the range 100°to 120 °. For the 110°to 120°bin, the F region at 300 km altitude is not sunlit. so the effects of solar illumination at these local times should be minimal. However, the data show that the densities in the dusk sector are statistically higher than those at other local times. This may be related to plasma convection. One possible explanation is that the plasma in the dawn and midnight sector of the polar cap has convected more slowly from the sunlit polar cap. Alternatively, it is possible that the plasma at these local times is convecting in a small convection cell that stays for only a short time in the sunlit hemisphere and therefore never reaches its steady state value. This is the condition that creates the polar hole at high latitudes in winter (Thomas and Andrews, 1969; Brinton et al., 1978), although the depth of the minimum observed here is much less. Sojka et al. (1981) performed a theoretical study of the high-latitude winter F region during solar minimum conditions and found that the depth of the polar hole is strongly dependent on UT or the longitude of the observations. Another possible explanation for the low densities at dawn and midnight is that the plasma in those local time sectors is convecting more rapidly than in the dusk sector. Rapid convection leads to Joule heating that increases the ion temperature and raises the scale height of NO+ (Banks et al., 1974). This increases the recombination rate at F-region altitudes and results in a smaller steady-state F-region density than would otherwise exist.

This study has also demonstrated the lack of influence of magnetic activity on the F region polar cap. Because magnetic activity is an indirect measure of auroral precipitation as well, this result suggests that ionization created by precipitation remains localized to the auroral zone. It has been shown that aurorally-enhanced plasma is transported from the dayside into the polar regions (Knudsen, 1974). However, since the noon sector was so poorly sampled in our data base, we could not examine these effects. At other local times, F region densities enhanced by auroral precipitation are more likely to be convected along the auroral zones back to the dayside, instead of into the polar cap.

The results presented here also show how the spatial structure in F region electron density changes with solar zenith angle. When the F region is sunlit the amplitude of the variations in F region peak density is about 10 per cent of the mean value. When the F region is not sunlit, the amplitudes grow to about 15 per cent. Several studies have examined the role of background electrical conductivities in the development of F-region irregularities (Volk and Haerendel, 1971). The background conductivity tends to short out polarization electric fields that are responsible for various structuring mechanisms. Although the theory was developed for small-scale irregularities, it is possible that the effect occur in the production of large scale F-region structures.

CONCLUSION

This study represents a new approach to studying the polar cap F region in that it isolates electron densities not produced in situ by local particle precipitation. The resulting electron densities are due to the combined effects of solar illumination, transport and chemistry. Because transport is important in the polar regions, these results provide a test for ionospheric models that is free of the complications caused by the presence of particle precipitation. The results presented here apply to solar minimum conditions. Similar measurements are being made on a regular basis at Sondre Stromfjord so that the opportunity to study polar cap electron densities during other phases of the solar cycle will be possible.

Acknowledgements. This work was supported by National Science Foundation Grant ATM-8717840 and the Lockheed Independent Research Program. The SRI International site crew at Sondre Stromfjord provided valuable assistance in maintaining and operating the radar during the experiments. The efforts of Ms. Carolyn Wen in software development and data reduction are gratefully acknowledged.

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, 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, 1978.

De la Beaujardiere, O., J. Holt and E. Nielsen, Early MITHRAS results: The electric field response to substorms, *Radio Sci.*, *18*, 981, 1983.

Knudsen, W. C., Magnetospheric convection and the high latitude F2 ionosphere, *J. Geophys. Res.*, *79*, 1046, 1974.

Robinson, R. M., and R. R. Vondrak, Measurements of E-region ionization and conductivity produced by solar illumination at high latitudes, *J. Geophys. Res.*, *89*, 3951, 1984.

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.*, *86*, 609, 1981.

Thomas, J. O., and M. K. Andrews, The trans-polar exospheric plasma, 3, A unified picture, *Planet. Space Sci.*, *17*, 433, 1969.

Tsunoda, R. T., High latitude F-region irregularities: A review and synthesis, *Rev. of Geophys.*, *26*, 719, 1988.

Volk, H. J., and G. Haerendel, Striations in ionospheric ion clouds, 1, *J. Geophys. Res.*, *76*, 4541, 1971.

Vondrak, R. R. and R. M. Robinson, Electrodynamics of auroral and polar cap arcs at very high latitudes, in *Electrodynamic Coupling in the Polar Clefts and Caps*, P. Sandholt and A. Egeland, Ed., NATO Advanced Study Institute Series, 1989.

AD-P006 315

Comparison of GPS and Incoherent Scatter Measurements of the Total Electron Content

A.J. Coster, E. M. Gaposchkin, L.E. Thornton,
M.I.T. Lincoln Laboratory
M. Buonsanto and D. Tetenbaum
M.I.T. Haystack Observatory

ABSTRACT

GPS and incoherent scatter (IS) measurements of the total electron content (TEC) taken during 1-3 March 1989 are compared. During this period, four different GPS satellites, SV nos. 6, 9, 11, and 12, were observed. The TI4100 GPS receiver at Millstone monitored these satellites continuously while they were in view. At the same time, the Millstone UHF radar was sequentially pointed in the direction of several of these satellites taking incoherent scatter measurements. The incoherent scatter measurements produce profiles of the electron density distribution. The IS profiles are then integrated to produce TEC measurements up to 800 km. The combined pseudo-range and phase data of the GPS system measure the group delay at both the L1 (1575.42 MHz) and L2 (1227.6 MHz) frequencies. This information is converted to TEC measurements along the line of sight to the satellite at 19000 km. The comparison of the GPS and incoherent scatter data gives us a new technique for estimating the number of electrons above 800 km. Our results were surprising. The 1-3 March time period was associated with interesting geophysical conditions. The experiment began during quiet geomagnetic conditions and at a time when the daily 10.7 cm solar flux values were moderate (162). On 1 March, a large electron content, sometimes greater than 30 TEC units, was observed in the ionosophere or plasmasphere above 800 km. This large TEC value is surprisingly high when compared to previous studies, taken during solar minimum conditions, which found only 2-6 TEC units of plasmaspheric content above 2000 km (Klobuchar, et al., 1978). During the middle of the experiment, the geomagnetic activity increased. An electron density trough was then observed to form towards the north and propagate southward. In addition, the total number of measured electrons above 800 km fell dramatically.

INTRODUCTION

At Lincoln Laboratory a transportable device is being developed which uses the Global Positioning System (GPS) to determine real-time ionospheric path delays. We are using a TI4100 GPS receiver which is capable of receiving the dual L-band frequencies (1227.6 and 1575.42 MHz) broadcast by each satellite. Up to 21 different GPS satellites will eventually be on station in near-circular, semi-synchronous orbits at altitudes of approximately 19000 km. The plan is that at least 4 satellites will be in view at any place, 24 hours a day (Spilker,1978).

Because our GPS receiver is located at the Millstone Radar site, we have the advantage that the GPS data can be compared directly with incoherent scatter (IS) measurements of the ionosphere. Incoherent scatter measurements are taken routinely with the UHF radar at Millstone Hill. During the period 1-3 March 1989, the UHF radar was sequentially pointed in the direction of several GPS satellites taking incoherent scatter measurements. The IS data was then processed to give an alternative determination of the total electron content (TEC) along the line of sight to each GPS satellite (Buonsanto,1989). The IS data is carefully calibrated at the beginning and end of each run using foF2 data from a colocated Digisonde operated by the University of Lowell. The IS radar provides electron density profiles from approximately 100 km to 1000 km or higher. Because of the decreasing signal-to-noise ratio, however, data at

91-09687

the greatest heights are not always reliable. For the present experiment, the TEC is derived by integrating the electron density profiles from 100 km to 800 km in height.

In principle, it should be possible to obtain the ionospheric path delay directly from the GPS group delay measurements at the two L-band frequencies. The delay introduced by the ionosphere is a function of both the TEC and frequency. However, although the GPS range data is accurate enough for the majority of user needs, the ionospheric correction cannot be determined directly from this data. This is primarily because of multipath effects and the L1/L2 biases.

Multipath effects in the group delay data account for approximately 6 ns of increase in the noise level. For comparison, a 2 ns increase in group delay amounts to an error in the path delay at L-band of 1 meter. At Millstone, the multipath effects have been reduced by use of absorbing material under the antenna and can be further reduced by data averaging (Sciegienny, private communication). The multipath problem is not as severe in the GPS phase advance measurements. The GPS phase data show an increase in the noise level on the order of only 50 ps (Dahlke et al.,1988). The trouble with the phase advance data is that they contain an unknown bias due to the integer cycle ambiguity of phase measurements, i.e. the number of full phase cycles between the satellite and receiver is not known. By combining the GPS phase data with the group delay data we may overcome the difficulties inherent with both data types.

Before going any further, we would like to illustrate these three different types of data. Figure 1 shows the GPS phase data, the GPS group delay data, and the Millstone UHF incoherent scatter data plotted as a function of time for one satellite. The increased noise level in the group delay data is apparent. To account for the unknown cycle ambiguity and bias, the mean value of the phase data is set to be the mean value of the group delay data. The IS data represents the integrated TEC up to an altitude of approximately 800 km along the line of sight to the satellite.

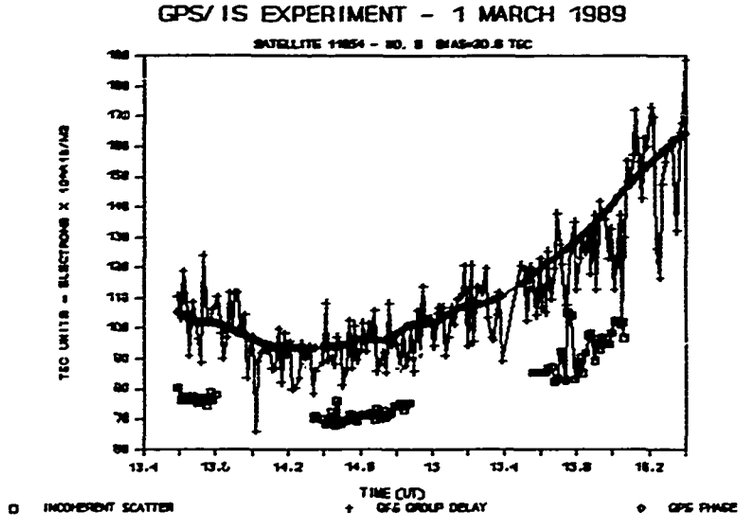


Figure 1. Illustration of Incoherent Scatter and GPS Phase and Group Delay Data.

The IS derived TEC should be less than that derived from the GPS system which measures out to 19000 km. Note that a total bias of 30.6 TEC units is evident between the GPS data and the IS data. This bias, which we will refer to as the IS bias, can be attributed to additional ionosphere above 800 km and to the combined satellite and receiver bias.

The combined satellite and receiver biases are known as the L1/L2 biases. The L1/L2 biases correspond to the additional differential delay between the two frequencies introduced by the satellite and receiver hardware. These biases must be measured (or estimated) and removed from the GPS group delay data in order to accurately determine the ionospheric path delay. These biases have

been studied extensively (Lanyi, et al.,1987, Dahlke, et al., 1988, Coster and Gaposchkin,1989) and the results of those studies will be presented in the next section.

In this paper we are going to discuss the comparisons of the GPS and incoherent scatter measurements of TEC taken during 1-3 March 1989. During this period we observed 4 different GPS satellites, SV nos. 6, 9, 11, and 12 (corresponding to SSC nos. 11054, 11783, 14189, and 15271). The GPS observing schedule from Millstone for these satellites is given in Table I. There is a 5 hour difference between local time at Millstone and universal time, so that the daytime experiments started at 8:30 am and the nighttime experiments started at 11:15 pm.

TABLE I. Experiment Times.

	Day (UT)		Night (UT)	
SV #.	1 March	2 March	2 March	3 March
6	13:35-16:23	13:28-16:19	4:17-7 28	4:13-7:24
9	13:35-18:08	13:28-16:19	4:17-9:45	4:13-9:41
11			4:17-9:45	4:13-9:41
12	15:30-18:08	15:24-18:03	5:39-9:45	5:37-9:41

Interesting geophysical conditions occurred during the 1-3 March time period. Although 1989 is generally associated with high solar flux conditions, the daily solar flux values during this time were only at moderate levels. The 90 day mean F10.7 cm flux value was 213 throughout, while the daily F10.7 cm flux values were, respectively, 169, 174, and 169. Prior to 2 March, the geomagnetic activity had been quiet for a period of about a week. The highest Kp value on 1 March was 3+, with the average Kp for that day being 3-. On 2 March, however, the Kp went from a 3+ value between 0-3 hours UT to a value of 6- between 3-6 hours UT. The Kp then fell to 4- from 6-9 hours UT. This means that during our nighttime pass on 2 March moderate geomagnetic storm conditions existed. The Kp was again moderately high during our daytime pass on 2 March with the Kp values being 3+ and 4. Finally, during the nighttime observations on 3 March, the Kp was 5 between 3-6 UT and 6- between 6-9 UT. Geophysical data for this period is summarized in Table II.

Table II. Geophysical Parameters.

DATE	F10.7(daily)	F10.7(avg)	Kp (3 Hour Values)
1 March 89	169	213	3+3 2 2-3+3-2 2+
2 March 89	174	213	3+6-4-3+4 3+3 3+
3 March 89	169	213	3 4 6-5-5-5 4 3

L1/L2 BIASES

To use the GPS group delay data accurately, one must first estimate the satellite (SV) and receiver L1/L2 biases. Coco et al.(1990) have shown that although the TI4100 receiver bias varies from unit to unit, the change over time in the bias is quite small. They found the day-to-day variation of the TI 4100 receiver bias over a 5 week period to be less than 0.5 ns. In this paper, we are going to assume that our receiver bias is neglible. Regardless of our assumption, however, the receiver bias should be common to all satellites observed, and it should be a stable value.

Pre-launch calibration values of the satellite, or SV (space vehicle) L1/L2 biases were measured by Rockwell. Since then, other studies have developed techniques for determining the SV biases using a multistation system of receivers (Dahlke, et al.,1988, Lanyi, et al.,1987) At Millstone, we estimate the SV biases using a single receiver at one site (Coster and Gaposchkin,1989). The estimated SV biases are listed in Table III. The ARL values are from Coco et al.(1990). The standard deviations are given in parenthesis. The ARL SV biases all have an associated standard deviation of 0.5 ns.

It is important to notice that the biases ARL determined remained fairly constant from the year 1987 to 1989. Our technique for determining the SV biases produced similar results, with the exception of satellite no. 6, which has a large standard deviation. Our method for determining these satellite biases does not properly account for the difference in electron content above

TABLE III. Satellite (SV) L1/L2 Biases in Nanoseconds of Differential Delay between L2 and L1.

SV PRN	Pre-Launch	ARL 87	ARL 89	MIT LL 89
3	-0.24	1.06	1.11	
6	2.33		0.29	4.41(1.79)
8	1.53			
9	-0.34	-0.34	-0.34	0.40(0.63)
11	3.05	3.16	3.16	2.94(0.54)
12	0.02	2.96	2.95	2.17(1.21)
13	1.62	1.79	2.36	

800 km as a function of location, and thus we assume that it is our estimate of the no. 6 bias that is in error. We are going to adopt the recovered 1989 ARL SV biases for our discussion here.

It is useful to present our adopted SV biases in units of TEC since those are the units that the IS Biases are reported in. The conversion between ns of differential delay and TEC units (equal to number of electrons x 10**16 m-2) is given below (Dahlke, 1986):

$$TEC = 2.85 * TDD, \text{ where} \qquad 1)$$

TDD = TTL2 - TTL1 (transit time of L2 - transit time of L1 in ns)

The 1989 ARL SV biases, given in TEC units, are listed in Table IV.

Table IV. SV Biases in TEC Units.

SV No.	ARL '89
6	0.83
9	-0.97
11	9.01
12	8.41

SATELLITE NO. 9.

We begin our discussion with satellite no. 9 since it has consistently been shown to have a small SV bias (-0.97 TEC), and a small standard deviation (0.61 TEC). Figure 2 shows the Millstone viewing geometry for this satellite in azimuth and elevation during 1-3 March 1989. The start of the nighttime pass is delimited by a small square and the start of the daytime pass by a small oval. During the night, the satellite moves from the northwest to the southwest and during the day it moves from the southeast to the northeast. Figures 3a-d show the combined incoherent scatter and GPS phase and group delay data for this satellite during the experiment. Note that on 1 March there is an IS bias of 33 TEC units (34 TEC units if we account for the SV bias). Recall that the IS biases include the bias due to the additional ionosphere above 800 km plus the satellite and receiver bias. In our discussion we will give the total measured IS bias followed by, in parenthesis, the IS bias minus the estimated SV bias. It is this second value which reflects our estimate of the number of electrons above 800 km. On 1 March, when the IS data measures 60-110 TEC units, the GPS system measures 90-140 TEC units. The implication of this is that as much as one-third of the TEC on 1 March is above 800 km, or that 5.4 m of delay at L1 can be attributed to electrons above 800 km. On the following day, 2 March, after the geomagnetic storm had begun, an IS bias of only 3.73 TEC units (4.7 TEC), or 0.61 m of delay at L1, is measured.

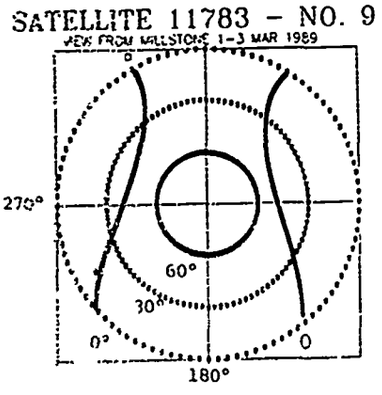


Figure 2. Geometry for SV # 9 from Millstone in Azimuth and Elevation.

On the night of 2 March, during the initial geomagnetic disturbance, the measured IS bias of satellite no. 9 was 8.14 TEC units (9.1 TEC). Smaller biases are expected at night due

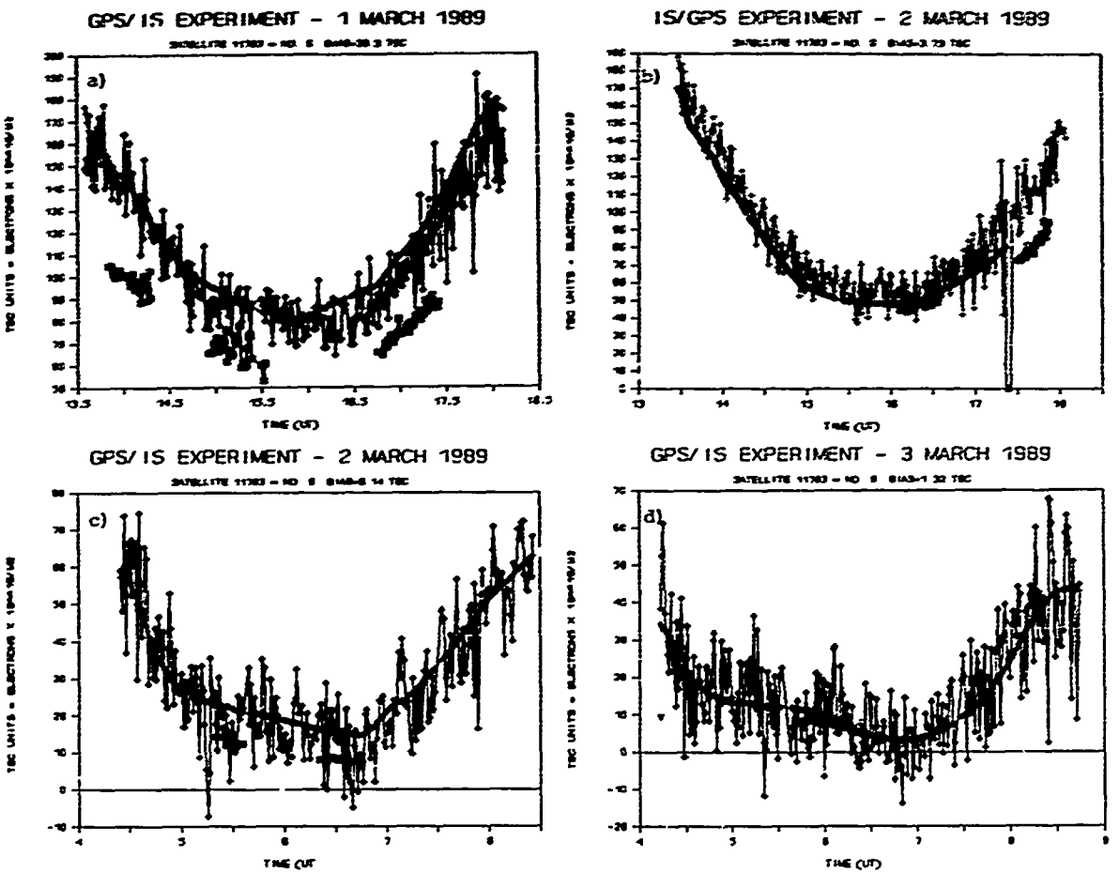
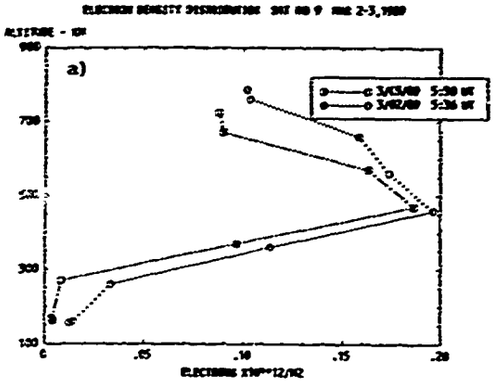


Fig. 3a-d. Comparison of GPS and Incoherent Scatter TEC Data for SV no. 9 on 1-3 March 1989. The IS data are demarked by a square, the GPS phase data by a cross, and the GPS group delay data by a diamond. Figs. 3a and b show the daytime data, while Figs. 3c and d refer to the nighttime data.

to the smaller electron content. Nevertheless, the total TEC measured by the GPS system, during the time for which we have GPS observations, varied between 16-25 TEC units, indicating that as much as half of the ionosphere on this night was above 800 km. The 3 March nighttime data, which was also taken during storm conditions, shows a bias of only 1.32 TEC (2.29 TEC) units. In fact, if the beginning and end of the IS data are compared on this day, we see that initially there was a bias of nearly 3 TEC units, while by the end of the pass, the IS bias is almost 0. This observation is confirmed by the associated IS profiles of the ionosphere shown in Figs. 4a-b. At the beginning of the pass, shown in Fig. 4a, a standard electron density distribution can be observed on both days. However, in Fig 4b by the 6:30 UT, the IS data on 3 March shows a very depleted electron density distribution. Note also that it is at this time that the smallest bias between the GPS and the IS data is measured. This suggests that what we are seeing a real ionospheric or plasmaspheric effect, such as an ionospheric trough, rather than a calibration or other type of problem.



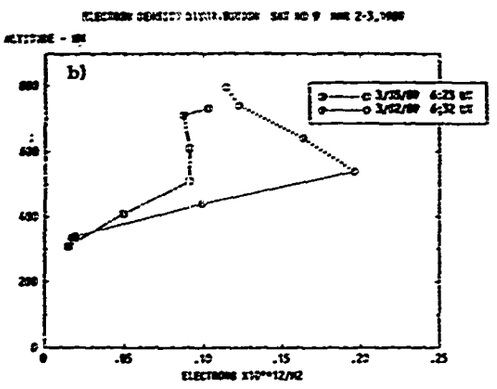


Figure 4a. **Figure 4b.**
IS Determined Electron Density Distribution as a Function of Altitude.

SATELLITE NO. 12.

On 1-3 March 1989, satellite no. 12 was observed to follow a path from the northwest to the southwest during the night and from the southeast to northeast during the day, see Figure 5. For this satellite, the estimated SV bias is approximately 6.73 TEC units. As can be seen in Fig. 6a, the measured IS bias was 27.4 TEC units on 1 March during the daytime. Subtracting the assumed satellite bias, we are left with 20.7 TEC units that can be attributed to the ionosphere above 800 km. The IS bias drops to 19.9 TEC (or 13.2 TEC once we subtract the satellite bias) on 2 March. In addition, we observe that the peak electron density is observed to about 118 TEC units on 1 March while on 2 March the peak is only 98 TEC units.

The ionospheric structure looks very smooth during both sets of daytime passes. This is in contrast to the nighttime data on 2 or 3 March. Both sets of nighttime data show disturbed ionospheric structure. On 2 March the observed IS bias was 25.3 TEC units (18.6 TEC). At approximately 8.2 UT, a bump in the GPS phase data can be observed followed by a rapid rise in the estimated TEC content. This is partly due to a change in elevation, although the associated elevation was only 23.5 degrees at this time. On 3 March, the GPS satellite is

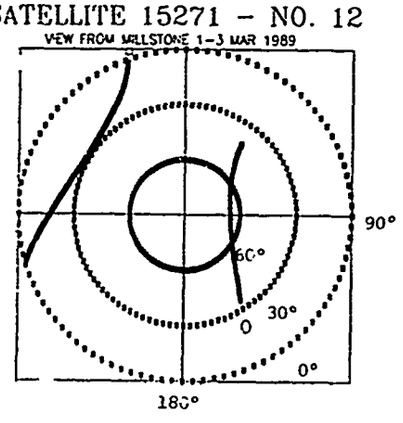
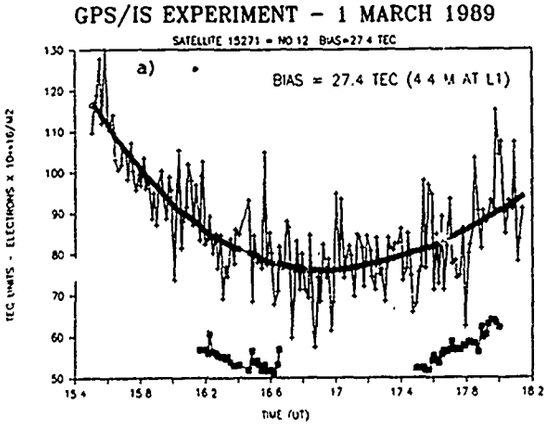
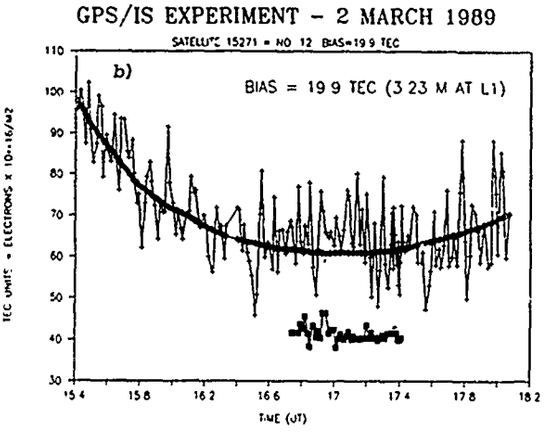


Figure 5. Geometry for SV # 12 from Millstone in Azimuth and Elevation.





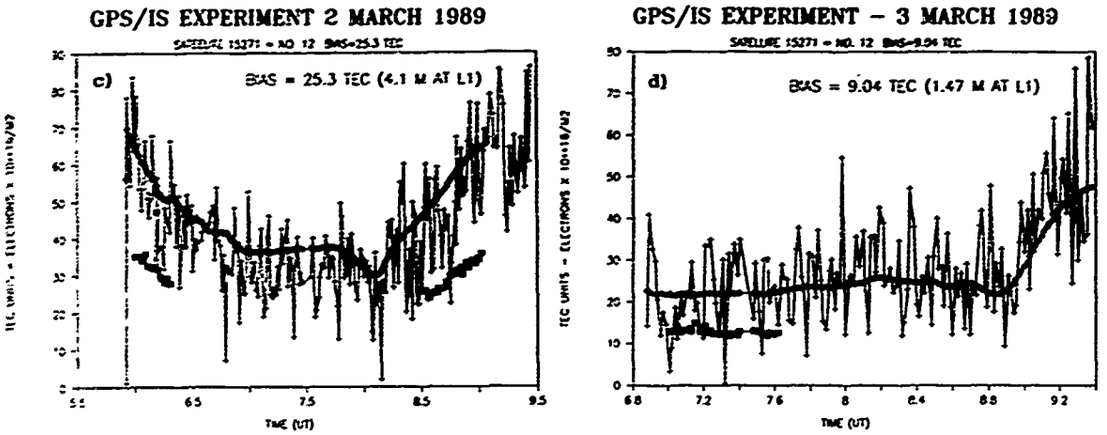


Fig. 6a-d. Comparison of GPS and Incoherent Scatter TEC Data for SV no. 12 on 1-3 March 1989. The IS data are demarked by a square, the GPS phase data by a cross, and the GPS group delay data by a diamond. Figs. 6a and b show the daytime data, while Figs. 6c and d refer to the nighttime data.

following the same observing geometry as on 2 March, except that it is viewed 4 minutes earlier. Nevertheless, the bump is not observed until approximately 8.9 UT, with an associated elevation of 10.4 degrees. In addition, on 3 March, the measured IS bias is only 9.04 TEC units (2.31 TEC). Recall that during the night, this satellite moves from the north to the south. We think that we are observing an electron density depletion or trough that was generated by the 2-3 March geomagnetic storm. The trough propagated southward from 2 March to 3 March.

SATELLITE NO. 6.

Satellite no. 6 travels from the northwest to the southwest during the night and remains at a fairly high elevation throughout the pass (see Fig. 7). During the daytime, the satellite travels from the southeast to the northeast. As can be viewed in Figs. 8a and b, the daytime bias changes relatively little from the 1st to 2nd of March. A bias of 30.6 TEC units (29.8 accounting for the satellite bias of .83 TEC) is seen on the 1 March. This bias drops to 26.4 TEC (25.6 TEC) during the day on 2 March. The bias satellite no. 6 sees on 1 March is somewhat misleading. If one looks only at the section of the pass that is repeated on the 2 March, the bias on 1 march is 38.5 TEC (37.7 TEC). The maximum TEC dropped from 170 on 1 March to 140 on 2 March.

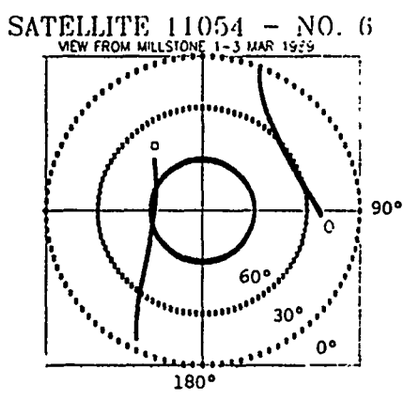


Figure 7. Geometry for SV # 6.

At night, the measured IS bias dropped from 12.6 TEC (11.8 TEC) on 2 March to 3.3 TEC (2.47 TEC) on 3 March. Again, the measured IS bias on 2 March is somewhat misleading. The earlier IS data taking during the 2 March nighttime pass has a bias of only 5.5 TEC.

It is looking towards the south where the bias gets larger. This is yet another indication that the electron density increases looking towards the south. It is also interesting to observe the flat slope of the 3 March data from 4.2 to 5.8 UT, and compare this shape to the 2 March data. During this time period, the TEC data remains less than 10 TEC (9.2 TEC), indicating the presence of a trough.

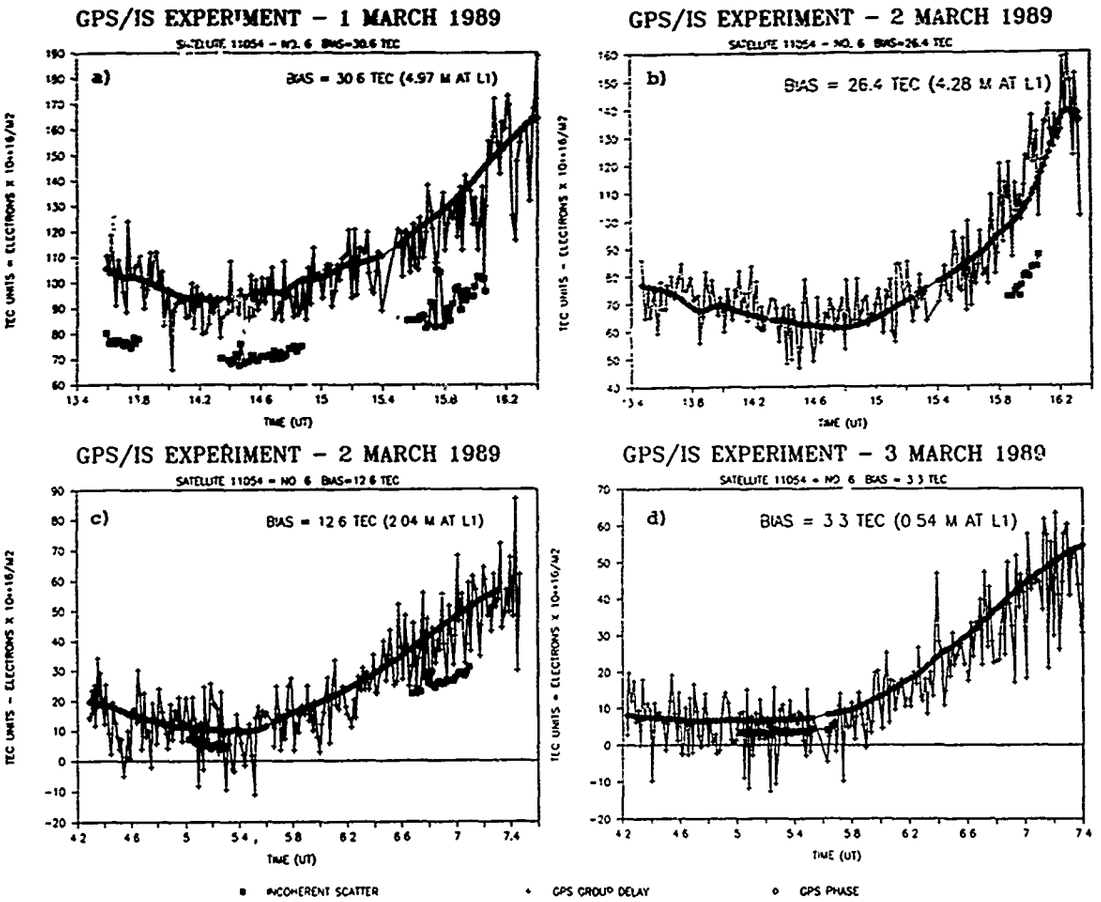


Fig. 8a-d. Comparison of GPS and Incoherent Scatter TEC Data for SV no. 6 on 1-3 March 1989. The IS data are demarked by a square, the GPS phase data by a cross, and the GPS group delay data by a diamond. Figs. 8a and b show the daytime data, while Figs. 8c and d refer to the nighttime data.

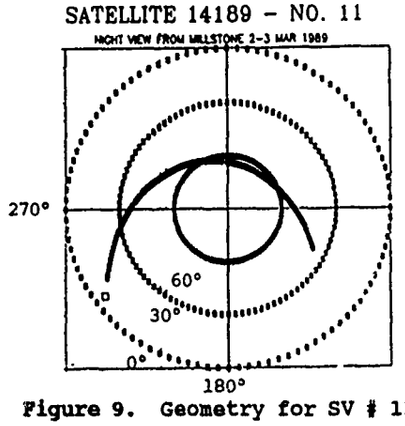
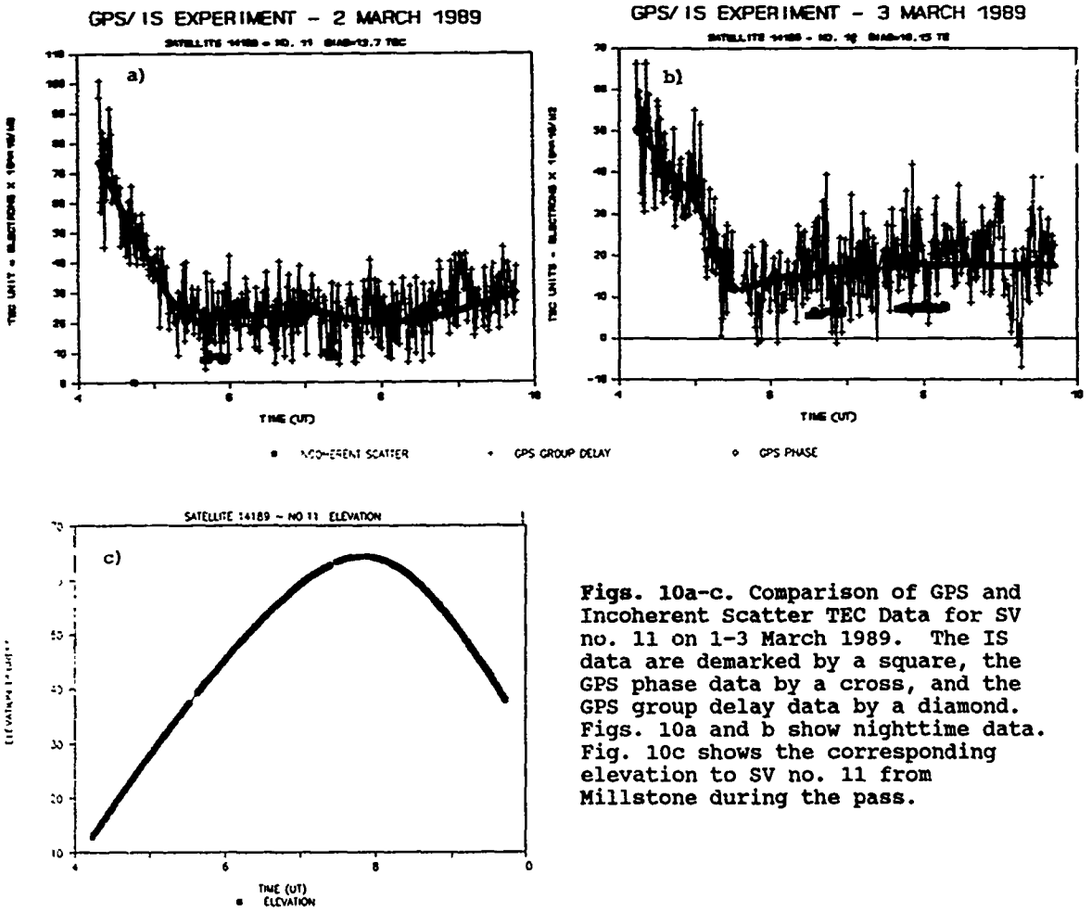


Figure 9. Geometry for SV # 11

SATELLITE NO. 11.

Satellite no. 11 was only viewed during the night observing sessions from Millstone. Nevertheless, it has some of the more interesting geometry as is indicated in Fig. 9. This satellite starts in the southwest, moves over head, and then heads towards the southeast. The measured SV bias for this satellite is 9.01 TEC. On 2 March a bias of 13.7 TEC (4.7 TEC) is seen in Fig 10a and on 3 March a bias of 10.15 TEC (1.15 TEC) is seen in Fig. 10b. Note the precipitous decline in the TEC observed in both sets of data. On both days the TEC flattens out at about 5.20 UT. Yet, on both days, the elevation is greater than 20 degrees by 4.6 UT and greater than 30 degrees by 5.20. We think this structure is yet another indication of the greater electron density towards the southwest.



Figs. 10a-c. Comparison of GPS and Incoherent Scatter TEC Data for SV no. 11 on 1-3 March 1989. The IS data are demarked by a square, the GPS phase data by a cross, and the GPS group delay data by a diamond. Figs. 10a and b show nighttime data. Fig. 10c shows the corresponding elevation to SV no. 11 from Millstone during the pass.

SUMMARY AND CONCLUSIONS

A new technique for estimating the number of electrons above 800 km has been validated through the comparison of the IS and GPS TEC data. The number of electrons that are detected to be above 800 km, sometimes as great as 30 TEC units, is surprisingly high when compared to previous studies of the plasmasphere which showed only 2-6 TEC units above 2000 km. (Klobuchar, et al., 1978). The previous studies were all taken during minimum solar flux conditions. Our estimates of the additional TEC above 800 km are given in Table V. All of the satellites estimate a smaller TEC above 800 km on 3 March than on 1 March, a decrease which varies from 18 to 32 TEC units.

Table V. Summary of Estimated TEC above 800 Km

	DAY (UT)		NIGHT (UT)	
	1 March	2 March	2 March	3 March
SV #				
6	29.8	25.6	11.8	2.5
9	34.3	4.7	9.1	2.3
11			4.7	1.6
12	20.7	13.2	18.6	2.3

The data for satellite no. 9 is slightly anamolous because it shows a large decrease in TEC during the day on 2 March. This is interesting because the observation was taken looking to the northeast at approximately 16.1-16.6 UT with an azimuth of 291 degrees and an elevation

between 38 and 47 degrees. A second incoherent scatter observation of this satellite, also measuring a small IS bias, was taken at 17.5 UT with an azimuth of 299 degrees and an elevation of 6.2 degrees. If we consider the geometry of this measurement, it is approximately 1.5 hours after satellite no. 6 observes approximately the same location (295 azimuth and 8 degrees elevation). In addition, satellite no. 12 looks at a similar part of the sky at a slightly higher elevation (azimuth 290 and elevation 60) directly after the first set of IS measurements of satellite no. 9. Neither satellite no. 6 nor satellite no. 12 observed a small IS bias on this day. Either we are observing the propagation of an electron density trough during the daytime, or we are observing changing satellite biases for at least one of the satellites. However, recall that satellite no. 9 had the smallest SV bias measured. The change in the measured IS bias is 10 nanoseconds, far greater than the estimated SV bias fforthis satellite. We therefore believe that an electron density trough is being observed moving southward during the day.

In conclusion, using the combination of our GPS receiver and Incoherent Scatter radar, we think that we observe a large electron content in the ionosphere and plasmasphere prior to the geomagnetic storm. As the result of the geomagnetic storm, this excess plasma is blown away and, during the following two day period, the electron density trough that is generated propagates southward. We think that the combination of the GPS and incoherent scatter data offers an exciting new technique for measuring the total electron content above 800 km.

ACKNOWLEDGMENTS

We are grateful to the University of Lowell Center for Atmospheric Research for providing the digisonde measurements and also to R. Abbot for providing us with the plots of the Millstone viewing geometry. We would also like to thank J. Klobuchar and D. Coco for their instructive comments. This work was sponsored by the Department of the Air Force. The views expressed are those of the authors and do not reflect the official policy or position of the U.S. Government.

REFERENCES

Buonsanto, M.J., Comparison of incoherent scatter observations of electron density and electron and ion temperature at Millstone Hill with the International Reference Ionosphere, J. Atmos. Terr. Phys., 51, no. 5, 441-467 (1989).

Coco, D. S., C. Coker, S. R. Dahlke, and J. R. Clynch, "Variability of GPS satellite Differential Group Delay Biases," submitted to Radio Science, March 1990.

Coster, A.J. and Gaposchkin, E.M., " Use of GPS Pseudo-Range and Phase Data dor Measurement of Ionospheric and Tropospheric Refraction, " Proceedings of ION GPS-89, Colorado Springs, Colorado, 27-29 September, 1989, 439-443.

Dahlke, R., Scott, D.S. Coco, C.E. Coker, C.E., "Effect of GPS System Biases on Differential Group Delay Measurements," ARL-TP-88-17, August 1988.

Klobuchar, J. A., M. J. Buonsanto, M. J. Mendillo, and J. M. Johanson, "The Contribution of the Plasmasphere to the Total Time Delay," in The Effect of the Ionosphere on Space and Terrestrial Systems, ed. John M. Goodman, vol. 2, p. 486-489, 1989.

Lanyi, G.E., Roth, T., Neilan, R.E., "A Comparison of Mapped and Measured Total Ionospheric Electron Content Using GPS and Beacon Satellite Observations," in The Effect of the Ionosphere on Communication, Navigation, and Surveillance Systems, ed. John M. Goodman, J. A. Klobuchar, R. G. Joiner, and H. Soicher, 135, 1987.

Sciegienny J., private communication, 1989.

Spilker J.J., Jr., GPS Signal Structure and Performance Characteristics, Navigation, # 2, 25, (1978)

AD-P006 316

MEASUREMENTS OF TRANS-IONOSPHERIC EFFECTS USING SIGNALS FROM GPS

G. J. Bishop, J. A. Klobuchar, and Santimay Basu
Ionospheric Physics Division
Geophysics Laboratory (AFSC)
Hanscom AFB, MA 01731 U.S.A.

J. R. Clynch, D. S. Coco, and C. Coker
Applied Research Laboratories
The University of Texas at Austin
P. O. Box 8029
Austin, TX 78713-8029

ABSTRACT

Effects of the ionosphere can be significant to modern RF systems which depend on trans-ionospheric signal propagation. Delay effects caused by ionospheric total electron content, (TEC), and scintillation effects due to small-scale ionospheric irregularities, are of particular importance since they can, for example, cause ranging errors and signal or target fades. Signals from the Global Positioning System (GPS) satellites will provide the opportunity to measure TEC and scintillation parameters on a continuous basis in several directions from anywhere on earth. GPS signals have been used since 1984 to make measurements of polar cap TEC and L-Band scintillation, at Thule, Greenland. Concurrent observations of amplitude fading at 250 MHz link these GPS measurements to a large data base of polar scintillation at that frequency. Examples of these data will be given which illustrate the extreme variability of the polar cap ionosphere, which can be a significant limitation to trans-ionospheric systems which must operate in this part of the world. This paper will also review ionospheric measurement capability of GPS receivers using the full GPS P-code capability, contrasted with newer, less costly, instruments which make much more limited measurements without requiring knowledge of the GPS P-codes.

INTRODUCTION

Effects of the ionosphere on radio signals which traverse it can represent one of the most important potential problems to the design and operation of modern RF systems which operate in the earth-space environment. Ionospheric measurement using GPS signals is the latest development in a progression of techniques, starting with observation of effects on the signals from the first low earth orbit satellites. Subsequently, monitoring signals from geostationary satellites of opportunity led to the development of a world-wide network of polarimeters monitoring TEC for the Air Weather Service (AWS). With GPS, it is possible to measure TEC, phase scintillation and amplitude scintillation simultaneously along multiple lines of sight. As the GPS constellation is completed it will be possible to make such measurements in four or more look directions, simultaneously, from any point on earth. Ionospheric measurements performed with several models of GPS receivers since 1979 have developed the technology base necessary for a new ionospheric monitoring system to replace the AWS polarimeter network, which is being phased out as there are progressively fewer geostationary UHF beacons to observe. This new Trans-Ionospheric Sensing System (TISS) will provide the AWS Space Forecast Center (SFC) with real-time reporting of TEC and scintillation, in multiple look directions, from stations throughout the world, [Bishop et al., 1989].

IONOSPHERIC EFFECTS ON SYSTEMS

There are several effects which the ionosphere can have on radio waves which must propagate through it. Modern transionospheric RF systems can be limited in performance

91-09860

91

by one, or more, of these effects and systems planners need to take their potential impacts into account, both in system operation and in new system design. Such systems include active communications and navigation satellites as well as both ground-based and potential space-based ranging systems.

The major trans-ionospheric effects on radio signals at typical system frequencies are: 1) delay of the signal, 2) Faraday rotation of linearly polarized signals, 3) rapid changes in signal delay, and 4) signal fading. These effects are caused by different ionospheric parameters. Specifically, 1) above is due to the total number of electrons along the signal path through the ionosphere, where the number of electrons is measured in an equivalent square meter column, and is defined as the Total Electron content, or TEC, with 1×10^{16} electrons/m^2 defined as one TEC unit; 2) is due to the product of the total number of electrons multiplied by the longitudinal component of the earth's magnetic field along the space object to observer path. Phase scintillation, 3), is due to the time rate of change of the electron content, and 4) is produced by diffraction from irregularities of scale size of approximately one kilometer, along the path. Depending upon the system design, these major effects have the potential to degrade system performance in varied ways, such as: data loss, ranging errors, and degradation of detection, tracking or imaging.

There are other effects which the ionosphere can have upon the characteristics of radio waves which propagate through it, such as distortion of the modulation envelope and angular refraction, or bending, of radio waves. However, these two effects are generally only potential problems for large percentage signal bandwidths, or low elevation angles, respectively.

While some of these effects have been studied since the beginning of the artificial satellite era in the late 1950's, adequate models of these effects on RF systems still do not exist. This lack of suitable models is due in part to the high temporal and geographic variability of the ionosphere, and partly due to the increasing radio signal accuracy requirements and special geometries of various RF systems which must operate through the ionosphere.

Most of the potential effects listed above are proportional, in some manner, to the total number of electrons through which the radio wave travels on its way from its transmitter to the receiver. A more complete discussion of the magnitude of many of these potential effects is available in the Handbook of Geophysics and the Space Environment, Sections 10.7, 10.8 and 10.9, [Jursa, 1985]. Table 1 summarizes the major trans- ionospheric effects and their interaction with systems.

Table 1. Trans-ionospheric effects on systems, (after Bishop et al., AGARD '89)

IONOSPHERIC PARAMETER	RADIO (RF) EFFECTS	SYSTEMS EFFECTS	POTENTIAL DEGRADATION OF:
TOTAL ELECTRON CONTENT (TEC)	SIGNAL DELAY	RANGE ERROR	TARGET LOCATION
	FARADAY ROTATION	SIGNAL LOSS	SURVEILLANCE
AMPLITUDE SCINTILLATION	FADES & ENHANCEMENTS	SIGNAL FADES TARGET FADES	MESSAGE CONTENT TARGET UPDATE
PHASE SCINTILLATION	RAPID CHANGES IN SIGNAL DELAY	RAPID CHANGES IN APPARENT RANGE	DETECTION TRACKING IMAGING

MEASURING IONOSPHERIC EFFECTS USING SIGNALS FROM GPS

GPS SYSTEM: WORLD-WIDE MULTI-DIRECTIONAL COVERAGE

The Navstar GPS operational constellation, when fully in place, will have several satellites visible at all times from any point on earth. In many cases, five or more satellites will be visible, [Runkle et al., 1989]. Because these several satellites will be seen in varying look directions from any given point, the geographic region of the ionosphere that will be monitored from a single station will be quite large.

This affords opportunity to observe gradients in ionospheric parameters or to locate boundaries of ionospheric regions.

As an example of the coverage that will be available, a map of proposed TISS sites is shown in Figure 1. The cross-hatched area around each station is the equivalent portion of the ionosphere which each TISS will be able to monitor, (350 km alt., 15^{o} elevation). The "hole" in coverage which occurs for each station is due to the 55^{o} inclination of the GPS satellites. Note that the stations closer to the geographic pole have this "coverage hole" almost overhead. Coverage of the North American Continent, especially in the important auroral and polar cap regions, is extensive, and is designed to provide the opportunity for near-real-time measurements of the ionospheric main trough, the scintillation boundary region, and the state of scintillation activity in the polar cap region. Latitudinal coverage is also provided from nearly the magnetic equator to the magnetic pole along the eastern part of the North American Continent.

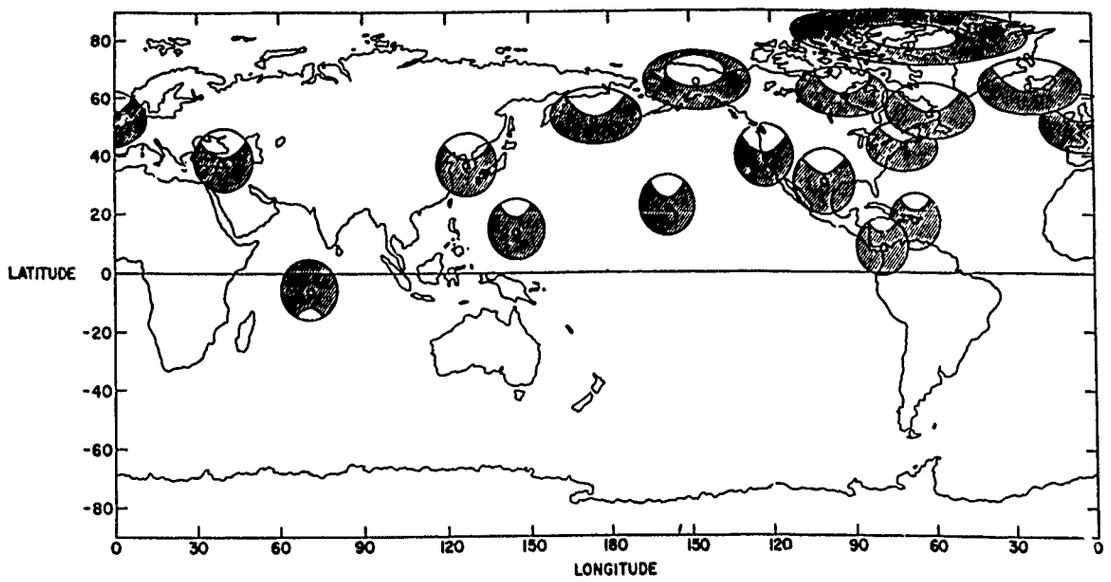


Figure 1. Proposed TISS Geographic Coverage, (after Bishop et al., AGARD 1989)

GPS FULL P-CODE RECEIVERS' MEASUREMENT CAPABILITY

The military user, who has access to the p-code modulation, and who can afford the greater expense of the p-code receivers, will have much greater capability to measure trans-ionospheric parameters. P-code receivers can operate at the high data bandwidths required to characterize ionospheric scintillation. Because of the differential nature of the measurement of ionospheric TEC and phase change, such measurements are not affected by the use of selective availability. Therefore, a non-military user may expect to employ p-code receivers for ionospheric measurement unless anti-spoofing is in effect.

Commercial GPS p-code receivers designed for navigation or geodetic applications may have limitations in their capability to make ionospheric measurements. Internal firmware changes, at least, will likely be required to accumulate and output ionospheric data. Such data may have sampling or bandwidth limitations, or may simply be limited by the capacity of the data port. Multi-channel receivers which operate via time-multiplexing will have a higher system noise floor, and thus a more limited capability to measure amplitude scintillation than receivers with independent channels. All receivers will require a TEC calibration mechanism to determine receiver offset and inter-channel bias, and to counteract drift.

CODE-FREE GPS RECEIVERS FOR IONOSPHERIC MEASUREMENTS

There is a need for an inexpensive receiving system using GPS signals to make measurements of TEC at many locations throughout the world to provide real-time inputs to monthly median ionospheric models as well as for real- time corrections for precise positioning and timing systems. These models are used, for example, in the radio astronomy community for correction to ionospheric differential group path errors in Very Long Baseline Interferometry, (VLBI), in precise time transfer by satellite, and in modern geodetic measurements using GPS radio waves. Single frequency measurements in VLBI, time transfer or geodesy are contaminated with the delays due to the group path of the ionosphere along each leg of the radio signal path. A direct, two-frequency method of determining the ionospheric group delay along the same satellite path, as that of the desired VLBI, time transfer, or geodetic measurement, is obviously the best means of correcting for ionospheric group delay, rather than using a monthly median ionospheric model. With a code-free GPS receiver the capability of making real-time dual frequency measurements of ionospheric group delay, along the same directions as the single-frequency GPS fully coded data, has become a reality.

Currently, GPS may not allow the non-military user to have access to the P code modulation which will allow a direct measure of the ionospheric differential carrier phase advance as well as differential group path delay. However, a code-free, dual frequency receiving system, based upon the ideas of MacDoran, (1984), has been developed by Davis, et. al., (1989) which makes measurements of differential group delay with a minimum time constant of 15 seconds. In order to attain an accuracy approaching one TEC unit, 1×10^{16} el/m^2 column, an integration time of several minutes is used by Davis et al. [1989] in the code-free mode of operation. For another receiver, the 'Rogue', extremely high precision, (much better than one TEC unit) has been documented by Srinivasan et al., [1989], using 5 minute averages. Most code-free receiving systems simultaneously determine the ionospheric group delay for all satellites within view, hence the TEC can be measured in several directions simultaneously. Another codeless GPS receiving system used for improving time comparison results has been described by Imae, et. al., [1988] and Imae, et. al., [1989].

Since the code-free receiver relies on the correlation of two pseudo-random, spread spectrum signals embedded in noise, received at the two GPS frequencies, a relatively long integration time is required to determine the TEC from the differential group delay technique. There are plans to have code-free units commercially available for routine use by groups which would not otherwise be able to correct directly for the effects on the ionosphere on their precise positioning or timing system. Other, partially code-free, receiving systems are also under development, Osborne, [1988].

GPS MEASUREMENTS OF POLAR CAP IONOSPHERIC EFFECTS

MEASUREMENTS OF POLAR CAP TEC AND L-BAND SCINTILLATION

One of the primary objectives of current ionospheric research is improved understanding of the various physical properties controlling the polar cap ionosphere. To this end a single-channel GPS receiver was used to make the first L-band observations of polar ionospheric effects at Thule, Greenland in 1984, [Klobuchar et al., 1985]. These data showed diurnal variation in background absolute TEC, with short-term variations in TEC that more than doubled the background values in less than ten minutes, (Figure 2), and occurrence of 1.2 GHz and 1.6 GHz scintillation generally coincident with these large TEC variations. In addition, coordinated GPS TEC measurements and optical observations of 6300 angstrom airglow emission, established a correlation between the large polar cap TEC variations and drifting F-layer ionization patches, [Weber et al., 1986].

Subsequent to that early campaign effort, nearly continuous observations have been obtained from Thule since October 1987. As will be discussed below, these measurements show increasing occurrence and magnitude of these TEC and scintillation effects, which occur predominately in the October-March period, [Bishop and Klobuchar, AGU, 1989].

MODELING TEC VARIATIONS CAUSED BY F-LAYER IONIZATION PATCHES

TEC variations associated with large ionospheric structures can be dramatically affected by varying the angle or velocity at which the line of sight transits the structure. Attempts to apply TEC data from one such measurement to a system with

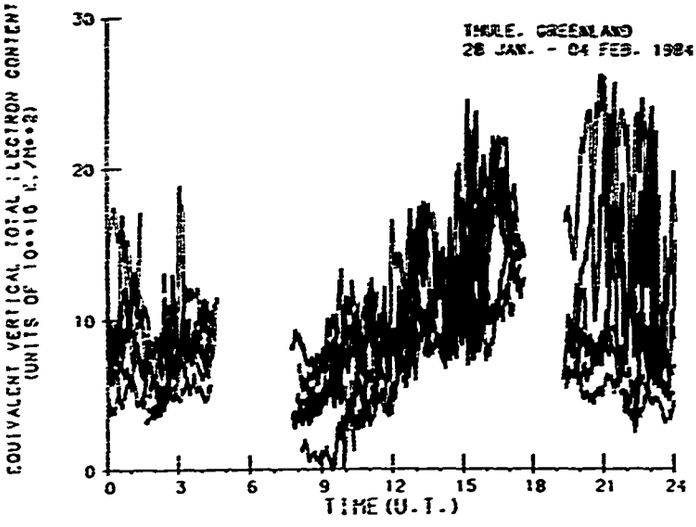


Figure 2a. Diurnal behavior of polar cap equivalent vertical TEC.

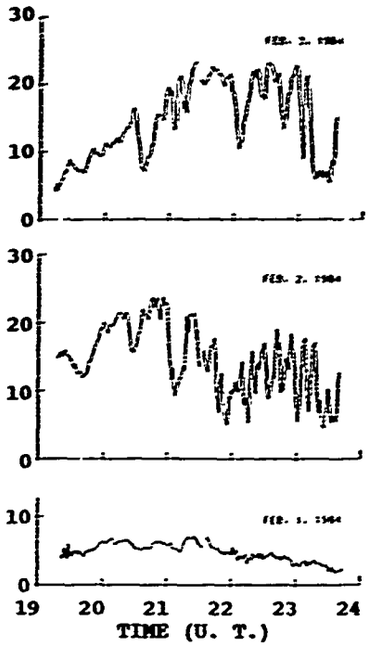


Figure 2b. Dynamic variability of polar cap TEC.

differing geometry must take these geometric effects into account.

The TEC variations described above, (Figure 2), associated with patches of enhanced F region plasma in the polar cap ionosphere were generally dominated by the drift velocity of the patches. Most of these GPS observations were taken at or above 45^{o}, where the line of sight moves at below 100 m/s. Weber et al., [1986], have noted that typical patch drift velocities are up to 1000 m/s. Were these same patches to be observed using satellites at 1000 km altitude, for example, the variations would be dominated by the velocity of the ionospheric penetration point (IPP) of the line of sight and would not show the same temporal TEC gradients. If the same observations were to be made at low elevation angles, such as below 20 degrees, the IPP velocity would be further increased and the TEC values would result from a significantly different set of integration paths through the structures.

Most measurements of ionospheric TEC variations have been made using signals from geostationary or high altitude satellites observed at high elevation angles. The short term temporal structure and gradients of TEC observed in these cases are generated primarily by ionization drift. However, TEC variations of concern to systems such as trans-ionospheric radars are often those dominated by IPP velocity, frequently at low elevation angles. For example, in many observing situations the space vehicles are at relatively low altitudes, such as 1000 km, and the IPP appears to move relatively fast. An object may remain visible for only 5 to 15 minutes in such a case. The rapid motion of the IPP is the primary source of changes in ionospheric effects over an observation window.

The variation in apparent TEC and TEC gradients with observation system reference frame may be examined using an analytical model, [Bishop and Klobuchar, URSI 1989], of a typical ionospheric "patch". Figure 3a shows such a typical patch, [after Weber et al., 1984], and Figure 3b shows the analytical reconstruction used to study apparent TEC variation. The analytical technique models the density profiles shown and integrates electron content along any desired raypath. The code allows a raypath to be established at any desired elevation angle and then plots the observed TEC as the patch is "drifted" past at any desired velocity. For contrast, the patch may be centered at any desired latitude relative to the observer and the raypath elevation may be varied to correspond to tracking an object at any given object altitude.

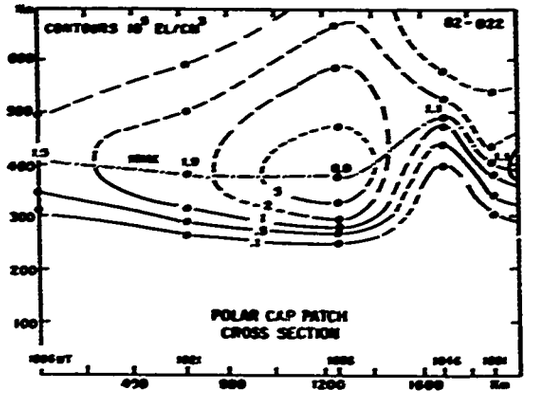


Figure 3a. Measured vertical profile of polar cap patch, drift velocity 700m/s, observed at Thule, Greenland, by Weber et al., [1984].

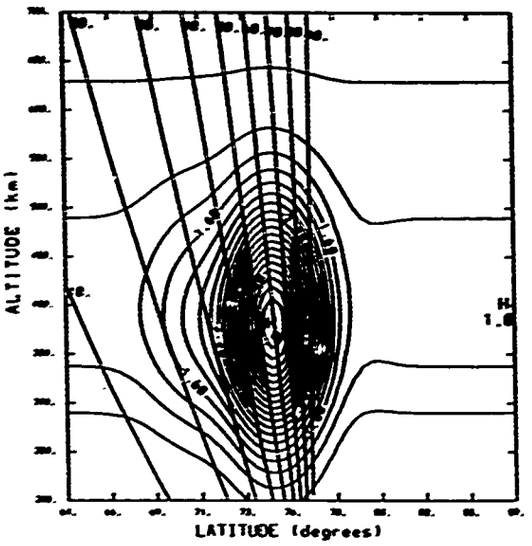


Figure 3b. Analytical reconstruction of patch shown in Fig. 3a, (with lines of sight at 10° intervals), used in simulation of TEC variation.

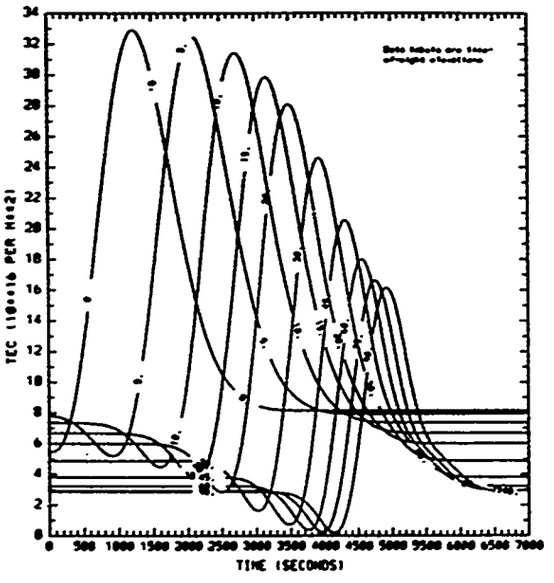


Figure 4a. Modeled total electron content profiles of constant elevation angle views of large scale polar cap ionization patch drifting at 600 m/s.

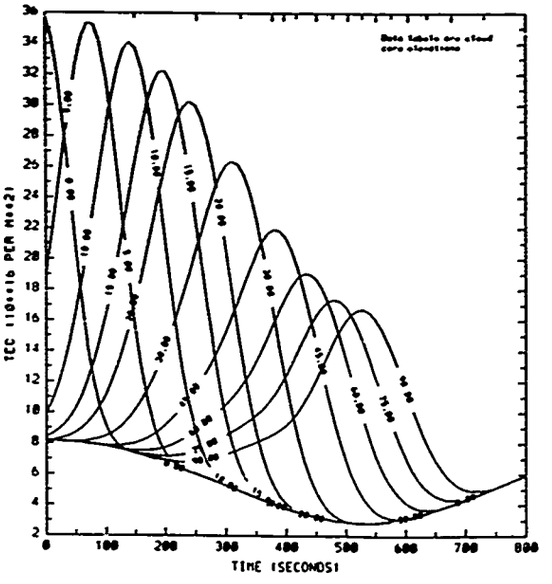


Figure 4b. Modeled TEC profiles, as from signals of a 1000 km altitude satellite, of a large scale polar cap ionization patch, located at fixed elevations with respect to the observer.

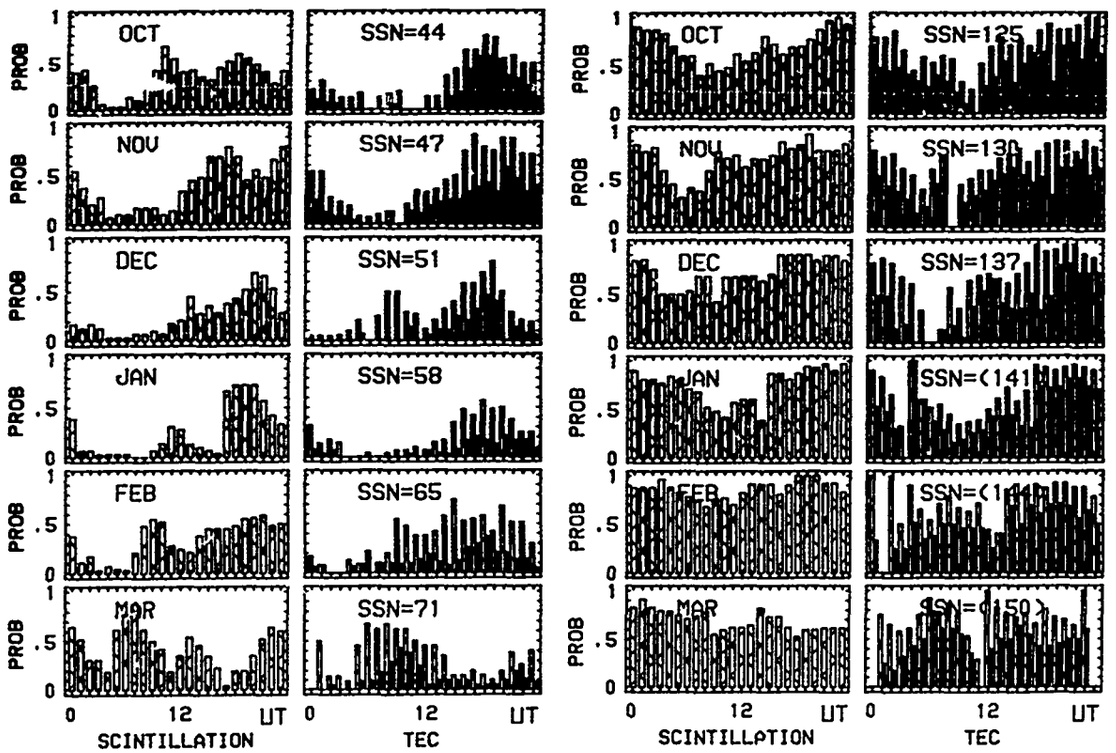
Figures 4a and 4b illustrate the large variability in TEC variation profile that a single patch structure may produce. In Figure 4a the analytical model was set to fix the look direction at a constant elevation as in observation of a geostationary or high altitude satellite. The figure gives a family of TEC profiles for the patch, drifted past at 600 m/s, and observed at 0, 5, 10, 15, 20, 30, 45, 60, 75, and 90 degrees elevation. In Figure 4b the look direction was moved in elevation angle

through the patch, as in observation of a satellite at 1000 km altitude. The family of TEC curves here shows the modeled TEC profiles for a set of 160 degrees scans from horizon through zenith, with the centroid of the patch located at latitudes corresponding to elevations of 0, 5, 10, 15, 20, 30, 45, 60, 75, and 90 degrees, as seen by an observer at 76.5 degrees (Thule) latitude.

POLAR TEC VARIATION LINKAGE WITH AMPLITUDE SCINTILLATION

Comparison of TEC and L-band amplitude scintillation measurements, both obtained using GPS signals during the January-February 1984 Thule campaign, showed a strong relation between TEC variations and scintillation enhancement. Other observations from that campaign linked the TEC variations to occurrence of significant 250 MHz amplitude scintillation observed using polar beacon satellite signals, [Weber et al., 1986]. Earlier campaign observations had linked similar diurnal occurrence of the patches with foF2 enhancements and with increases in 250 MHz amplitude scintillation [Buchau et al., 1985; Weber et al., 1984].

Nearly continuous GPS observations have now been obtained from Thule since October, 1987. These new observations provide the first substantial data base characterizing the diurnal occurrence and size of polar cap TEC variations, [Bishop et al., SPI Conference Proceedings, 1989]. These data confirm the effects observed during the short 1984 campaign, and show an increase in their occurrence and magnitude with increasing solar activity. A quick-look activity assessment was made of all the



Scintillation: Activity approximating >= 5 dB 15 - minute fade depth index

TEC activity:
crosshatched: >= structures of 1 TEC unit amplitude, 5 minute period
solid: >= gradients 1 TEC unit/5 minutes, typical 10 minute period

Figure 5. Occurrence of polar cap TEC disturbances and amplitude scintillation at 250 MHz, comparing the 1987-88 and 1988-89 winter seasons at Thule, Greenland.

chart recordings of this two-year, L-band and 250 Mhz, data set. Figure 5 charts the probability of occurrence of activity, on an hourly basis, for the winter months of this data set. The 250 Mhz scintillation activity level in Figure 5 corresponds to activity exceeding the 5 dB - 15 minute fade depth index used by Basu et al., [1987]. The two TEC activity levels charted in Figure 5 represent, respectively, small disturbances - on the order of $1x10^{16}$ electrons/m^2 (one TEC unit) over periods less than 5 minutes, and larger disturbances (solid bars) with gradients exceeding 1 TEC unit/ 5 minutes and typical periods of ten minutes.

In the 1987-88 data from Figure 5, a clear diurnal pattern, confirming earlier observation, may be seen. The March 1988 data show a shift of disturbances to an earlier UT period which may be transitional between earlier patterns and the more uniform diurnal occurrence of scintillations reported by Buchau et al. [1985], to prevail for sunspot activity above the 80 to 100 level. The 1988-89 data clearly confirm this uniform occurrence, except in the TEC data where two gaps occur, which are due to lack of available data.

The diurnal patterns in the 1987-88 data show clear linkage between the occurrence of TEC variations and 250 MHz scintillation in the Thule region. Differences between the patterns for a given month may be attributable to the limitations in available GPS data, and the fact that the look directions to the GPS satellites and the polar beacon satellites only occasionally coincide. As multiple-direction GPS measurements (Figure 6, below), show, there can be high activity observable in one direction from Thule, while little or no activity is visible in another. The apparent linkage between occurrence of significant TEC variation and 250 MHz amplitude scintillation presented here suggests that it may be reasonable to infer the monthly occurrence of polar cap TEC disturbances based on the occurrence of the scintillation, since this is now well-documented over a solar cycle, [Basu et al, 1987].

EXTREME VARIABILITY OF POLAR CAP TOTAL ELECTRON CONTENT

One of the advantages of using GPS to measure the ionosphere is that multiple satellites will be available simultaneously. Four - direction GPS observations, obtained at Thule in October 1988, show that during TEC disturbances there is dramatic spatial variation in the magnitude, period and occurrence of the TEC variations. An example of this, from Clynch et al., [1989], is shown in Figure 6a. The ionospheric locations of the lines-of-sight for these four satellites are shown in Figure 6b. Three of these lines-of-sight were fairly close and yet saw significantly different detail in the TEC variations. The fourth started from over 1000 km west-southwest of Thule, and only encountered the irregularities as it approached Thule at the end of the hour. It is interesting to note that PRN 8 appears to exit the region of irregularities, and enter a region of high background TEC, just east of Thule. This example illustrates the high degree of spatial and short-term temporal variability in the polar cap ionosphere. These data emphasize not only the difficulty of using models in the high latitudes, but also the difficulty of using reference data from "close" lines-of-sight in these regions.

Figure 7 gives three monthly comparisons of TEC activity levels between the 1987-88 and 1988-89 winter seasons at Thule. For each month, 5 examples have been selected to illustrate the largest "peak to valley" TEC changes occurring within a half hour or less. It may be noted that the magnitude of the TEC variations increased significantly between these two years - an increase by at least a factor of three. (Also note that the plots are relative, slant TEC, referenced to zero for ease of examination. Azimuth and elevation of the line-of-sight is indicated by symbols "X" and "+", respectively.) Since 1988-89 is at least a year before the projected peak of the current solar cycle, greater TEC activity may yet be observed in 1989-90.

SUMMARY

Signals from the GPS satellites provide the means to continuously measure TEC and scintillation, in multiple directions, from anywhere on earth. Because of this multi-directional capability, the geographic region of the ionosphere that will be monitored from a single station will be quite large, and it is possible to monitor spatial gradients in ionospheric parameters over distances up to 2000 km, and to locate boundaries of ionospheric regions. Such information, provided in real-time, will be a substantial improvement over point references for models of ionospheric parameters. The Trans-Ionospheric Sensing System will use GPS signals in this fashion to provide the AWS Space Forecast Center with real-time reporting of TEC and scintillation, in multiple look directions, from stations throughout the world. Economical, code-free

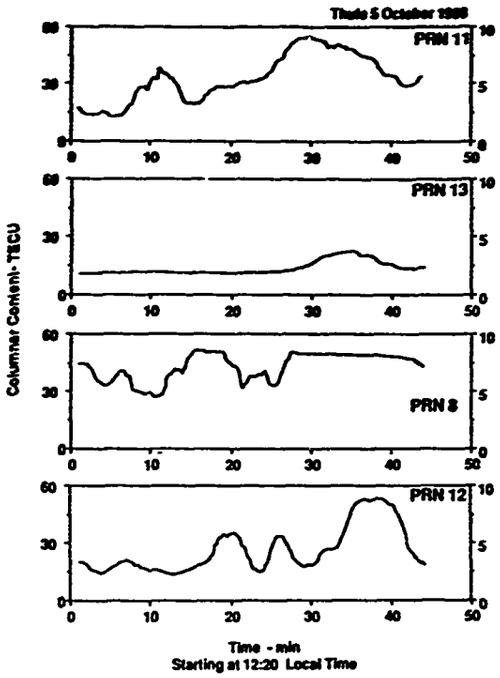


Figure 6a. Vertical absolute TEC from four different GPS SV's

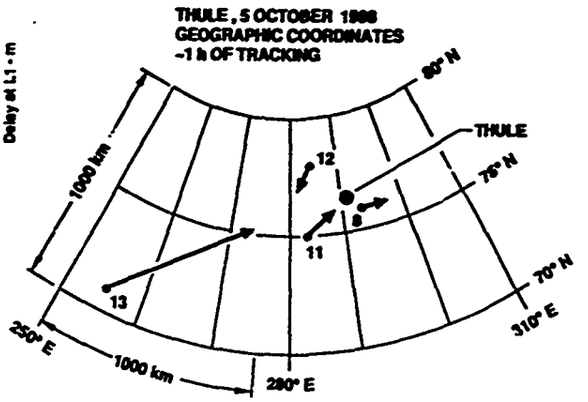
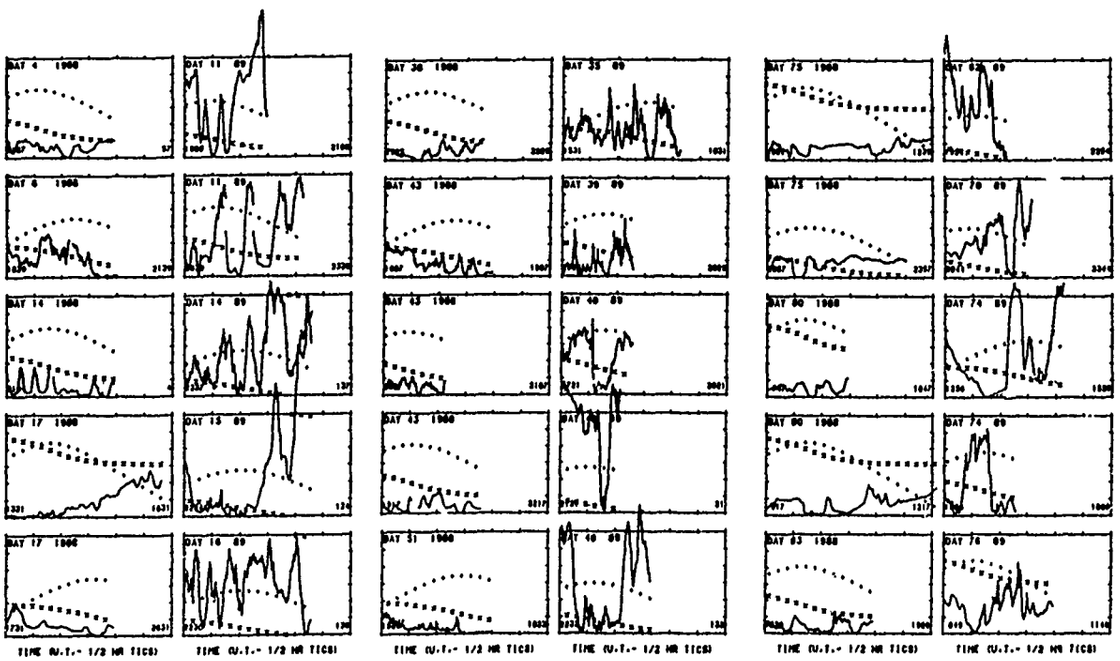


Figure 6b. Ionospheric intercept tracks



ORDINATE: Relative slant TEC, 0-60; Az. deg. (X), 0-360; El. deg. (+), 0-90

Figure 7. Comparison of magnitude of more significant TEC variations observed at Thule, Greenland during January - March 1988 and 1989.

GPS receivers will enable non-military users to measure TEC regardless of GPS operating mode. Several years of measurements in the polar cap have demonstrated the value of GPS' multi-directional, multi-parameter capability. These measurements have linked polar cap TEC disturbances with scintillation occurrence, and have shown significant spatial, diurnal, monthly and solar cycle variability of TEC that set it apart from the relatively smooth spatial gradients and temporal cycles typically found in midlatitudes. It is noted that TEC variations associated with large ionospheric structures can be dramatically affected by varying the angle or velocity at which the line of sight transits the structure. Attempts to correlate TEC measurements from satellites at greatly differing altitudes, or to apply TEC data from one such measurement to a system with differing geometry must take these geometric effects into account.

ACKNOWLEDGEMENT

The authors wish to thank the Danish Commission for Scientific Research in Greenland for permission to conduct ground experiments at Thule AB.

REFERENCES

Basu, Santimay, E. MacKenzie, and S. Basu, Ionospheric constraints on UHF/VHF communications links during solar maximum and minimum periods, in The Effect of the Ionosphere on Communication, Navigation and Surveillance Systems, J. M. Goodman, ed., Ionospheric Effects Symposium, 569, 1987.

Bishop, G. J., J. A. Klobuchar, and D. W. Blood, Polar Cap Total Electron Content Variations: Linkage to 250 Mhz Scintillations, Physics of Space Plasmas (1988), T. S. Chang, ed., 1988 Cambridge Workshop in Geoplasma Physics on Polar Cap Dynamics and High Latitude Ionospheric Turbulence, Vol 8, SPI Conference Proceedings, Scientific Publishers Inc., 1989.

Bishop, G. J., J. A. Klobuchar, Lt Col A. E. Ronn, and Capt M. G. Bedard, A Modern Trans-Ionospheric Propagation Sensing System, Operational Decision Aids for Exploiting or Mitigating Electromagnetic Propagation Effects, NATO AGARD Electromagnetic Wave Propagation Panel Symposium, San Diego, 15-19 May 1989, paper 31, AGARD Conference Proceedings no. 453.

Bishop, G. J., and J. A. Klobuchar, Polar Cap Observations of Ionospheric Effects at L-Band During March 1989, presented at 1989 Fall Meeting of the American Geophysical Union, 4-8 December 1989.

Bishop, G. J., and J. A. Klobuchar, Modeling Total Electron Content Variations Associated with Polar Cap F Layer Patches at Low Elevation Angles, presented at the National Radio Science Meeting of the International Union of Radio Science (URSI), Boulder, CO, 4-6 January, 1989.

Buchau, J., E. J. Weber, D. N. Anderson, H. C. Carlson, J. G. Moore, B. W. Reinisch, and R. C. Livingston, Ionospheric Structures in the Polar Cap: Their origin and relation to 250-Mhz scintillation, Radio Sci., 20 325 (1985).

Clynch, J. R., D. S. Coco, C. coker, and G. J. Bishop, A Versatile GPS Ionospheric Monitor: High Latitude Measurements of TEC and Scintillation, in Proceedings of ION GPS-89, The Institute of Navigation, Washington, DC, 445, September 1989.

Davis, D., M. Weiss, M. Vidmar, A Codeless Ionospheric Calibrator for Time Transfer Applications, in Proceedings of ION GPS-89, The Institute of Navigation, Washington, DC, 455, September 1989.

Imae, M, W. Lewandowski, C. Thomas, and C. Miki, A Dual Frequency GPS Receiver Measuring Ionospheric Effects Without Code Demodulation and Its Application to Time Comparisons, Proceedings of 20th Annual Precise Time and Time Interval (PTTI) Applications and Planning Meeting, Dec 1988.

Imae, M., C. Miki, and C. Thomas, Improvement of Time Comparison Results by using GPS Dual Frequency Codeless Receivers Measuring Ionospheric Delay, Proc. 21st Annual Precise Time and Time Interval (PTTI) Applications and Planning Meeting, Dec 1989.

Klobuchar, J. A., G. J. Bishop, P. H. Doherty, Total Electron Content and L-Band

Amplitude and Phase Scintillation Measurements in the Polar Cap Ionosphere, Propagation Effects on Military Systems in the High Latitude Region, NATO AGARD Electromagnetic Wave Propagation Panel Symposium, Fairbanks, 3-7 June, 1985, paper 2.2, AGARD Conference Proceedings no. 382.

Jursa, A.S., ed., Handbook of Geophysics and the Space Environment, available from NTIS, 5285 Port Royal Road, Springfield VA 22161, 1985.

MacDoran, P. F. , ISTAC, Inc. press release, 13 January 1984.

Osborne, A., private communication, 1988.

Runkle, M. T., and T. M. Larkin, GPS Program Status, in Proceedings of ION GPS-89, The Institute of Navigation, Washington, DC, September 1989.

Srinivasan, J. M., T. K. Meehan, and L. E. Young, Code and Codeless Ionospheric Measurements with NASA's Rogue GPS Receiver, in Proceedings of ION GPS-89, The Institute of Navigation, Washington, DC, September 1989.

Weber, E. J., J. A. Klobuchar, J. Buchau, H. C. Carlson, R. C. Livingston, O. de la Beaujardiere, M. McCready, J. G. Moore and G. J. Bishop, Polar Cap F-layer Patches: Structure and Dynamics, J. Geophys. Res., 91, 12, 121, 1986. [AFGL-TR-86-0074 (31 Mar 86) ADA 175241]

Weber, E.J., J. Buchau, J.G. Moore, J.R. Sharber, R.C. Livingston, J.D. Winningham, and B.W. Reinisch, F-layer Ionization Patches in the Polar Cap, J. Geophys. Res., 89: 1683 (1984).

AD-P006 317

IONOSPHERIC DELAY MEASUREMENTS USING GPS SATELLITE SIGNALS

J.D. Holmes and R.L. Meyer
Texas Instruments Incorporated
P.O. Box 869305 M/S 8443
Plano, Texas 75086

Abstract

The ionospheric group delay associated with GPS satellite signals is usually compensated by use of a wavelength squared formula. Using this relationship the typical GPS receiver minimizes the ionospheric effects by tracking both the L1 and L2 signals and correcting the navigation solution with the two frequency measurements independent of the ionosphere or carrier frequency. Normally, these compensations are performed in the navigation software and the ionospheric group delay values are not specifically characterized or accessible to the user. This level of functionality is sufficient if the operator desires only navigation data. However, the ionospheric delay data may be useful for other purposes and could be provided to the user if the receiver and navigation software were appropriately designed. A receiver employing distinct tracking loops for Range and TEC dynamic estimation can directly supply the navigation function with ionosphere corrected ranges and interested observers with the ionospheric delay parameters directly from the L-band tracking process. Further, this receiver can track the time varying nature of the ionosphere and provide continuously corrected ranges to the navigation function in times of highly distressed disturbances. A GPS receiver manufactured by Texas Instruments for the Space Systems Division of the Air Force Systems Command (on contract F04701-83-C-0076) has this capability. This receiver is capable of tracking L1 and L2 simultaneously on six satellites and, as a byproduct of its design, outputs L1 excess delay and delay rate directly. This receiver has been used to obtain L1 excess delay data spanning several months during 1989 through measurements from L1 and L2 on GPS satellites. This data, taken at a Texas Instruments facility in Plano Texas, is presented.

Introduction

The excess ionospheric delay measurements presented were made at Plano, Texas (33.0656 N, 96.6948 W) between July 1989 and February 1990. Satellites of opportunity were tracked at mid-afternoon local time to sample the delay at a point near the diurnal maximum. The data from November and February were collected to sample early morning values. Martin (1) provided the following equation relating the L1 excess delay in meters to the TEC:

delta L = [-b/(4 x pi**2 x f**2)] * I * [csc(E**2 + 20.3 **2)]**1/2

where L = excess delay in meters
b = .6 x 10**3 = constant in MKS unit
f = carrier frequency in hertz (at L1 = 1575 MHz)
I = vertical electron content in electron/meter**2
E = elevation angle in degrees

The 50 meters full scale on the plots of L1 excess delay corresponds to a TEC of approximately 3 x 10**18 electrons/meter**2 along the line of sight to the GPS satellite.

The maximum in the 11 year sunspot cycle has not arrived, yet the data taken coincide with a high level of solar activity. The delay measurements, therefore, reflect near maximum expected values. The ionospheric characteristics observed are considered to be those prevalent along the line-of-sight to each SV. Satellite azimuth and elevation from Plano are provided to allow easy co-ordination with SV position. Data was taken at one minute intervals and fitted to a fourth order polynomial. The coefficients of these polynomials are supplied along with the mean time and end points of validity. A sample set of the data along with the curve fit for each SV are provided for September 14 and 15, 1989.

91 9 4 127

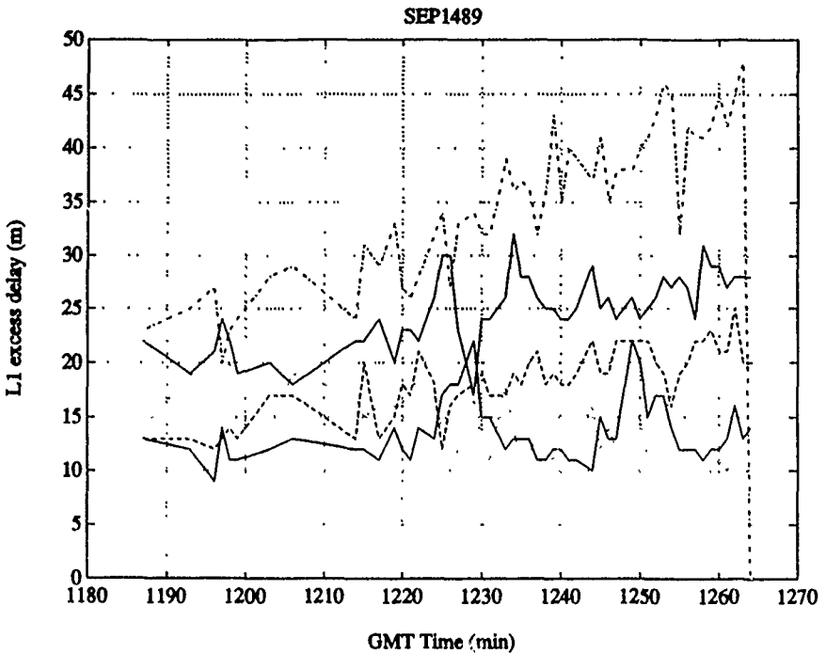
91-09685

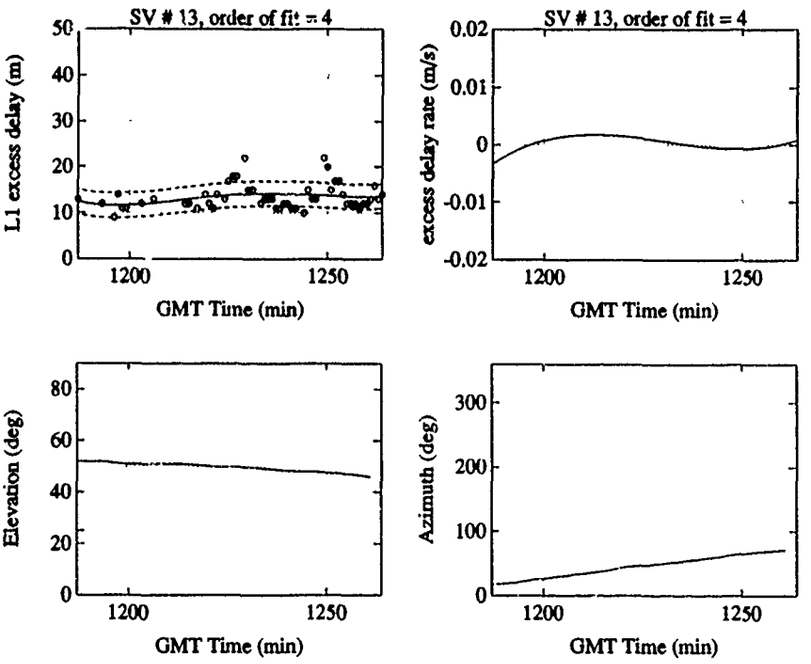
COEFFICIENTS of fourth order polynomial fit for L1 excess delay due to TEC effects (time units in hours)

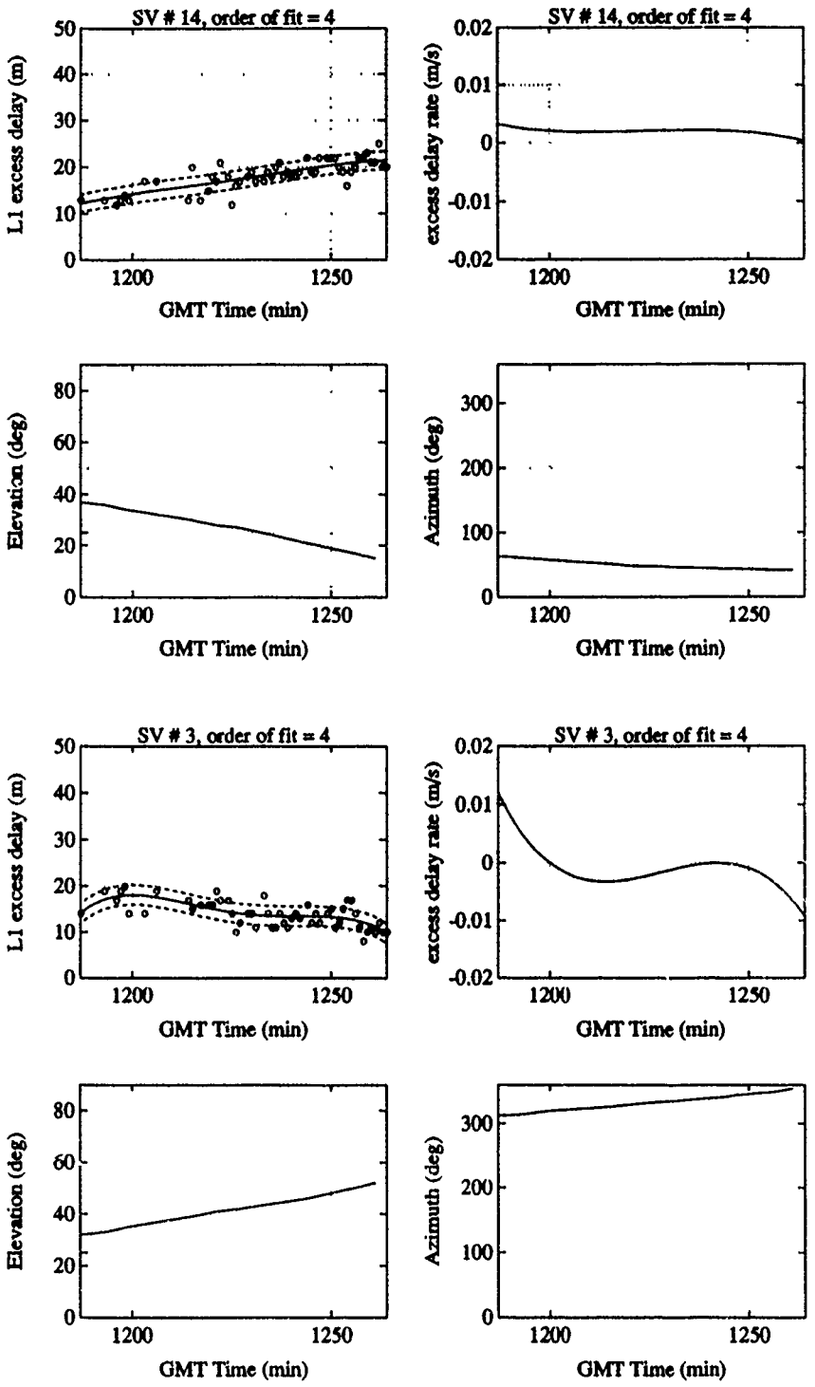
SV#	x^4	x^3	x^2	x^1	x^0	Mean Time	Std Dev	Start time	End time
COEFFICIENTS for JUL1489									
SV# 6:	-152.316186	-41.912354	43.540262	21.875439	20.443701	22.483333	2.50	22.0	23.0
SV# 9:	80.729551	-0.562365	-11.531875	3.379847	8.063147	22.508333	1.52	22.0	23.0
SV#11:	-77.654478	61.296571	17.650953	-11.920607	7.129722	22.508333	1.83	22.0	23.0
SV#12:	-219.286814	11.750669	48.049272	-4.609511	7.736251	22.508333	1.63	22.0	23.0
SV#13:	24.974406	37.747518	-7.018992	-8.014483	8.370745	22.508333	1.36	22.0	23.0
SV#14:	38.978805	6.807033	-2.909719	-8.516000	11.724391	22.508333	1.88	22.0	23.0
COEFFICIENTS for JUL2089									
SV# 6:	-7.008525	5.365286	4.267901	-0.626488	6.282076	19.975000	0.44	19.5	20.5
SV# 9:	14.184122	-2.001992	0.949526	-4.760955	8.524017	19.975000	0.51	19.5	20.5
SV#11:	-4.391403	2.415337	-1.227946	-4.639246	8.736526	19.975000	0.45	19.5	20.5
SV#13:	-52.304936	-39.879317	12.098305	-9.118807	18.677418	20.000000	2.77	19.5	20.5
COEFFICIENTS for AUG0789									
SV# 6:	10354.612155	343.218286	-366.522116	19.093512	27.948986	21.266667	1.93	21.1	21.5
SV# 9:	3165.378894	334.026314	-100.324500	-8.018098	13.829751	21.283333	2.03	21.1	21.5
SV#11:	-3199.677550	-196.116633	181.057287	7.665200	8.963873	21.283333	2.02	21.1	21.5
SV#12:	-1291.291477	182.345909	17.239097	-11.585944	11.082358	21.283333	1.38	21.1	21.5
SV#13:	4837.660961	-439.214532	-194.218062	19.455177	10.411933	21.283333	1.01	21.1	21.5
COEFFICIENTS for SEP1189									
SV# 3:	1.093358	0.321554	-0.741252	-3.260794	12.357399	20.700000	2.16	19.5	21.9
SV# 9:	-26.648471	-93.670745	27.837721	36.690408	29.799025	19.983333	3.12	19.5	20.5
SV#11:	-0.930590	-1.215413	3.266246	9.431891	22.402868	20.525000	1.96	19.5	21.6
SV#12:	-6.532398	-3.347672	7.088533	16.837703	30.751592	20.466667	3.01	19.5	21.4
SV#13:	0.053391	-0.668835	1.777244	1.787224	11.852594	20.700000	1.85	19.5	21.9
SV#14:	0.974728	1.523919	-2.174570	1.705431	15.042168	20.691667	1.86	19.5	21.9
COEFFICIENTS for SEP1489									
SV# 3:	-58.535449	9.880237	13.402405	-7.494891	14.220069	20.425000	2.13	19.8	21.1
SV#11:	42.808471	-10.671766	-14.449179	9.217845	24.480652	20.425000	2.43	19.8	21.1
SV#12:	8.336324	-10.989530	2.344328	20.677758	31.649949	20.416667	2.97	19.8	21.1
SV#13:	19.139103	-6.735333	-9.81786	3.667478	13.863905	20.425000	2.69	19.8	21.1
SV#14:	-7.328592	-0.835391	1.950481	7.496180	17.224210	20.425000	1.92	19.8	21.1

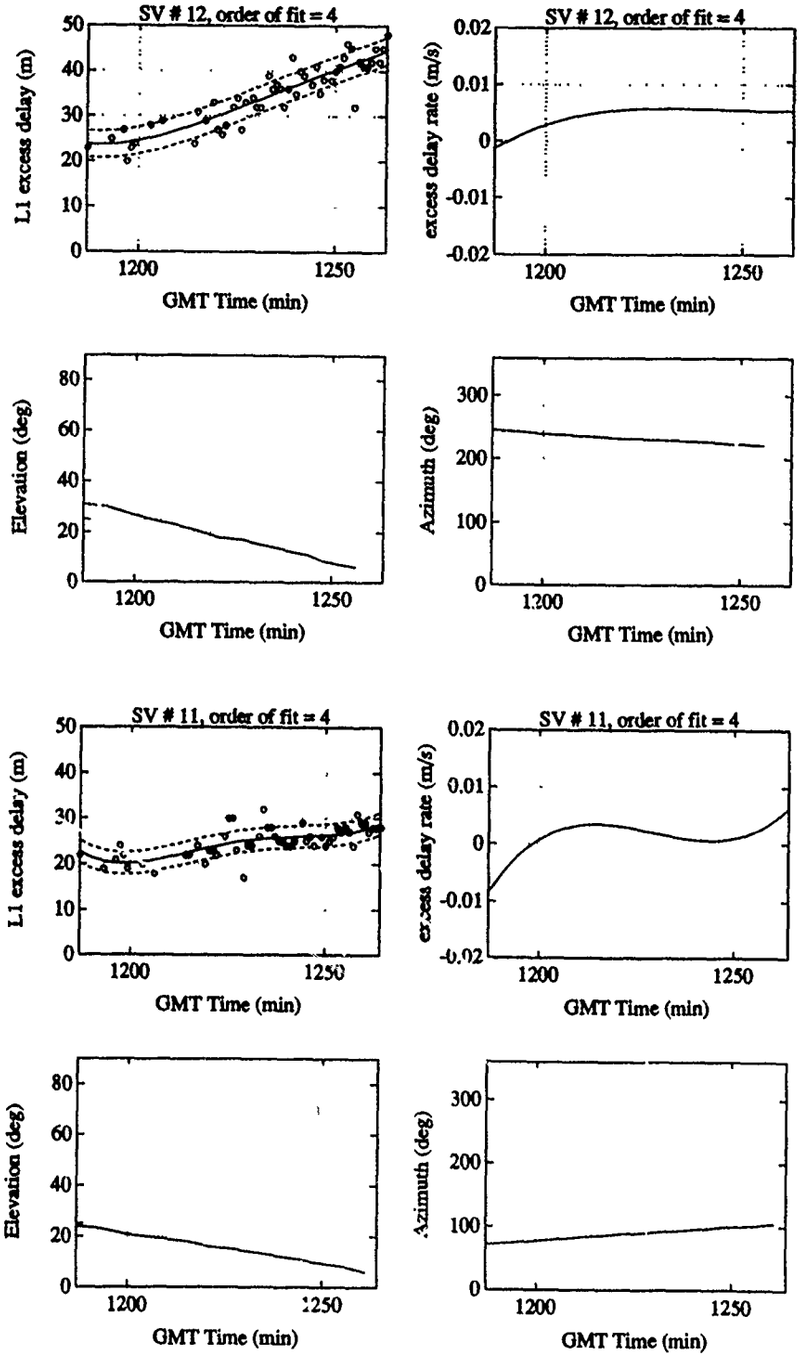
COEFFICIENTS of fourth order polynomial fit for L1 excess delay due to TEC effects (time units in hours)

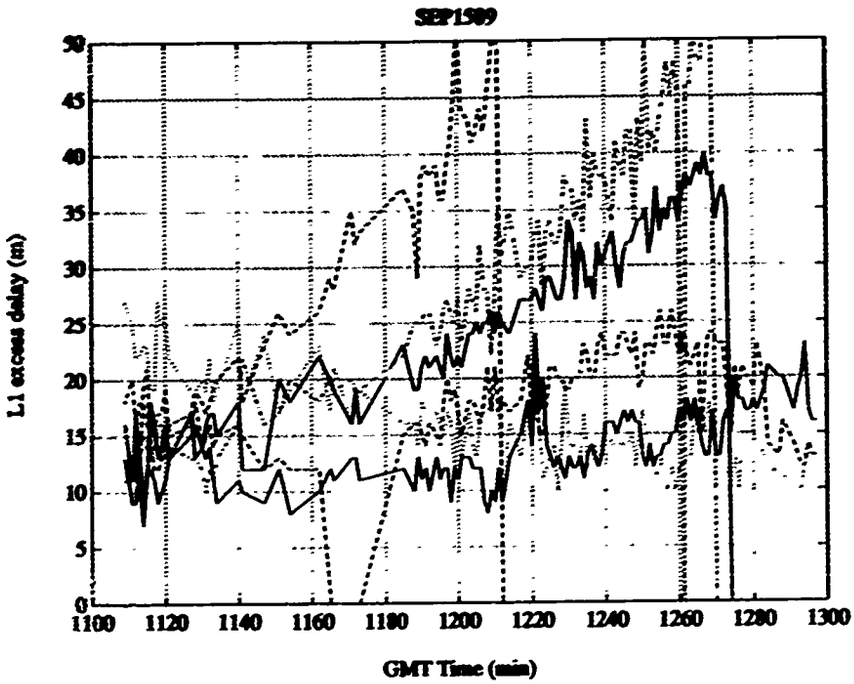
SV#	x^4	x^3	x^2	x^1	x^0	Mean Time	Std. Dev	Start time	End time
COEFFICIENTS for SEP1589									
SV# 3:	-0.190670	-0.155725	1.340540	-3.331461	15.521571	20.050000	2.18	18.5	21.6
SV# 9:	11.201729	3.349786	-2.845377	18.331062	27.397640	19.333333	2.77	18.5	20.2
SV#11:	-1.974616	0.482141	5.359402	8.743881	21.222661	19.850000	1.94	18.5	21.2
SV#12:	1.210371	0.713023	2.385382	13.120698	24.044123	19.816667	2.48	18.5	21.2
SV#13:	0.124587	-0.173801	1.254136	2.466686	12.073726	20.050000	2.32	18.5	21.6
SV#14:	-1.538372	-4.292288	2.843046	8.336606	16.886577	20.050000	1.87	18.5	21.6
COEFFICIENTS for OCT0689									
SV# 3:	11.074209	-2.028535	-4.477463	-2.008117	18.092017	19.525000	1.61	18.8	20.2
SV#12:	184.141518	79.413728	-19.382910	9.229199	40.988232	19.283333	1.70	18.8	19.7
SV#13:	-44.100293	6.090459	21.104636	1.734755	15.401203	19.525000	2.45	18.8	20.2
SV#14:	15.020479	3.336566	-7.583172	9.999513	29.296615	19.516667	1.96	18.8	20.2
COEFFICIENTS for OCT1189									
SV# 3:	29.348652	91.827712	-0.553850	-21.047199	12.681842	19.908333	1.66	19.4	20.4
SV#13:	-205.445280	-57.977916	44.951507	21.465930	21.766759	19.908333	2.05	19.4	20.4
SV#14:	-547.960800	259.888076	47.437558	7.976981	29.985133	19.608333	1.75	19.4	19.8
COEFFICIENTS for NOV1589									
SV# 2:	-3.857631	8.217452	8.769335	3.251534	5.031582	11.691667	1.84	11.0	12.4
SV# 6:	-1.704601	1.091662	0.746561	-0.574626	3.176724	11.833333	1.37	10.8	12.9
SV# 9:	-3.130573	-0.928464	2.819764	0.319082	2.053309	11.925000	1.53	11.0	12.9
SV#11:	-2.950154	-1.596976	3.872516	0.873094	2.458192	11.883333	1.70	10.8	13.0
SV#12:	193.540664	28.489754	-34.624398	-3.961845	5.630631	12.616667	1.44	12.3	13.0
SV#13:	-25.195034	8.286065	12.340051	-1.122827	2.869276	12.291667	1.33	11.6	13.0
COEFFICIENTS for FEB0790									
SV# 3:	-49090.909091	3328.671329	858.566434	-0.557498	22.867188	15.091667	1.73	15.0	15.2
SV# 6:	-28.750635	5.610689	29.018451	-5.668976	6.486678	15.800000	1.77	15.0	16.6
SV# 9:	-0.874582	0.429799	3.065970	1.989405	11.954319	16.500000	2.75	15.0	18.0
SV#12:	0.015908	-0.161459	0.711978	-2.424616	16.774866	16.883333	2.53	15.7	18.1
SV#16:	-2.855188	0.332479	8.286224	8.350046	10.909810	16.266667	2.38	15.0	17.5
SV#17:	-0.445185	0.467166	1.207764	1.446180	6.358201	16.541667	1.80	15.0	18.1
SV#19:	-1.576651	1.553871	1.868383	-5.868934	15.033059	17.050000	1.77	16.0	18.1

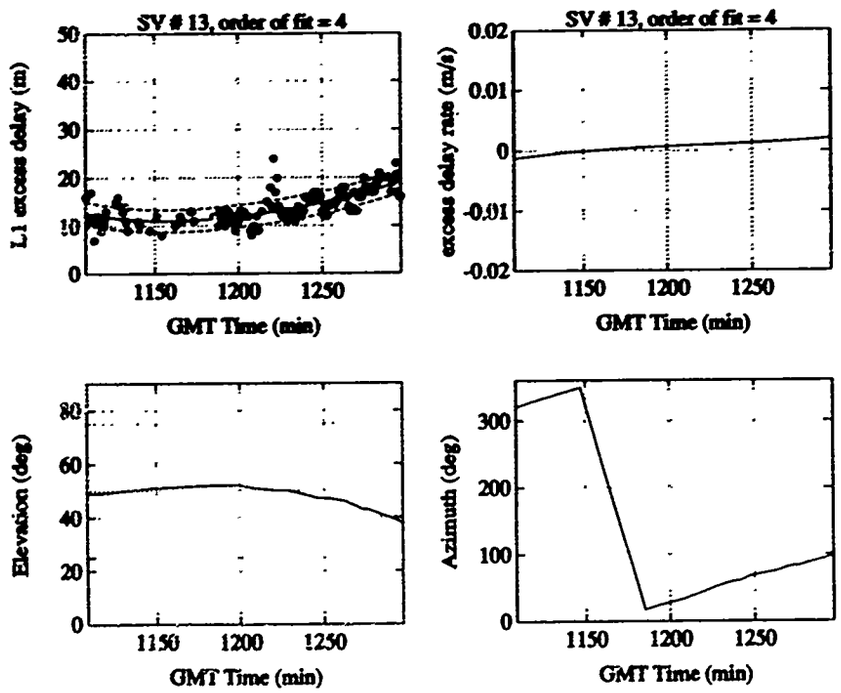

SEP1489
L1 excess delay (m)
GMT Time (min)

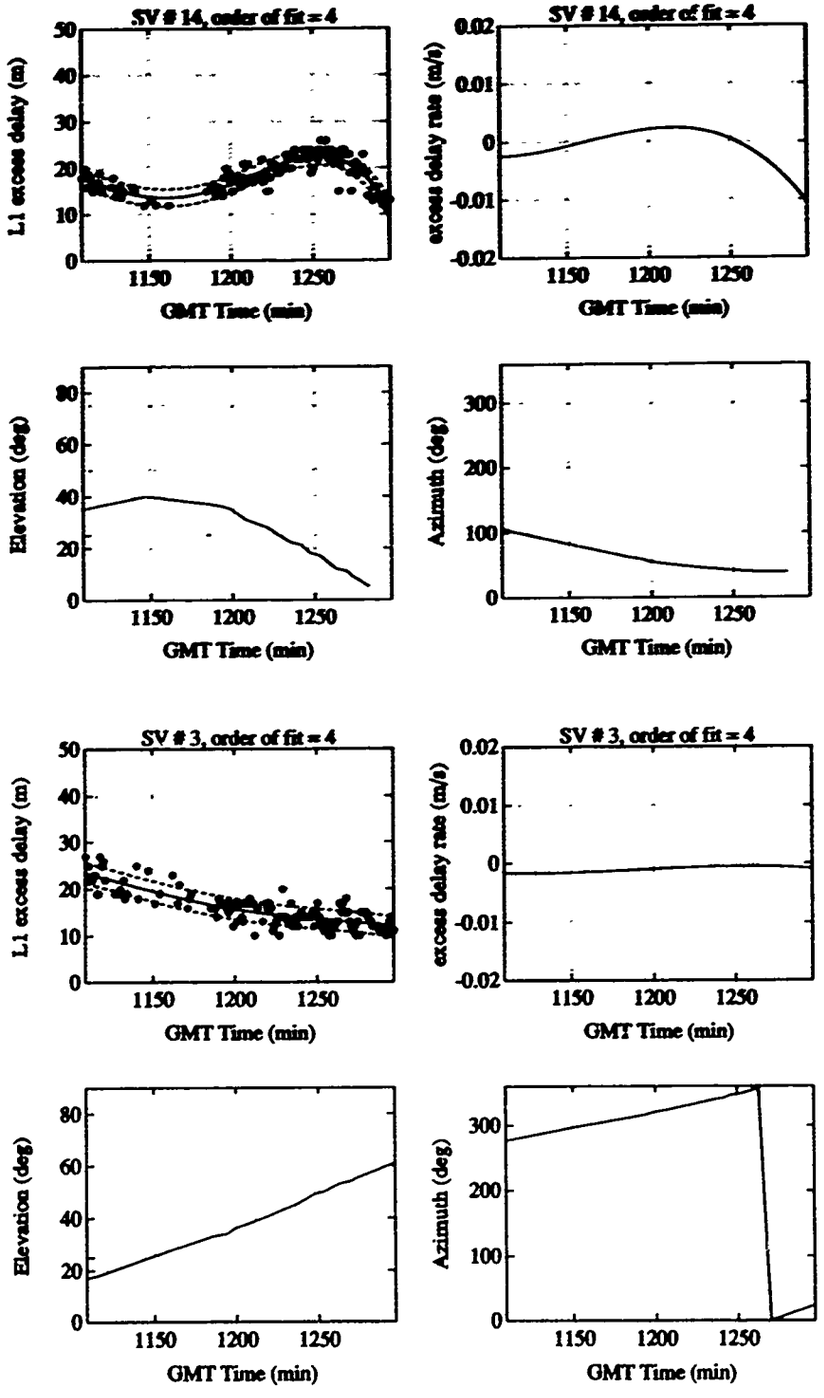

SV # 13, order of fit = 4
L1 excess delay (m)
GMT Time (min)
SV # 13, order of fit = 4
excess delay rate (m/s)
GMT Time (min)
Elevation (deg)
GMT Time (min)
Azimuth (deg)
GMT Time (min)

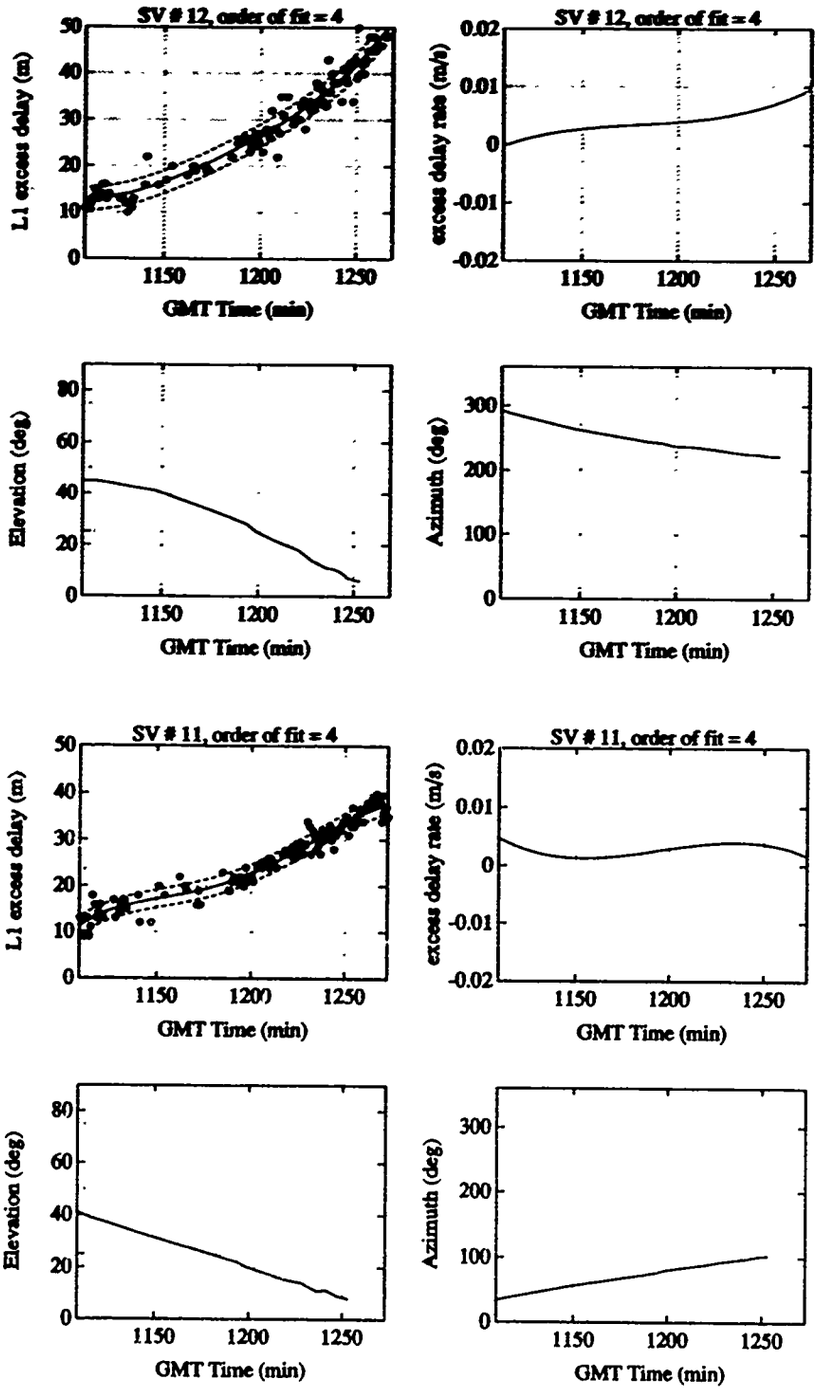

SV # 14, order of fit = 4
L1 excess delay (m)
GMT Time (min)
SV # 14, order of fit = 4
excess delay rate (m/s)
GMT Time (min)
Elevation (deg)
GMT Time (min)
Azimuth (deg)
GMT Time (min)
SV # 3, order of fit = 4
L1 excess delay (m)
GMT Time (min)
SV # 3, order of fit = 4
excess delay rate (m/s)
GMT Time (min)
Elevation (deg)
GMT Time (min)
Azimuth (deg)
GMT Time (min)

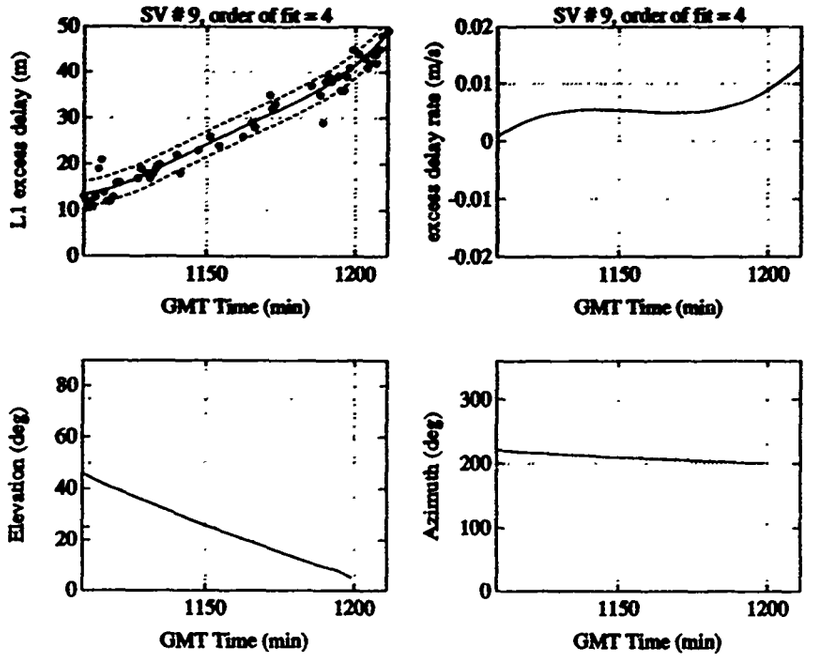

SV # 12, order of fit = 4
L1 excess delay (m)
GMT Time (min)
SV # 12, order of fit = 4
excess delay rate (m/s)
GMT Time (min)
Elevation (deg)
GMT Time (min)
Azimuth (deg)
GMT Time (min)
SV # 11, order of fit = 4
L1 excess delay (m)
GMT Time (min)
SV # 11, order of fit = 4
excess delay rate (m/s)
GMT Time (min)
Elevation (deg)
GMT Time (min)
Azimuth (deg)
GMT Time (min)


SEP1589
L1 excess delay (m)
GMT Time (min)


SV # 13, order of fit = 4
L1 excess delay (m)
GMT Time (min)
SV # 13, order of fit = 4
excess delay rate (m/s)
GMT Time (min)
Elevation (deg)
GMT Time (min)
Azimuth (deg)
GMT Time (min)


SV # 14, order of fit = 4
L1 excess delay (m)
GMT Time (min)
SV # 14, order of fit = 4
excess delay rate (m/s)
GMT Time (min)
Elevation (deg)
GMT Time (min)
Azimuth (deg)
GMT Time (min)
SV # 3, order of fit = 4
L1 excess delay (m)
GMT Time (min)
SV # 3, order of fit = 4
excess delay rate (m/s)
GMT Time (min)
Elevation (deg)
GMT Time (min)
Azimuth (deg)
GMT Time (min)


SV # 12, order of fit = 4
L1 excess delay (m)
GMT Time (min)
SV # 12, order of fit = 4
excess delay rate (m/s)
GMT Time (min)
Elevation (deg)
GMT Time (min)
Azimuth (deg)
GMT Time (min)
SV # 11, order of fit = 4
L1 excess delay (m)
GMT Time (min)
SV # 11, order of fit = 4
excess delay rate (m/s)
GMT Time (min)
Elevation (deg)
GMT Time (min)
Azimuth (deg)
GMT Time (min)



Conclusion

The currently accepted models for ionospheric effects predict maximum excess delays during the peak of the sunspot cycle. We obtained measurements similar to the maximums predicted even though the current sunspot cycle has not yet reached its peak. It is hoped that this data will help further the understanding of these effects and provide insight and support data to others who may have taken measurements coincident with these.

Acknowledgments

Support for the development of the GPS receiver used to collect this data was provided by the Space Systems Division of the Air Force Systems Command (on contract F04701-83-C-0076). The authors thank the Air Force Systems Command for their support in allowing the publicaton of this data and Dr. Troy Fuchser for his assistance in data reduction, providing the plots and polynomials.

References

1. E.H. Martin, GPS User Equipment Error Models, Institute of Navigation conference, June, 1977.

AD-P006 318

Ionospheric Predictions -- II

N.C. Gerson

Laboratory for Physical Sciences
4928 College Avenue
College Park, MD 20740

March 2, 1990

I. Objective

The objective of this paper is to stimulate discussion on the topic of ionospheric predictions the better to determine whether additional effort is warranted, and if so, what and where.

II. Definitions

Commonly used terms in this field are listed in Table 1. Note that models and climatology usually do not include time variability or dynamics. Nevertheless it is the dynamics that causes most if not all of the problems in operational systems. Several other definitions are needed. Thus the ionosphere may be regarded as comprising four subregimes: mid-latitude, polar cap, equatorial band and auroral - sub auroral. The size, location and characteristics of these subregimes are dynamic and a function of the energy deposited upon the planet by the solar plasma.

III. Users

The climatology and models are used by different groups in various fashions (See Table 2). System planners employ ionospheric climatology and models in the same manner as an architect uses tropospheric climatology. Architects use tropospheric climatology in defining structures which will last most of the time. HF system planners use ionospheric climatology in defining terminal specifications: (e.g., transmitter power, frequency range, antenna type, etc.) for systems which will operate most of the time. It is almost immaterial which program is used since all are based on a similar climatology and all provide about the same results.

For example, systems designed on the basis of such prediction programs prepared by the USSR, UK, Canada, or the U.S. (HF MUFES or IONCAP) all work. The commonality rests upon the same ionospheric climatology. Also, a system designer, like a bridge builder, invariably includes a "safety factor". Thus, HF systems for a given purpose have about the same specifications irrespective of the program used.

Operators, in contrast to designers, encounter ionospheric weather. Thus, the dynamic effects degrade system operation. The degradations so plagued the communicator that he strove for alternate circuits: e.g. scatter circuits. He was in the forefront of establishing satellite links and now transoceanic fiber optics circuits.

The radar operator also suffers from the dynamics of the ionosphere, his procedure is analogous to that of a searchlight operator who seeks to identify and track birds by first directing his beam onto a lake surface and then illuminating the birds. He then views the images back scattered from the birds again in the lake surface. Unfortunately the lake is not

91 7 1 19

91-09684

TERMINOLOGY

TERM	BASIS	PURPOSE
PREDICTIONS	CLIMATOLOGY	SYSTEM SPECIFICATIONS
MODELS	SEMI-EMPIRICAL ASSUMPTIONS	STANDARD REFERENCE IONOSPHERE
FORECASTS (1 - 600 MINS)	TEMPORAL EXTRAPOLATION	IMPROVE ANALYSES OR OPERATIONS
NOWCASTS	SPATIAL EXTRAPOLATION	INSTANTANEOUS SNAPSHOT OVER DESIGNATED REGION (ATTAINABLE ??)

USER VIEWS

USER	RELIANCE	SATISFACTION	SAFETY HATCH
PLANNER	PREDICTIONS	ADEQUATE, MID LATITUDES	
COMMUNICATOR	FORECAST	DEFICIENT	DUALLING, RETRANSMIT, RESCHEDULE, SPREAD SPECTRUM
HF RADAR OPERATOR	FORECAST	INADEQUATE	TIME AND TRACK INTEGRATION

NOTE 1: FORECAST DEPENDS ON SOLAR FORECASTS

NOTE 2: MILLISECOND PREDICTIONS IMPOSSIBLE

a specular reflector but a rippling surface traversed by waves of different periods, velocities and amplitudes. These dynamic features distort the images. Some amelioration of these effects is possible by recognizing that the first derivative of the tracks should be continuous.

Other users are also adversely affected by ionospheric dynamics. Forecasts for systems requiring millisecond stabilities are impossible. Changes of this order cannot be forecast.

In short the variable ionosphere is the chief culprit in system degradation. To forecast its changes, much more knowledge is required on the hydrodynamics of the lower ionosphere and the hydromagnetics of the upper ionosphere. Motions arise because of the introduction of energy sources and sinks which give rise to thermal and pressure gradients. Further, electric forces are present at higher latitudes. A long-term research effort in ionospheric dynamics is needed, before we can forecast its turbulence and motions. This aspect is properly the province of the scientific community.

IV. Requirements

For practical purposes the requirements of system operators are listed in Table 3. Efforts are needed in three areas: models, computational techniques and engineering.

Models should include standard deviations (of electron density, height of maximum, half thickness, etc.), and should incorporate a suitable magnetic index (Ae) representative of the energy deposition upon the planet. The standard deviation, together with similar tolerances for absorption, noise intensity, antenna lobe patterns, etc. would allow the error budget to be traced through the computational sequences and thus improve the credence or computed "reliability".

Computer programs require a complete reevaluation based upon modern, not 1960, procedures. Computer graphics must be utilized. User friendly and user understandable outputs are essential. It is time to stop delivering to potential users a one meter stack of computer printout topped by a medicine man proclaiming "Only I can explain it." Some possible outputs are shown in figures 1-6 where MUF footprints are shown for Greenwich, U.K. during periods of high and low sunspot number. Similar outputs can be provided for the LUF or other outputs of the program.

Plots to illustrate results from the ICED model could not be provided because of a lack of the source code. However, such plots should be made and compared with figures 1-6 to determine whether significant improvements have been made. The figures show that MUF footprints for HFMUFES and IONCAP at low sunspot number are similar. Greatest differences between HFMUFES and IONCAP occur during periods of high sunspot number. However, even in this case, the practical circuit designer would probably provide the same specifications for a given system irrespective of the program used.

Engineering aspects should be locally produced and controlled as subroutines. They must interface with the ionospheric program but not tamper with the models involved. They should be transportable and tailored to a particular user.

The figures show that the frequency support provided by the ionosphere (i.e. the ionospheric bandwidth) is much wider during the years of solar maximum than minimum. During solar minimum, all users are constrained towards the lower end of the HF band, where a minor solar event can cause considerable absorption and circuit disruption for all users.

DESIDERATA

MODELS	COMPUTATION	ENGINEERING
DYNAMIC, STATIC	MORE RAPID MORE INNOVATIVE USER FRIENDLY	ALL SUBROUTINES, LOCALLY PRODUCED ANTENNA
REGIME EQUATORIAL, AURORAL, POLAR CAP	USER UNDERSTANDABLE EXPLICIT DOCUMENTATION DOD INTEROPERABLE	NOISE

Further, interference problems become much more severe because of the packing of all users in a small segment of the HF band. Practical systems must consider these effects and incorporate adaptive arrays prior to this time period.

At the present time, for design purposes, it is almost immaterial which ionospheric program is used. Whether ICED will provide materially improved design specifications remains uncertain.

Some comments must be made about dynamic models. Both the troposphere and ionosphere represent atmospheric fluids on a rotating planet. The future state of any fluid can only be determined if we know: a) its present state and b) the applied forces and energy sources and sinks. The troposphere contains a relatively dense network of observational sites. Further, snapshots of the tropospheric disk can be obtained. However, even with these advantages, forecasts are difficult in some areas. The ionosphere contains a much more sparse observational network. Airglow and ultraviolet snapshots of the ionospheric disk may become available, but their interpretation is not yet clear. Despite the valiant efforts of Akasofu, Fedder, Rostoeker, Schunk, Carlson, etc. a decade or more will still be required before adequate ionospheric forecasts become available. Under these circumstances, timely and accurate forecasts for each layer remain as distant goals. Efforts of this type should be funded by the research and not the operational community.

V. Suggestions

A possible approach is shown in Table 4. It suggests that the entire function of predictions be assigned to the Air Weather Service which now has a charter to examine and improve atmospheric and ionospheric models. Also, it is obvious that any prediction or forecast of the ionospheric state must rely on a forecast of solar conditions. The latter forecasts are a decade or more away.

The computational and engineering aspects of the program can be introduced today without any additional research. While some model improvements are possible, we should note that during the 1970s and 1980s, about 15 papers were published respectively, without any notable improvements in the current prediction programs.

VI. Conclusions

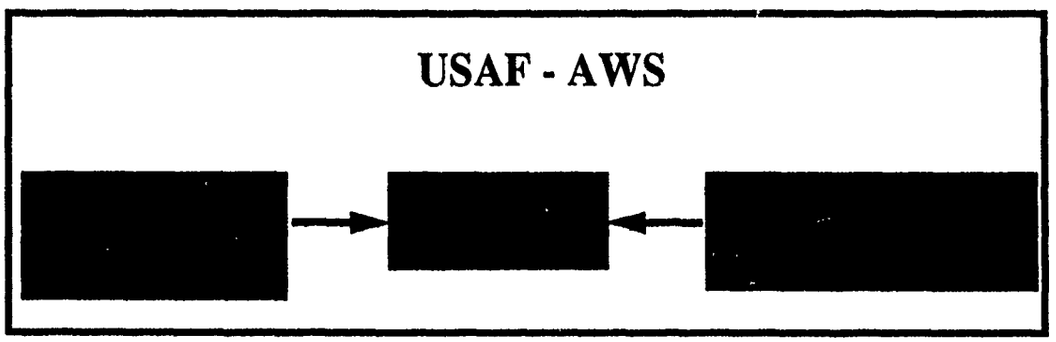
At a time of shrinking resources and increasing responsibilities, DoD must do more with less. An internal group which will generate, document and maintain the HF design programs is essential. The policy of allowing numerous spigots to nourish eager contractors of different competencies may well be over. DoD must invest internally for its own future, and must attempt to retain some long term corporate memory.

On an overall basis, we must recognize that although long term research efforts are needed before credible ionospheric forecasts can be made, research of this type must be funded by the scientific (NASA, NSF, ONR, AFOSR) and not the operational community. Secondly, improvements in the present predictions are possible and should be assigned to the Air Weather Service with the proviso that it consider all DoD needs.

It is also possible to conclude from figures 1-6 that many HF systems will require adaptive arrays during the years of solar minimum. They should be implemented now to overcome the expected decrease in the ratio of signal / (noise +

SUGGESTED OPTIMAL SOLUTION

- **DOD GENERATED, DOCUMENTED, AND MAINTAINED PROPAGATION PROGRAMS (NO INDUSTRIAL OR GOVERNMENT CONTRACTORS)**



interference).

We must also consider the future. At present different DoD elements are establishing oblique sounder grids with different equipments, outputs and separations. This haphazard approach is wasteful of bandwidth and dollars. A far more effective approach would include a grid of calibrated sounders of common design and common output format suitable for computer input. Grid spacings of 2-3 Mm would be desired insofar as possible. These data would be extrapolated globally using the same spherical harmonic expansions now incorporated in most prediction programs. Their representation using computer graphics would provide a global nowcast. Current needs of the Navy, Air Force and other DoD elements could be satisfied simultaneously with such a network. It is time to optimize total DoD needs.

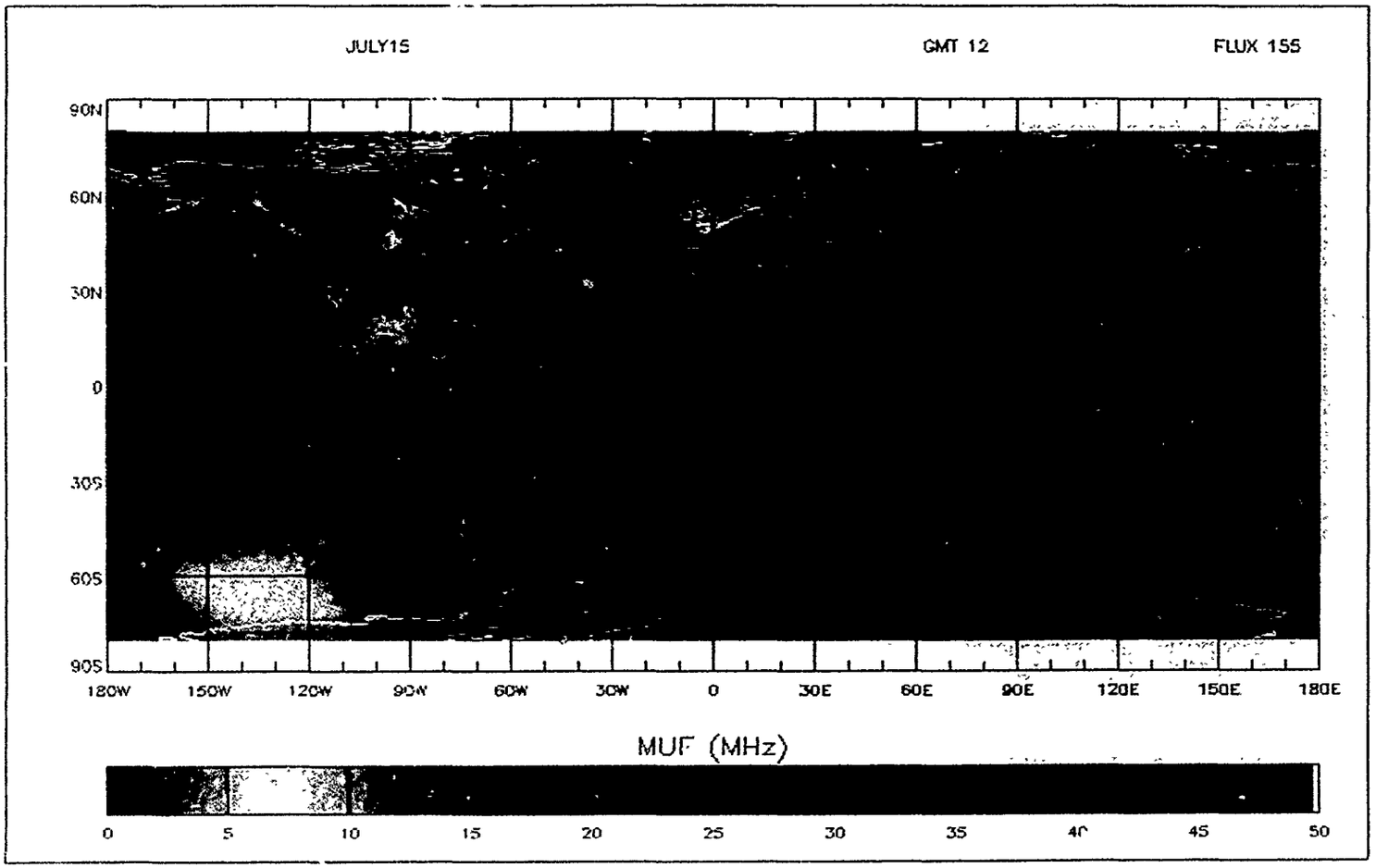


Figure 1 MUF footprint, Greenwich, R=110, HF MUFES.

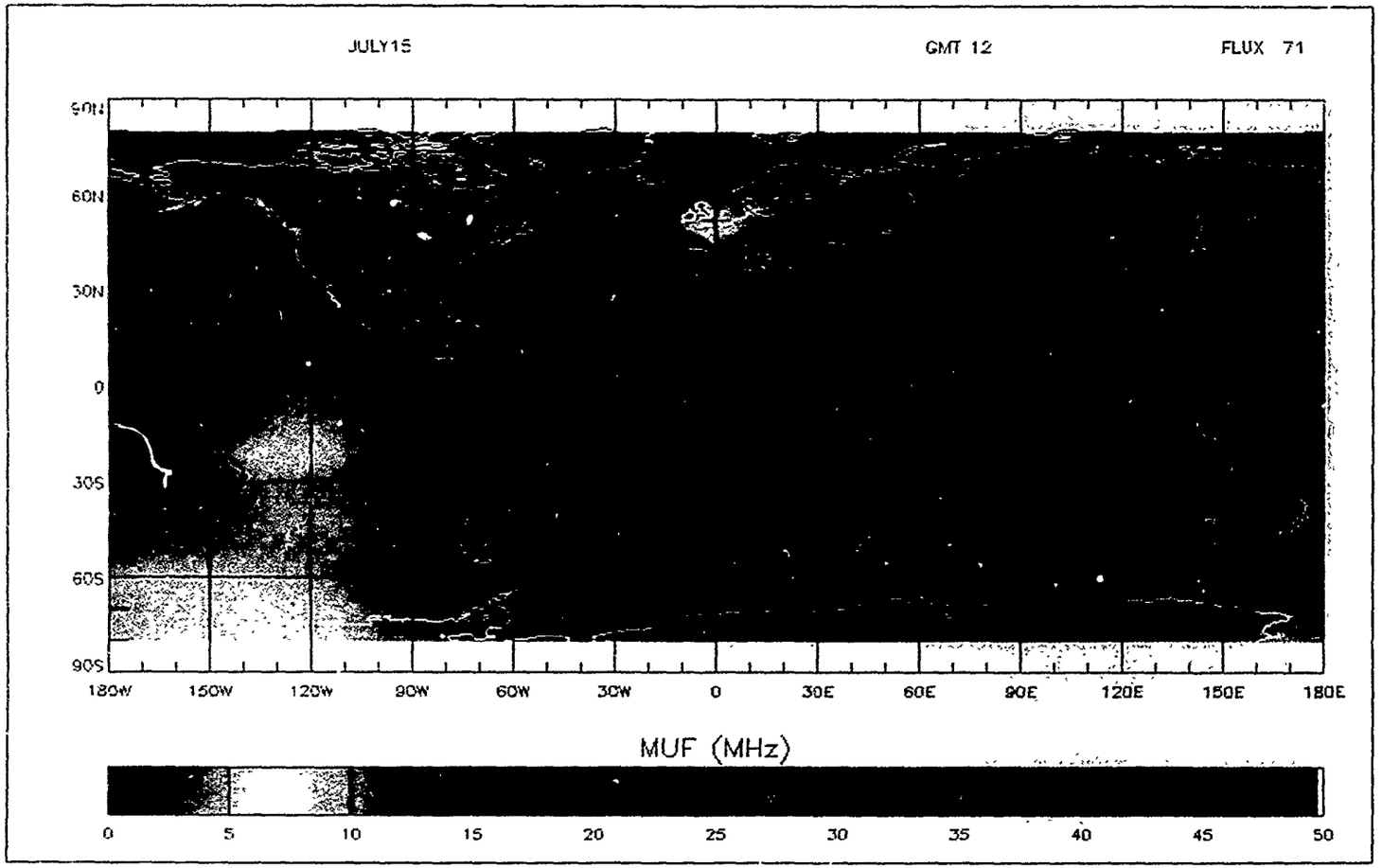


Figure 2 MUF footprint, Greenwich, R=10, HF MUFES.

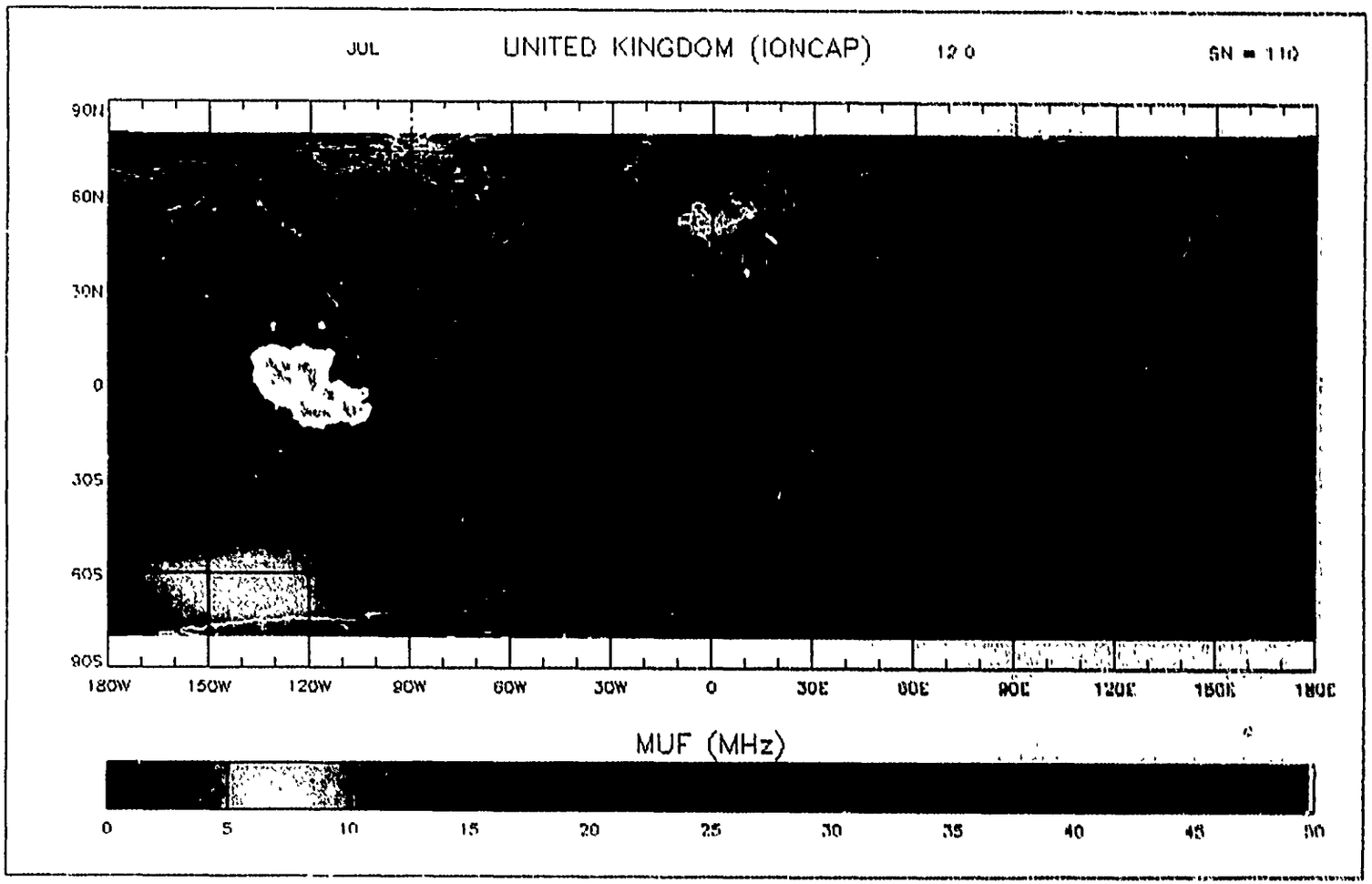


Figure 3 MUF footprint, Greenwich, R=110, IONCAP.

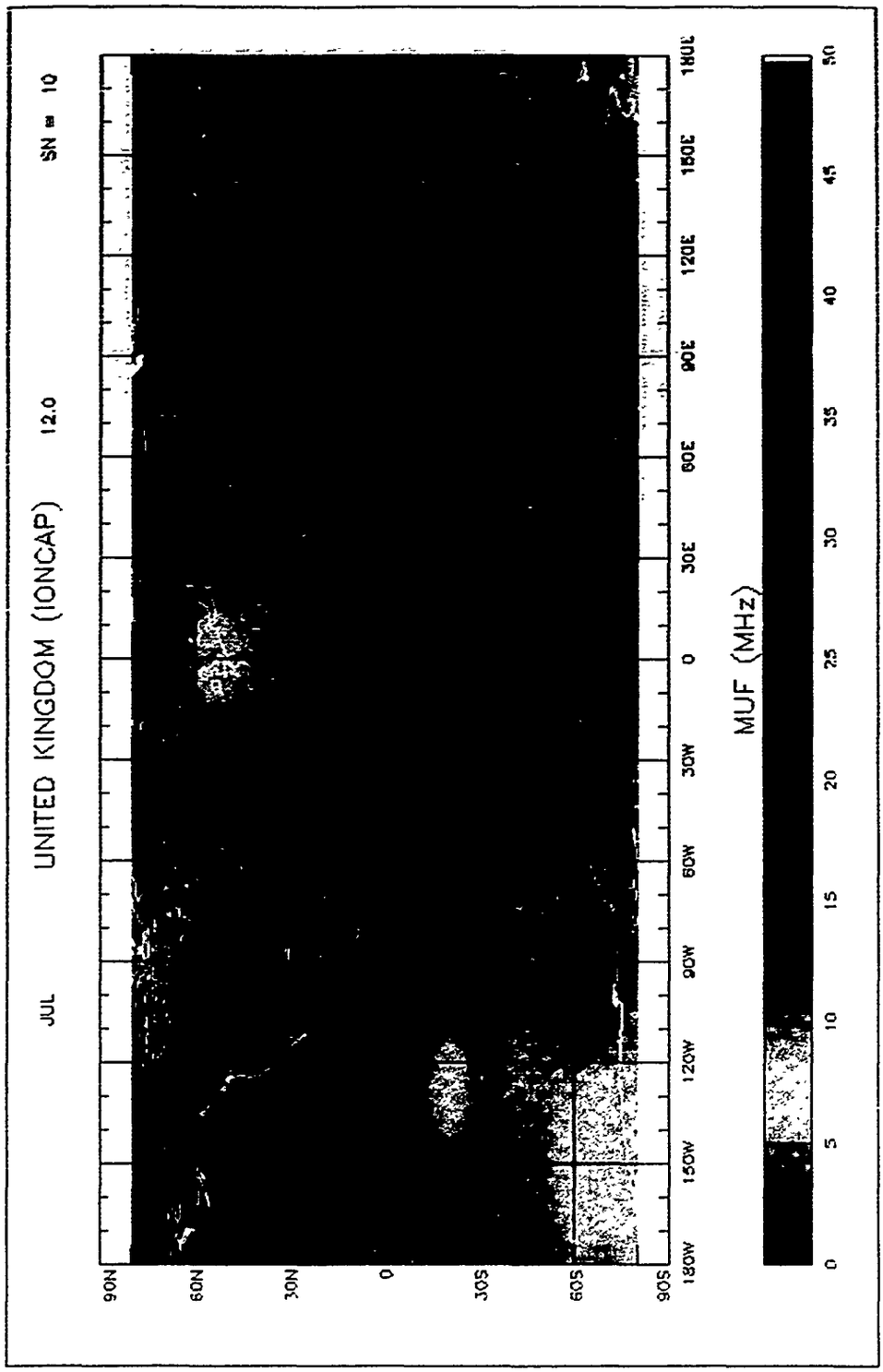


Figure 4 MUF footprint, Greenwich, R=10, IONCAP.

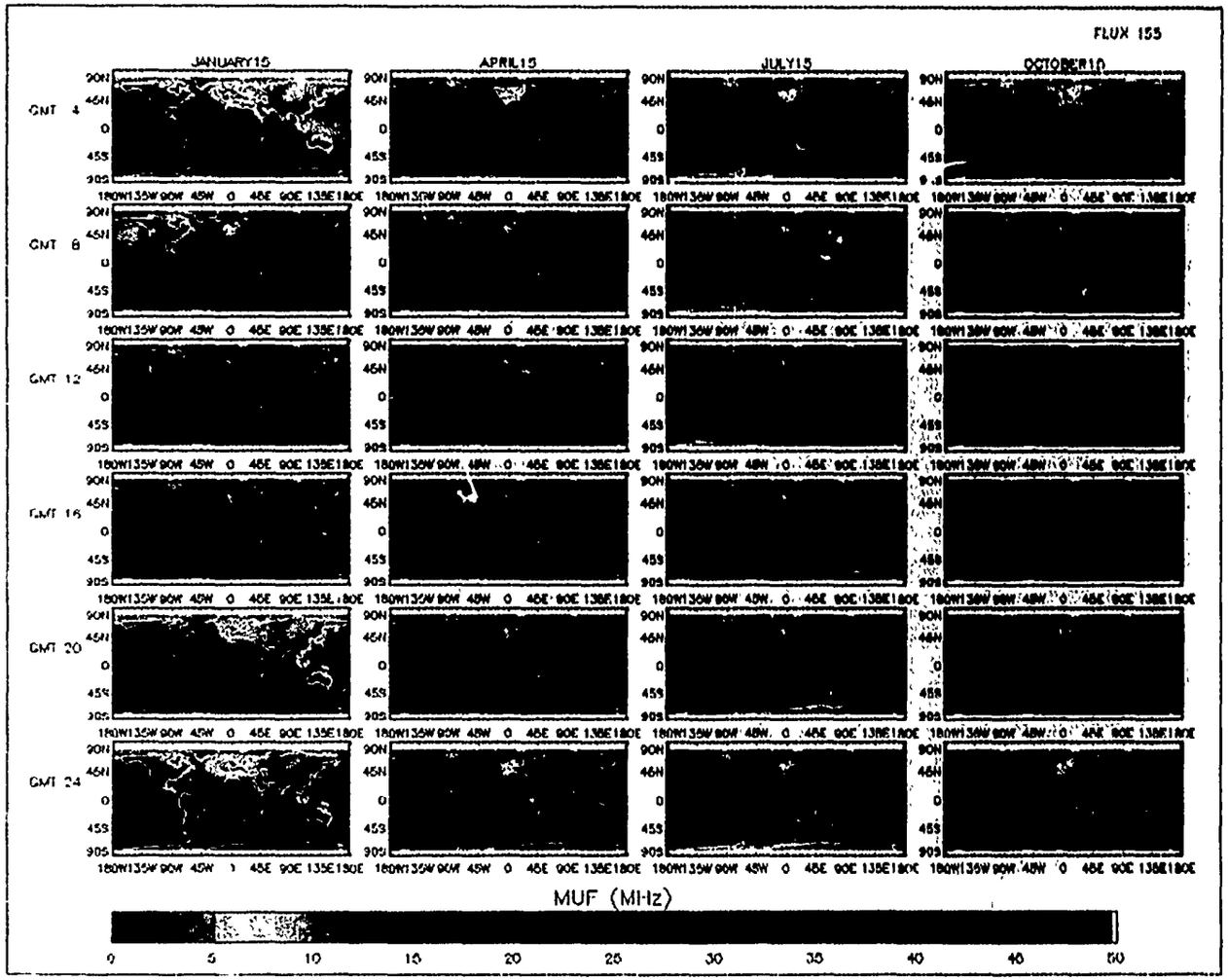


Figure 5 MUF footprints by season and hour, R=110, Greenwich, HF MUFES.

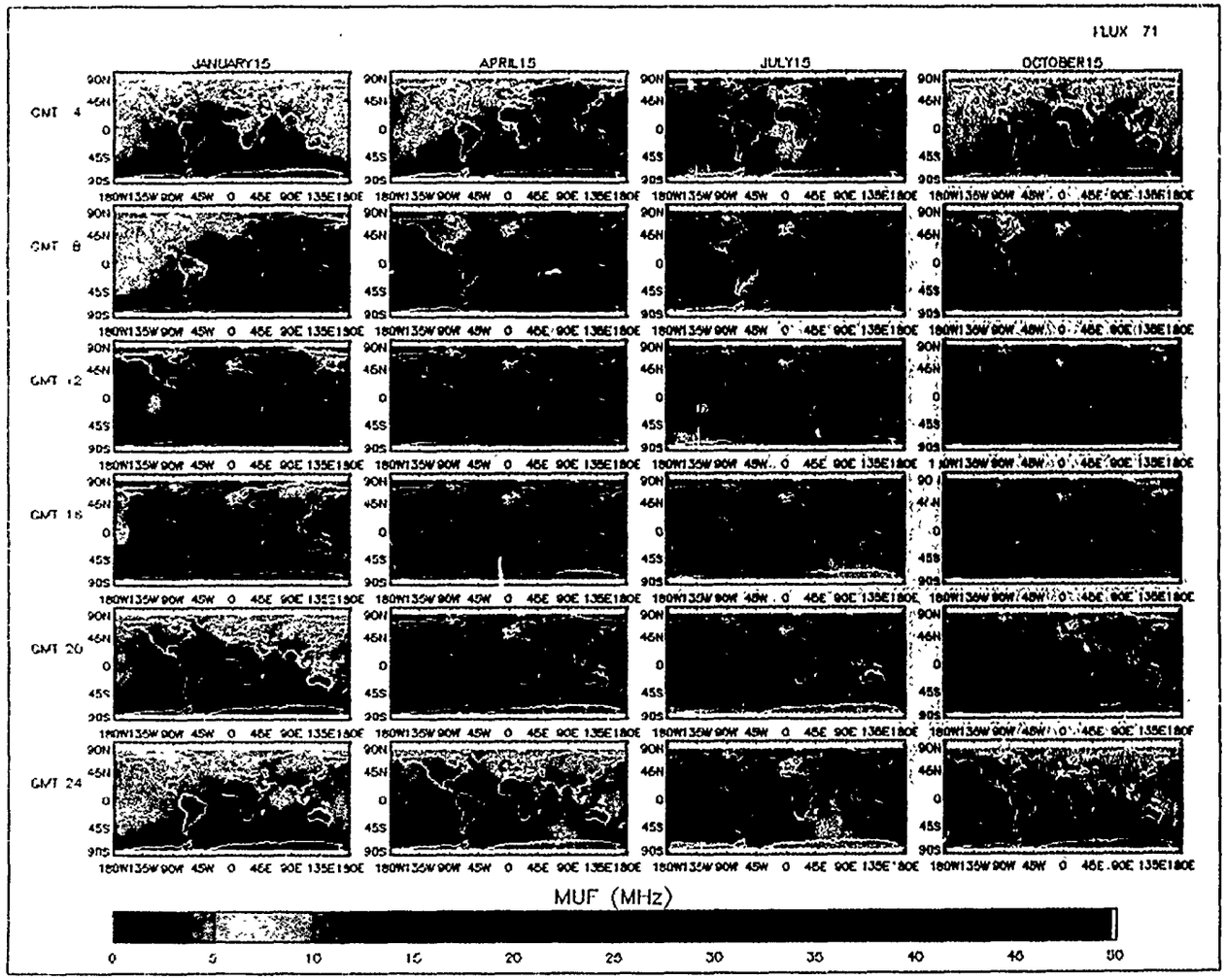


Figure 6 MUF footprints by season and hour, R=10, Greenwich, HF MUFES.

AD-P006 319

91-09682

STATUS OF THE NAVAL OCEAN SYSTEMS CENTER'S LONG WAVE PROPAGATION CAPABILITY

Jerry A. Ferguson
Naval Ocean Systems Center
San Diego, CA 92152-5000

INTRODUCTION

This paper describes a new software version of the Naval Ocean Systems Center's Long Wavelength Propagation Capability (LWPC). This program applies the concept of the earth-ionosphere waveguide to very low frequency (vlf: 10 kHz to 30 kHz) and low frequency (lf: 30 kHz to 60 kHz) radio propagation. The propagation capability has been incorporated into a FORTRAN program which sets up calculation of mode parameters along arbitrary propagation paths or for user defined operating areas. In the latter case, the paths are automatically selected to account for areas of low conductivity. Previous dependence on VMS has been eliminated. Substantial improvements have been made in storage and identification of the program outputs. A set of programs have been developed for graphically displaying the output from the program.

The first version of the LWPC was a collection of separate FORTRAN programs linked together in operation by a command procedure written in Digital Equipment Corporation's VMS operating system command language (Ferguson and Snyder, 1989a,b). The initial version, designated LWPC-0 in this paper, was originally a test of a concept, making full integration of the constituent programs premature. The core of the new LWPC, designated LWPC-1, is embodied in a single FORTRAN program named LWPM (short for Long Wavelength Propagation Model). This program implements all of the features of the original VMS version. In addition, LWPC-1 includes a number of auxiliary FORTRAN programs which provide summaries of the contents of files and can be used to update and graphically display the contents of its output files.

The propagation model implemented in the LWPC-1, which is identical to that in LWPC-0, treats the space between the earth's surface and the lower ionosphere as a waveguide. The upper boundary of this waveguide is characterized by an isotropic conductivity which increases exponentially with height. The exponential increase is defined by a log-linear slope and a reference height. The LWPC incorporates a model which defines an average value of the slope and reference height which depends on frequency and diurnal condition. The height of the nighttime ionosphere over the polar caps is lower than it is at middle and equatorial latitudes. This model has been compared extensively with available measurements. The lower boundary of the LWPC is based on the Westinghouse Geophysics Laboratory conductivity map (Morgan, 1968). Both boundaries of the earth-ionosphere waveguide are known to have possibly large uncertainties in certain regions and seasons.

A typical usage of the LWPC is to generate geographical maps of signal availability for coverage analysis. The program makes it easy to set up such problems by automating most of the required steps. The user specifies the transmitter location and frequency, the orientation of the transmitting and receiver antenna and the boundaries of the operating area. The program automatically selects paths using a coarse resolution of 15 degrees to ensure that the operating area is fully covered and then using a fine resolution of 3 degrees to ensure that all significant low conductivity areas are included. The diurnal conditions and other relevant geophysical parameters are then determined along each path. After the mode parameters along each path are determined, the signal strength along each path is computed and is then mapped onto a grid overlaying the operating area. The data in the grid include a user specified (constant) standard deviation of the signal throughout the grid. The signal strengths in this grid can be displayed by other programs to be described later.

91

LWPC-1 incorporates a number of improvements over LWPC-0 including reduced dependence on the operating system, better integration of the intermediate data storage and modification of the input to use control strings coupled with input parameters. An option for previewing the propagation paths to be processed by the program is included. A map with major land masses graphically filled is included in this preview option as well as in the geographical coverage displays. The main program, LWPM, incorporates essential elements of several of the programs used in LWPC-0: PRESEG (which sets up the paths), MODEFNDR (which determines the initial solutions for each set of segments), the Segmented Waveguide program (which determines the solutions for segments of similar type), FASTMC (which does the mode summation along the path) and OPAREA (which sets up the data for plotting coverage area maps). While LWPC-0 generates a log file and a mode parameter data file for each path which is processed, LWPC-1 generates only one log file and one data file, making file management easier.

Three noise models are currently available: 'ITSN' names the ITS noise model of Zacharisen and Jones (1970), 'NTIA' names the noise model of Spaulding and Washburn (1985) and 'DECO' names the DECO Westinghouse thunderstorm-based noise model (Maxwell et al., 1970). The NTIA model is identical to the new CCIR noise model (CCIR, 1986). The model named ITSN is the implementation of CCIR 322 (CCIR, 1963) which maps the basic noise map parameter, Fam, in Universal Time. The model named NTIA is the new noise model developed using additional measurements and has since become the new CCIR model described in CCIR Report 322-3 (CCIR, 1986). These two models are based on surface mappings of measurements at a limited number of sites. The DECO model of atmospheric noise is limited to the VLF regime (10 to 30 kHz) and uses a data base of thunderstorms and does propagation calculations from these thunderstorms to the receiver sites. This model was calibrated using the same measurement sites which were used to develop the other two models. Because the model uses propagation calculations instead of polynomials, it is a much longer running model and is used infrequently.

PROGRAM CONTROL

The LWPC-1 uses character strings for program control and to specify options. The control strings have the same meaning and use amongst all of the programs. On input, most control strings may be abreviated. For example, the string "TX-DATA" can be entered in upper or lower case and can be shortened to "TX-D". If a control string is shortened too much, it will not be recognized and execution will stop. To make it easier to read control strings which are composed of more than one word, dashes are used to separate the words, such as the above mentioned "TX-DATA". A blank in the first column of a line of data causes the string to be treated as a comment line, allowing the user to annotate run streams for documentation and to provide prompts for editing.

After the necessary control strings and their associated data are specified, a specific control string named "START" is used to initiate the calculations. The program LWPM first creates a status file. This file is named with the extension STA and contains a list of the bearing angles and lengths of all of the paths. As the calculation of mode parameters along each path is completed, the parameters are written to a file named with the extension MDS and the corresponding entry in the status file is updated with the date and time of completion and the CPU time used. Calculations for successive paths continues automatically. If the computer run aborts for some reason, then the run can be restarted simply by resubmitting the original command file, after the error has been corrected. The program checks for the existence of the status file and continues execution at the first entry which does not have a date and time entry. When all paths have been processed, the program calculates the field strength along each path using the parameters specified for the transmitting antenna and the receiver. These data are written to a file named with the extension LWF and the status file is updated. If the propagation paths were set up automatically by user specification of an operating area, the program uses the data in the LWF file to generate a file named with the extension GRD and the status file is updated. The GRD file contains values of the signal strength and its standard deviation in a grid of latitude vs. longitude which covers the operating area. This last file may now be used in a program named PLOT_GRD to obtain geographical displays of the signal levels, signal to noise levels, or, together with other GRD files, signal to jammer levels.

OUTPUT DATA

An important improvement in LWPC-1 is in the handling of data files. Compared to LWPC-0, more information is stored in these files and the formats have been developed to allow for future enhancements. The new files are all written in unformatted form and a special program named SCAN has been written to print out summaries of the parameters stored in them. Some parameters share common usage throughout the full set of output data files. Whenever appropriate, these parameters from one file are passed on to subsequent types of files in order to provide continuity and some audit trail information. The first record of all data file types contains the same information. This first record contains an 8 character string to be used to record archive information. It also contains three strings, named MDS-FILE-ID, LWF-FILE-ID and GRD-FILE-ID. Each of these strings contains the date the file was written, a randomly generated, 3 character string to uniquely identify the file and the full file name including the directory tree. The presence of these strings in each type of file provides information regarding the history of the data which was used to produce the file. For instance, each GRD file identifies the LWF file which provided the signal strength data and the MDS file which provided the mode parameter data used to calculate the LWF data. Generally, the program identification will be 'LWPM-V10'. This first record also contains the list of parameters to identify the propagation paths for the data sets.

A utility program named SCAN is used to produce summaries of the data in any of the output files. Examples of the output from this program using the sample problem are shown below.

SAMPLE PROBLEM

To illustrate the capabilities of LWPC-1 a sample case is presented. The run stream is shown in Figure 1. An operating area named "Mediterranean" is used to define the propagation paths. The root file name is "SAMPLE" so that the following files were created: SAMPLE.sta, SAMPLE.mds, SAMPLE.lwf and SAMPLE_mediterranean.grd. The transmitter parameters were retrieved from a specification file. The preview option was exercised to produce the plot shown in Figure 2. A preview option was then "turned off" by indenting the control string one space.

The second execution of the program produced the MDS, LWF and GRD files named above. Figure 3 shows the output from the SCAN program for the first path in the MDS file. The output identifies the program which was used to generate the file and the date that it was run. The data under the heading which begins with "mds" is a summary of the parameters for each segment along the path. Figure 4 shows the output from the SCAN program for the LWF file. The parameter under the heading "nc" identifies the number of field components in the mode sums. The data under the heading which starts with "bearng" identify the paths by their bearing angle and shows the length of each path. The signal strength as a function of distance from the transmitter along the northernmost (bearing = 24 degrees; upper panel) and southernmost (bearing = 72 degrees; lower panel) paths are plotted in Figure 5. Along the bottom of each graph is a summary of the important path segmentation data, namely, the height of the ionosphere and the ground conductivity. The beginning of each segment is indicated by a small diamond in the curve representing the ground conductivity. In the northern path, the rapid change in signal strength that occurs near 4 Mm is due to the effect of Greenland.

Figure 6 shows the SCAN output for the GRD file. This output identifies the boundaries of the operating area. The values under the headings "nlat" and "nlon" are the number of latitudes and longitudes used to define the grid. Sample contour maps generated by PLOT_GRD using the run stream in Figure 7 are presented in Figures 8 and 9. The first of these figures shows contours of signal strength and the second shows contours of signal to atmospheric noise ratio. The rapid drop in signal strength as paths cross Greenland is seen in the closely spaced and nearly vertical contour lines in the eastern part of the operating area.

DISTRIBUTION

The source code and associated data sets are available from the Defense Technical Information Center.

REFERENCES

CCIR, World Distribution and Characteristics of Atmospheric Radio Noise, Rept. 322, Documents of the Xth Plenary Assembly, Geneva, 1963.

CCIR, World Distribution and Characteristics of Atmospheric Radio Noise, Rept. 322-3, Documents of the XVIth Plenary Assembly, Geneva, 1986.

Ferguson, J. A., F. P. Snyder, Long-Wave Propagation Capability Program Description and User's Guide, NAVOCEANSYSCEN TD 1449, 1989a; available from DTIC: ABD130808.

Ferguson, J. A., F. P. Snyder, The NAVOCEANSYSCEN's Long Wavelength Propagation Capability, NAVOCEANSYSCEN TD 1518, 1989b; available from DTIC: ABD133690.

Maxwell, E. L., D. L. Stone, R. D. Croghan, L. Bell, A. D. Watt, Development of a VLF Atmospheric Noise Prediction Model, Westinghouse Georesearch Laboratory Rept. 70-1H1-VLF NO-R1, 1970.

Morgan, R. R., World-wide VLF Effective Conductivity Map, Westinghouse Electric Corp. Rept. 8013F-1, 1968.

Spaulding, A. D., J. S. Washburn, Atmospheric Radio Noise: Worldwide Levels and Other Characteristics, U. S. Dept. of Commerce NTIA Rept. 85-173, 1985.

Zacharisen, D. H., W. B. Jones World Maps of Atmospheric Radio Noise in Universal Time by Numerical Mapping, U. S. Dept. of Commerce, Office of Telecommunications Rept. OT/ITSRR 2, 1970.


```
$ run lwpc-dir:lwpm
case-id      OMEGA coverage of the Mediterranean
tx           SAMPLE
tx-data      OMEGA-D lwp:[dat]xmtr.list
ionosphere   lwpm day
op-area      Mediterranean 30 10 45 -45
lwf-vs-dist  ,11000,,
print-swg    2
print-lwf    2
 gcpath
 preview
map-area     LANT rect 20 100 80 -50 7 5
map-type     conductivity
start
quit
```

Figure 1: Sample run stream for LWPM

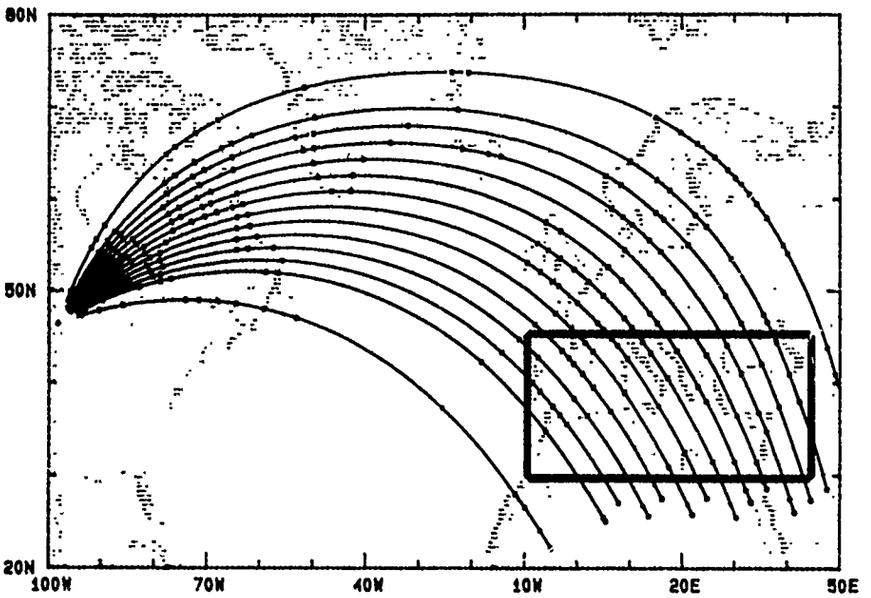


Figure 2: Output from "PREVIEW" option.

```
file_id:  11DEC89 LSF [.lwp.tst]SAMPLE.mds
          ***
          ***
prgm_id:  LWPM-V10
case_id:  OMEGA coverage of the Mediterranean
prfl_id:  LWPM Day
xmtr_id                tlat    tlon    freq
OMEGA-D                46.4    98.3    10.2
path_id                lat1    lon1    lat2     lon2
Mediterranean          30.0    10.0    45.0    -45.0
bearng    rlat     rlon     rrho
  24.0    99.0      0.0       0.
```

mds	lat	lon	rho	azim	dip	bfield	sigma	epsr	beta	hprime
4	46.4	98.3	0.	13.6	74.6	0.573	1.E-02	15.	0.30	74.0
4	49.8	96.0	420.	17.5	77.6	0.582	1.E-03	15.	0.30	74.0
4	54.6	91.9	1020.	27.2	81.2	0.588	3.E-04	10.	0.30	74.0
4	55.5	91.0	1140.	30.0	81.8	0.588	3.E-03	15.	0.30	74.0
4	57.1	89.3	1340.	35.7	82.8	0.587	4.E+00	81.	0.30	74.0
4	64.8	77.7	2400.	84.3	85.0	0.570	3.E-04	10.	0.30	74.0
4	65.6	76.0	2520.	89.9	84.9	0.567	4.E+00	81.	0.30	74.0
4	67.2	72.2	2760.	99.5	84.6	0.560	3.E-04	10.	0.30	74.0
4	68.6	67.9	3000.	107.0	84.1	0.554	4.E+00	81.	0.30	74.0
4	72.1	51.6	3720.	119.5	82.6	0.534	1.E-05	5.	0.30	74.0
4	73.7	23.7	4640.	124.9	80.6	0.515	1.E-04	10.	0.30	74.0
4	73.6	20.5	4740.	125.3	80.4	0.513	4.E+00	81.	0.30	74.0
4	68.7	-15.1	6100.	130.2	77.3	0.501	3.E-03	15.	0.30	74.0
4	67.2	-19.8	6360.	131.6	76.6	0.500	1.E-03	15.	0.30	74.0
4	65.9	-23.0	6560.	132.7	76.0	0.499	4.E+00	81.	0.30	74.0
4	64.4	-26.1	6780.	134.0	75.4	0.498	3.E-03	15.	0.30	74.0
4	63.0	-28.7	6980.	135.3	74.7	0.498	1.E-02	15.	0.30	74.0
4	62.2	-30.1	7100.	136.1	74.3	0.497	3.E-03	15.	0.30	74.0
4	61.5	-31.3	7200.	136.7	74.0	0.497	1.E-02	15.	0.30	74.0
4	60.4	-32.8	7340.	137.7	73.5	0.497	3.E-03	15.	0.30	74.0
4	59.4	-34.2	7480.	138.6	72.9	0.496	3.E-02	15.	0.30	74.0
4	57.9	-36.0	7680.	140.0	72.1	0.495	1.E-02	15.	0.30	74.0
4	54.9	-39.2	8060.	142.7	70.5	0.494	3.E-02	15.	0.30	74.0
4	54.0	-40.1	8180.	143.5	69.9	0.493	1.E-02	15.	0.30	74.0
4	43.8	-47.5	9440.	151.5	62.1	0.479	4.E+00	81.	0.30	74.0
4	40.9	-49.1	9780.	153.4	59.3	0.472	1.E-02	15.	0.30	74.0
4	40.1	-49.6	9880.	153.9	58.4	0.470	4.E+00	81.	0.30	74.0
4	36.9	-51.1	10260.	155.7	54.8	0.461	3.E-03	15.	0.30	74.0
4	34.9	-52.0	10500.	156.8	52.2	0.455	3.E-03	15.	0.30	74.0
	99.0									

Figure 3: SCAN Output for MDS files

```
file_id:  11DEC89 LSF [.lwp.tst]SAMPLE.mds
          11DEC89 XKL [.lwp.tst]SAMPLE.lwf
          ***
prgm_id:  LWPM-V10
case_id:  OMEGA coverage of the Mediterranean
prfl_id:  LWPM Day
xmtr_id                 tlat    tlon   freq
OMEGA-D                 46.4    98.3   10.2
path_id                 lat1    lon1   lat2    lon2
Mediterranean           30.0    10.0   45.0   -45.0
nc nrpt bearng  rhomx   rlat    rlon    rrho    pwr     dist incl headng talt ralt
 1  551   24.0 10500.   99.0     0.0      0.     10.      0.   0.    0.0  0.0  0.0
 1  551   30.0 11000.   99.0     0.0      0.     10.      0.   0.    0.0  0.0  0.0
 1  551   33.0 11000.   99.0     0.0      0.     10.      0.   0.    0.0  0.0  0.0
 1  551   36.0 11000.   99.0     0.0      0.     10.      0.   0.    0.0  0.0  0.0
 1  551   39.0 10500.   99.0     0.0      0.     10.      0.   0.    0.0  0.0  0.0
 1  551   42.0 10500.   99.0     0.0      0.     10.      0.   0.    0.0  0.0  0.0
 1  551   45.0 10500.   99.0     0.0      0.     10.      0.   0.    0.0  0.0  0.0
 1  551   48.0 10000.   99.0     0.0      0.     10.      0.   0.    0.0  0.0  0.0
 1  551   51.0 10000.   99.0     0.0      0.     10.      0.   0.    0.0  0.0  0.0
 1  551   54.0  9500.   99.0     0.0      0.     10.      0.   0.    0.0  0.0  0.0
 1  551   57.0  9500.   99.0     0.0      0.     10.      0.   0.    0.0  0.0  0.0
 1  551   60.0  9000.   99.0     0.0      0.     10.      0.   0.    0.0  0.0  0.0
 1  551   63.0  9000.   99.0     0.0      0.     10.      0.   0.    0.0  0.0  0.0
 1  551   72.0  8500.   99.0     0.0      0.     10.      0.   0.    0.0  0.0  0.0
```

Figure 4: SCAN Output for LWF files

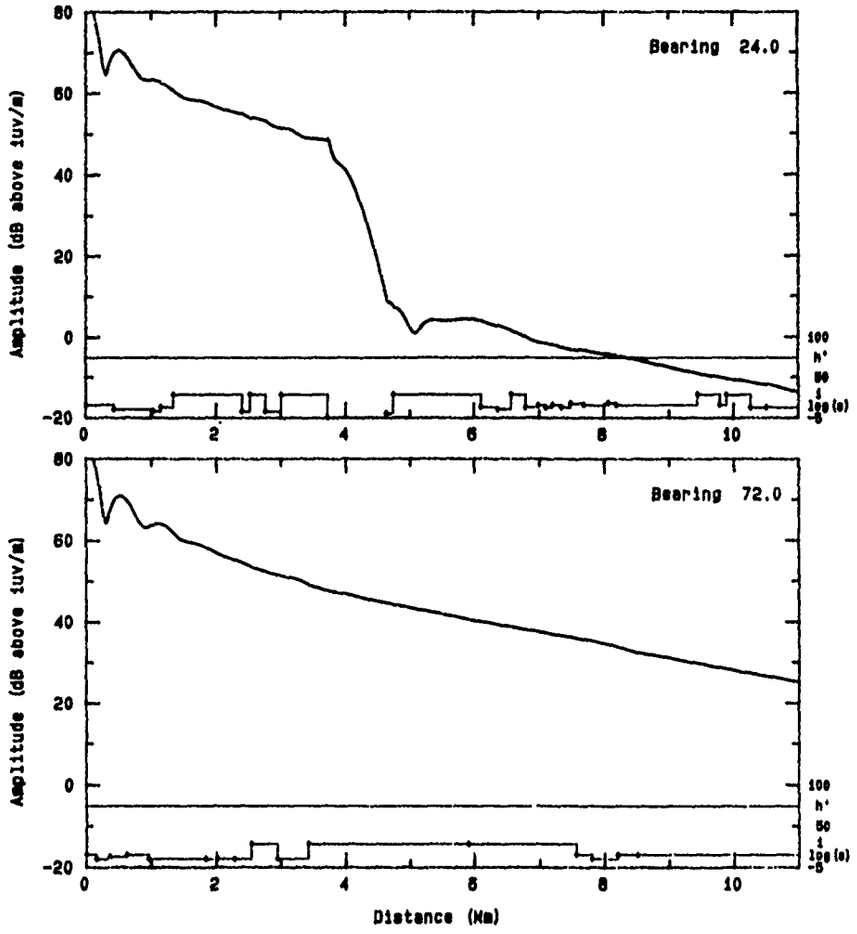


Figure 5: Output from PLOT_LWF for two paths in Figure 2.

```
file_id:  11DEC89 LSF [.lwp.tst]SAMPLE.mds
          11DEC89 XKL [.lwp.tst]SAMPLE.lwf
          11DEC89 XKL [.lwp.tst]SAMPLE_mediterranean.grd
prgm_id:  LWPM-V10
case_id:  OMEGA coverage of the Mediterranean
prfl_id:  LWPM Day
xmtr_id                tlat    tlon   freq
OMEGA-D                46.4    98.3   10.2
path_id                lat1    lon1   lat2    lon2
Mediterranean          30.0    10.0   45.0   -45.0
area_id               xlat1   xlon1  xlat2   xlon2  nlat nlon
Mediterranean          30.0    10.0   45.0   -45.0    13   45
nc power  incl  headng  talt  ralt  mn/dy/yr:UT    bandw  adjny  stndev
 1   10.    0.     0.0   0.0   0.0  00/00/00:0000     0.    0.0     3.0
```

Figure 6: SCAN Output for GRD files

```
$ run lwpc-dir:plot_grd
case-id      OMEGA coverage of the Mediterranean
tx           SAMPLE
op-area      Mediterranean       30 10 45 -45
map-area     Mediterranean rect 30 10 45 -45 6 3
map-type     land coast
a-noise      ntia July 18 1000
cntr-range   ,,3
ta-level     50
plt-s        1
plt-s/n      1
start
quit
```

Figure 7: Sample run stream for PLOT_GRD

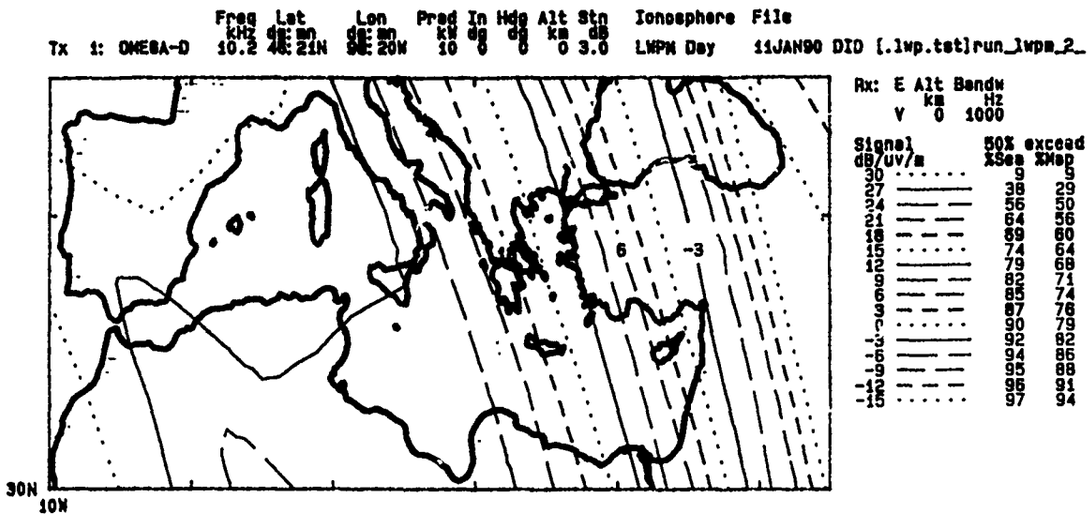


Figure 8: Signal contours for the operating area of Figure 2.

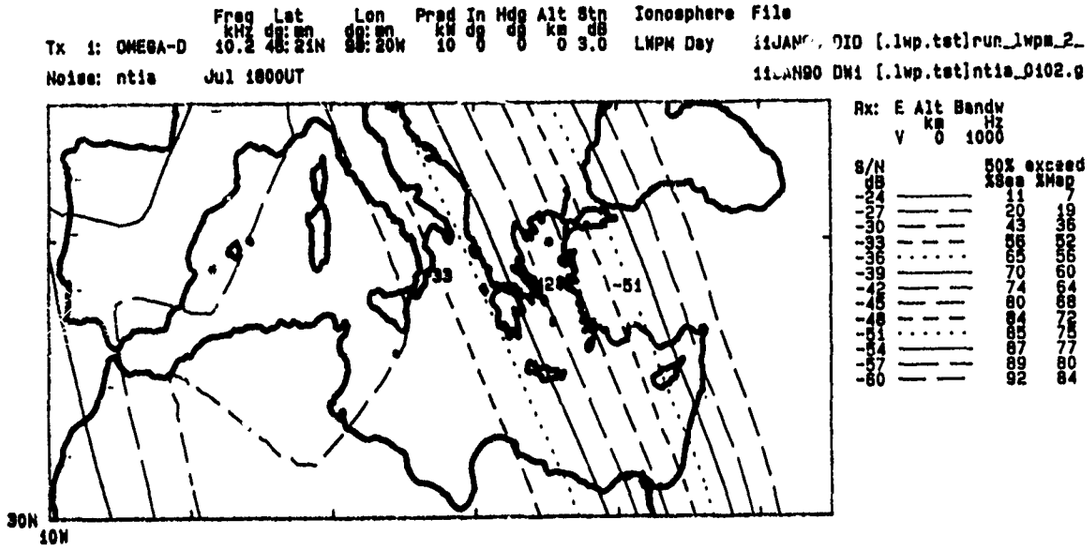


Figure 9: SNR contours for the operating area of Figure 2.

AD-P006 320

3-D MODELING OF SUBIONOSPHERIC VLF PROPAGATION IN THE PRESENCE OF LOCALIZED D-REGION PERTURBATIONS ASSOCIATED WITH LIGHTNING

W. L. Poulsen, T. F. Bell, and U. S. Inan
Space, Telecommunications and Radioscience Laboratory
Stanford, California, 94305

91-09683

ABSTRACT

A theoretical model of multiple-mode, subionospheric VLF wave propagation in the presence of localized perturbations of the nighttime D-region has been developed. Such perturbations could be produced, for example, by lightning-induced electron precipitation associated with a characteristic type of phase or amplitude perturbation in VLF signals known as "Trimpi" events. Our model assumes that the ionospheric perturbation is slowly varying in the horizontal plane and that mode-coupling within the region of the perturbation is therefore negligible. However, the model does assume mode-coupling along the paths between the transmitter and receiver, the transmitter and the perturbation, and the perturbation and receiver. The model accounts for (1) effects of perturbations with finite extent in the dimension transverse to the great circle (GC) path between transmitter and receiver, and (2) effects of perturbations which lie off the GC path as well as on it. The formulation used for the numerical calculations depends significantly on the mode refractive index of the ambient earth-ionosphere waveguide and the mode refractive index in the region of the perturbation. In the calculations, values for the mode refractive index are determined from the electron-density-versus-altitude profiles of both the ambient and perturbed ionospheres. The results of varying the location of a perturbation along the GC path as well as off the path in the transverse dimension, varying the horizontal scale of the perturbation, and varying the vertical density profile of the perturbation are examined. Values for changes in the amplitude and phase of a received signal were obtained from the model and compared with amplitude and phase measurements of signals received at several receiver sites, operated by the VLF group of Stanford University's STAR Laboratory, from various VLF transmitters during precipitation events. Using realistic values for the ground and ionospheric profile parameters, values of the shift in the amplitude and phase of the signal similar to those measured on signals from several VLF transmitters were obtained using this model. For example, a cylindrically symmetric perturbation of 5 λ in horizontal extent due to a 0.2 second burst of precipitating electrons of $\sim 2 \times 10^{-3}$ ergs/cm^2-s flux density can produce amplitude changes of ~ -0.3 dB and phase changes of $\sim 2^\circ$ in a signal with a long-distance (800 λ), all-sea-based path between transmitter and receiver. Results from the model suggest that the ratio of the shifts in signal phase and amplitude can be used to determine the distance of the perturbation from the GC path. We conclude that the 3-D theory appears capable of explaining VLF signal measurements showing the "Trimpi" signature.

INTRODUCTION

A three-dimensional method of modeling VLF propagation in the Earth-ionosphere waveguide and how it is affected by the appearance of a localized, *D*-region disturbance has been developed. Such disturbances can be produced, for example, by lightning-induced electron precipitation (LEP) events, which in turn produce perturbations, known as "Trimpi" events, in the phase and/or amplitude of a received VLF signal [*Helliwell et al.*, 1973; *Lohrey and Kaiser*, 1979].

Figure 1 from *Poulsen et al.* [1990] shows the configuration of the problem from a side view and a top view representation. In two-dimensional models [*Tolstoy*, 1983; *Ferguson and Snyder*, 1987], the situation depicted in Figure 1*a* is assumed to be unvarying in the horizontal dimension transverse to the direction of propagation. Thus, a disturbance always occurs directly over the great circle (GC) path between the transmitter and the receiver and

91 9 12

extends to infinity on either side. Our 3-D model accounts for the fact that a typical disturbance is of finite extent in the transverse (as well as along the path) directions. It also accounts for effects of disturbances lying partially or completely off the GC path.

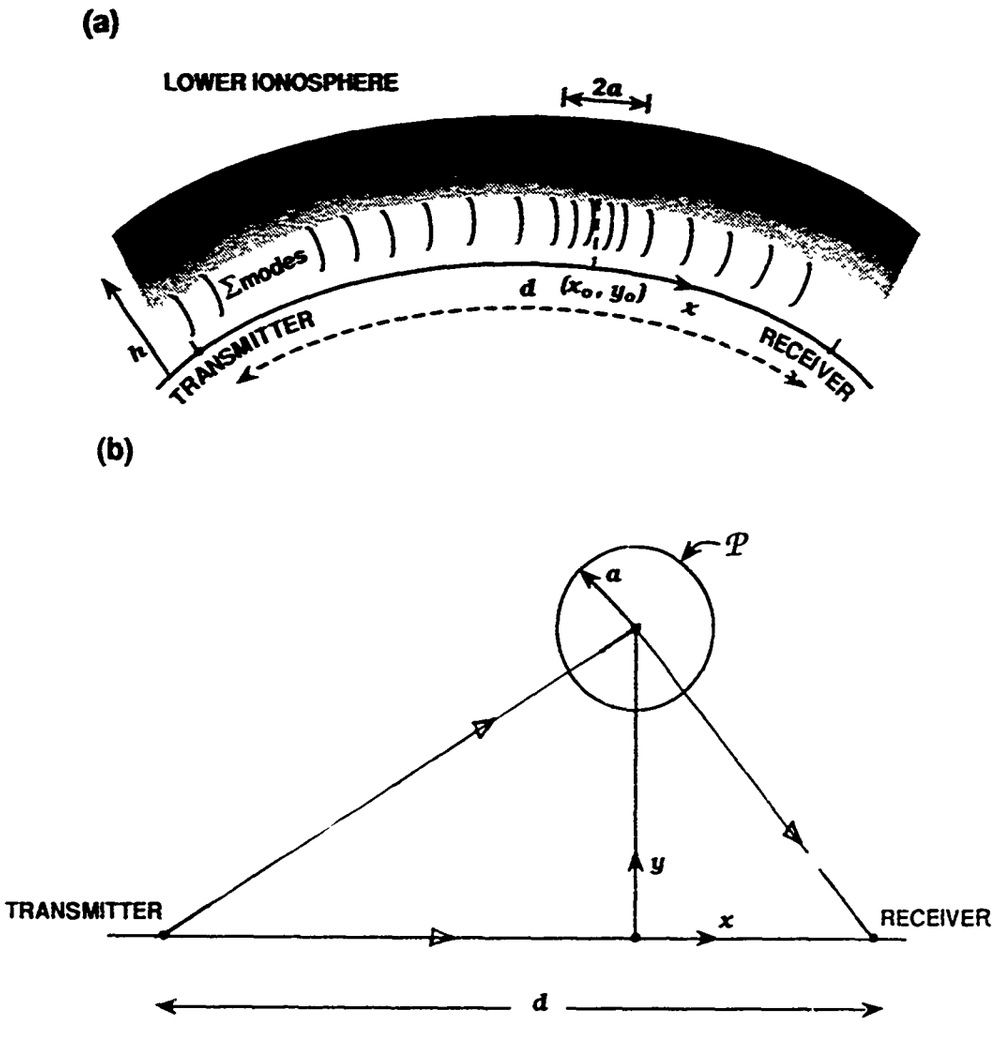


Figure 1. (*a*) Side view representation of the Earth-ionosphere waveguide between a transmitter and receiver separated by a distance d along the surface of the Earth. The change in electron density with altitude h of the lower ionosphere is represented by the change in shading density. Also represented is a density enhancement region or perturbation of the ambient ionosphere such as those generated by lightning-induced electron precipitation bursts. Such a perturbation, appearing transiently, scatters some of the signal impinging on it and causes a temporary perturbation in the total signal measured at the receiver. (*b*) A plan view, seen from above, of the situation depicted in Figure 1*a* showing the three-dimensional configuration of the problem.

For the case of a single dominant waveguide mode, it was found that [*Poulsen et al.*, 1990] the effect of a localized electron density enhancement in the vertical ionospheric density profile is important, i.e., that rather than treating the disturbance as a reflective scatterer, the actual density profile of the disturbance compared to the ambient density profile must be considered in determining the effect that a particular disturbance will have on the propagating VLF wave. Figure 2 shows a vertical density disturbance profile (as well as a typical ambient density profile) that might occur at the center of a disturbance produced by an LEP event at $L \sim 2.5$ [*Inan et al.*, 1988]. The height profile of enhanced ionization within the disturbed region is expected to be important in the multiple-mode case as well, especially since each mode will be scattered a different amount by the same disturbance.

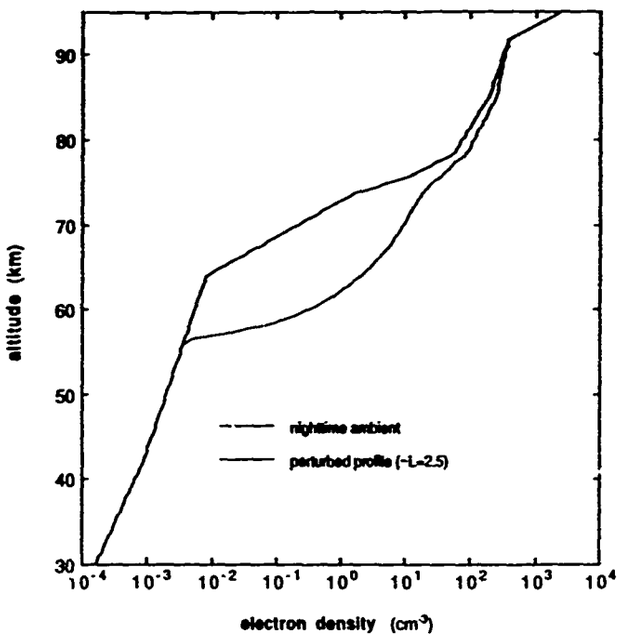


Figure 2. Plot of the electron density distribution versus altitude at the center of a disturbance produced by an LEP event induced by whistlers propagating at $L \sim 2.5$. A typical value of 200 ms for the duration of the lightning discharge and subsequent LEP burst has been assumed in the profile generation. The ambient density profile used in the study is also shown.

THE MODEL

The methodology used in the present 3-D model is depicted in Figure 3. The Long Wave Propagation Capability (LWPC) developed by the Naval Ocean Systems Center [*Pappert and Snyder*, 1972; *Ferguson and Snyder*, 1987] is used to calculate the complex signal strength of the wave as it travels along the direct path between the transmitter and receiver ('direct path'), and along the paths from the transmitter to the disturbance ('leg 1') and from the disturbance (or scatterer) to the receiver ('leg 2'). (The LWPC code incorporates a two-dimensional model to calculate signal amplitude and phase along the GC path between the transmitter and receiver and takes into account multiple-mode propagation, changes in ground conductivity along the path, and user-defined density-versus-altitude ionospheric profiles.) Meanwhile, the single-mode 3-D method described in *Poulsen et al.* [1990] is used to calculate the amount of signal scattered by the disturbance towards the receiver on a mode-by-mode basis, assuming that no conversion among modes occurs within the disturbed region. (This implies that a WKB approximation holds;

in other words, that the density profile is slowly varying over distances of $\sim 1\lambda$ in the horizontal directions.) The methodology can be summarized as follows:

1) LWPC is used along the GC path from the transmitter to the receiver to find the total electric field arriving along the 'direct path' (E_{direct}).
2) LWPC is used along 'leg 1' from the transmitter to the disturbance to find the value of the electric field for each mode n arriving at the disturbance region (e_n).
3) The signal strength scattered towards the receiver is calculated for each mode n, one mode at a time, based on the strengths of each of the e_n arriving at the region.
4) LWPC is then used along 'leg 2' from the disturbed region to the receiver, using the signal strength values calculated in the previous step as starting values for LWPC (rather than the usual antenna excitation values), to find the total electric field scattered by the disturbance which arrives at the receiver ($E_{scattered}$).
5) The direct and scattered electric fields are summed to obtain the total perturbed value of the electric field ($E_{total} = E_{direct} + E_{scattered}$).

The total perturbed value of the electric field at the receiver E_{total} is compared to the total ambient electric field at the receiver E_{direct} to calculate the change in amplitude and change in phase produced by the disturbance. These steps are repeated as the center location, the horizontal extent, and the vertical density profile of the disturbed region are varied. These values can then be compared with values of amplitude and phase perturbations measured on actual transmitter-receiver paths.

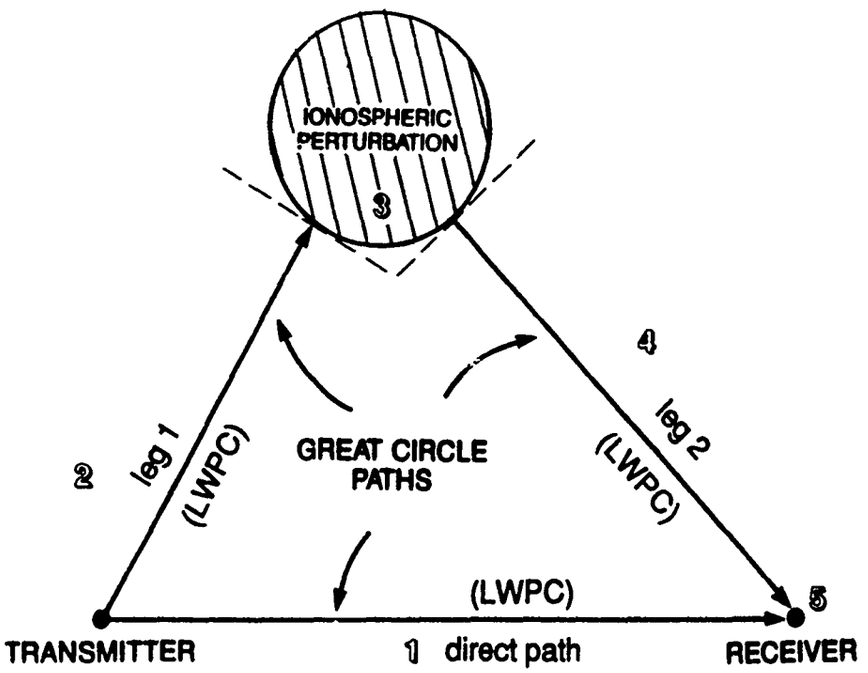


Figure 3. A depiction of the methodology used to calculate the total perturbed value of the electric field observed at the receiver. The numbers correspond to the numbered items given in the text. The Long Wave Propagation Capability (LWPC) is used along the three great circle paths shown, and the single-mode 3-D methodology explained in *Poulsen et al.* [1990] is used on a mode-by-mode basis within the ionospheric perturbation region. See text.

RESULTS

An example of the application of the new 3-D model is presented in Figures 4 and 5 for the propagation path between the NSS transmitter (Annapolis, Maryland) and Stanford University for one particular example of a disturbance. The center of the disturbed region for this case was assumed to have a vertical profile as shown in Figure 2. This was scaled back to the ambient profile values in the transverse directions proportional to a gaussian function to produce a cylindrically symmetric transverse profile with ~50 km radius. Figure 4 shows the amplitude and phase plots along the three 'legs' as discussed above. Notice that along the scattered path ('leg 2'), a larger number of modes of comparable amplitude (produced at the scatterer) result in a more complicated amplitude variation than along 'leg 1'. Also note that the amplitude scale for 'leg 2' is much smaller than for the other two legs since the scattered components for each mode are smaller than the incident ones for the disturbance profile used here (Figure 2).

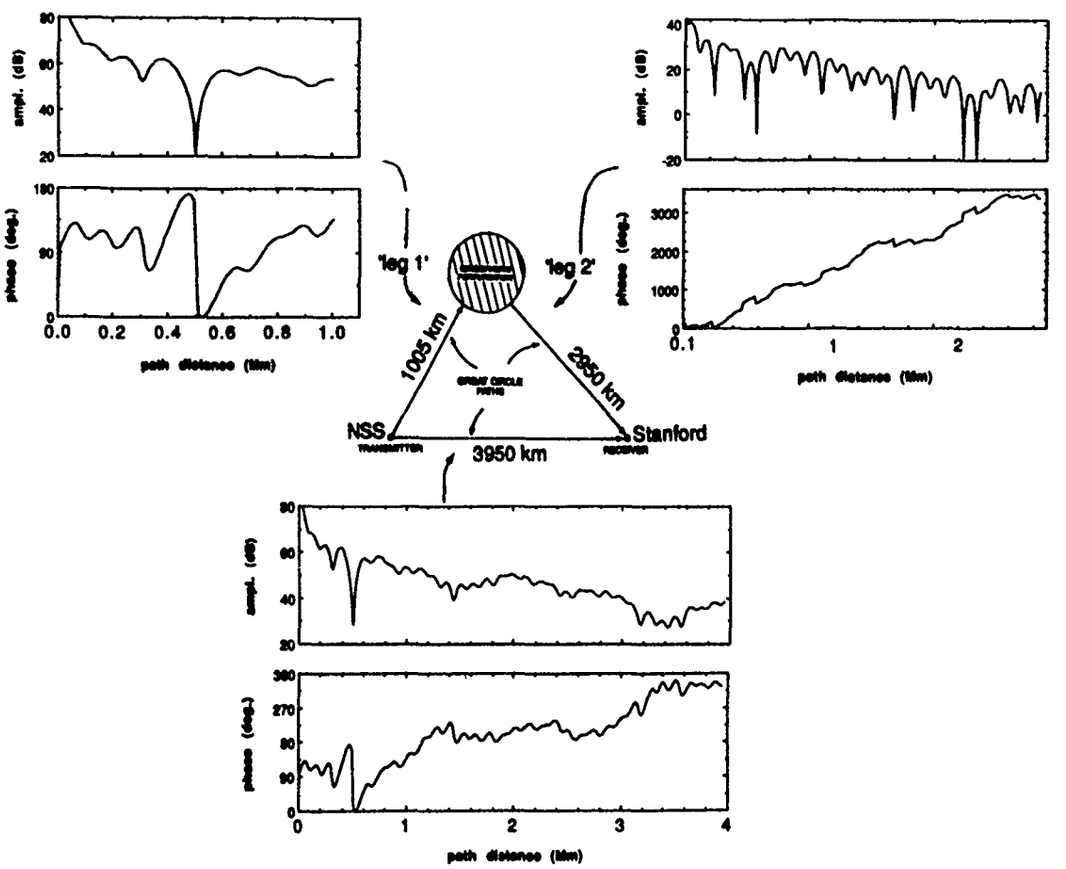


Figure 4. Amplitude and phase plots of a signal propagating along the three great circle paths shown for one particular example of a disturbed situation. The radius of the disturbance in this case is 50 km and the center of the disturbance is located 1000 km along the great circle path from NSS to Stanford and 100 km away from this 'direct path' on its north side. The lengths of the three paths are given in the center part of the figure (Note: not drawn to scale).

Figure 5 shows the relative signal strength and phase of each propagating mode at various points along the three paths of propagation for the specific example indicated (NSS to Stanford; disturbance of 50 km radius centered 100 km north of a point 1000 km along the GC path from NSS). The value in brackets beside each phasor diagram indicates the full-scale value of the relative signal strength (for that diagram). As was noted in the previous paragraph, the phasor diagrams show larger amplitudes of higher-order modes being scattered by the disturbance that subsequently propagate along 'leg 2'. Even at the receiver, there are substantial amounts of these higher-order modes still present. For this example the change in total received amplitude of the signal is ~0.12 dB, accompanied by a phase change of ~0.1 degrees. These values are within the range of measured amplitude and phase changes that occur on the NSS-Stanford path in LEP events [*Wolf*, 1990] (A bar graph showing the percentage of amplitude Trimpi events falling within spaced decibel ranges measured at Stanford on the NSS signal is given in Figure 5).

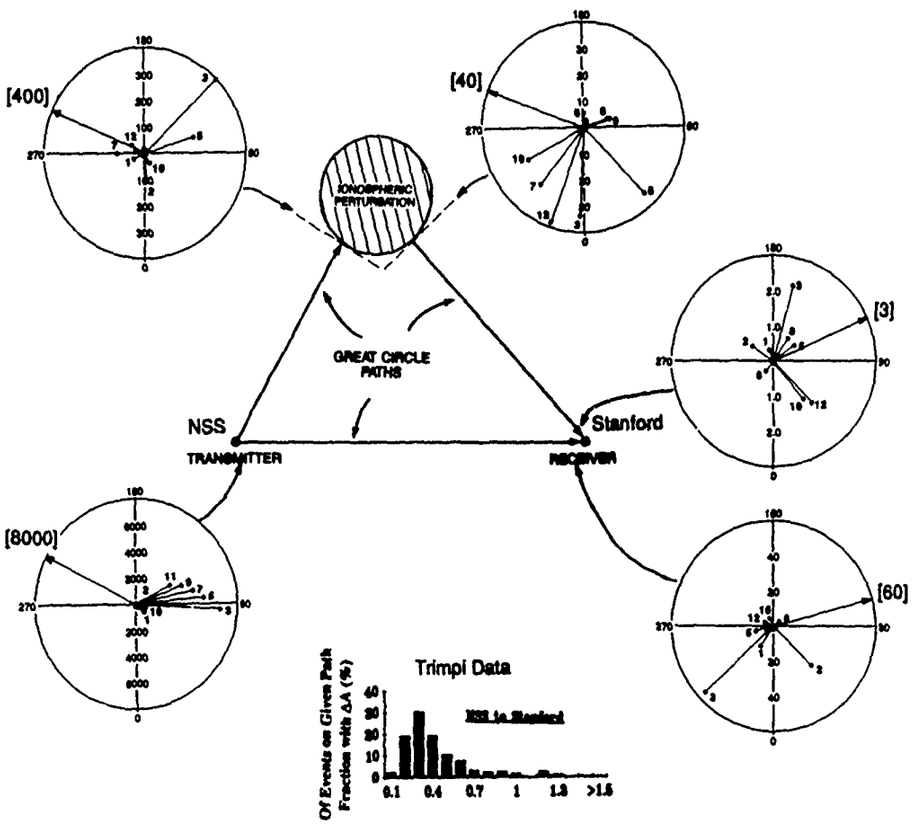


Figure 5. Phasor diagrams of the relative signal strength and phase of each mode of the propagating signal located at the points indicated by an arrow along the three propagation paths for the example described in Figure D. The value in brackets by each diagram indicates the full scale value of the relative signal strength for that diagram. Also plotted is a distribution bar graph showing the percentage of amplitude Trimpi events falling within spaced decibel ranges typically measured at Stanford on the NSS signal.

While the results in Figures 4 and 5 are for one particular size and location of the disturbance, the amplitude and phase changes at the receiver are expected to depend strongly on the location of the receiver as well as the disturbed region. This is illustrated in Figure 6, which shows an amplitude and phase plot of a two-dimensional (LWPC-only) calculation of the signal strength along the NSS to Stanford path for an ambient case and a case with a disturbance 100 km wide located at a point 3 Mm from NSS (and assumed to lie directly over the GC path and extend to infinity in the transverse directions). The amplitude and phase changes produced by the disturbance in this example change very little at some of the points along the path beyond the disturbance, while at other locations, changes larger than 10 dB may be observed. Similarly, for a fixed receiver point, some locations of the disturbance are expected to produce larger signal changes than other locations.

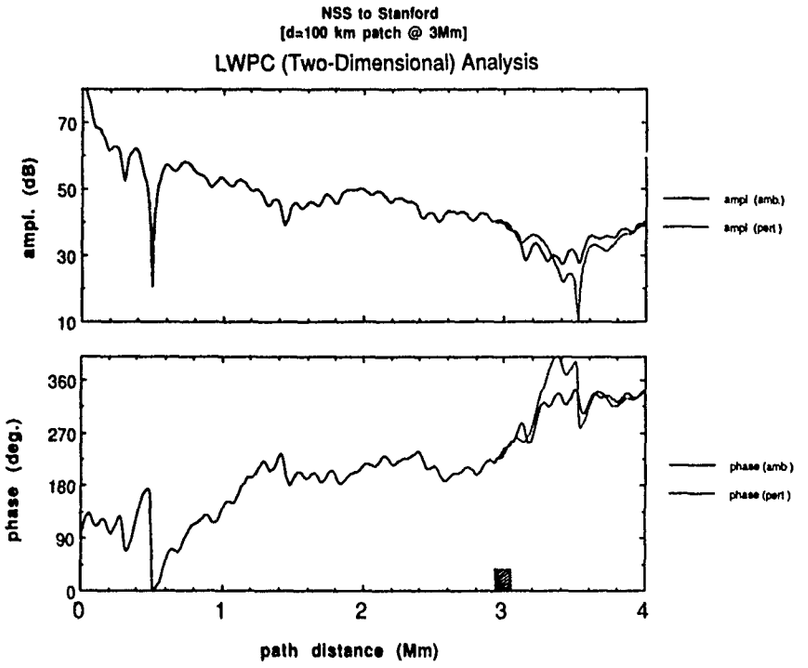


Figure 6. Amplitude and phase plots of a two-dimensional (LWPC-only) calculation of the signal strength along the NSS to Stanford path for an ambient case and a case with a disturbance 100 km wide centered at a point 3 Mm from NSS, having the disturbed electron density profile shown in Figure 2.

CONCLUSIONS

In summary, a new methodology to account for the effect of localized, off-great-circle path disturbances on propagating VLF waves has been developed. It utilizes the LWPC software developed by the Naval Ocean Systems Center and the results of *Poulsen et al.* [1990]. Preliminary results of the model are in agreement with observations of LEP event effects along typical paths. The model is applicable to a wide range of disturbance types in the lower ionosphere that can be modeled by a vertical density profile and that satisfy the WKB approximation. And this model can be generalized to account for mode coupling within the disturbed region itself (non-WKB cases).

REFERENCES

Ferguson, R. L., and F. P. Snyder, The segmented waveguide program for long wavelength propagation calculations, *Tech. Doc. 1071,* Naval Ocean Systems Center, San Diego, Calif., 1987.

Helliwell, R. A., J. P. Katsufrakis, and M. L. Trimpi, Whistler-induced amplitude perturbations in VLF propagation, *J. Geophys. Res., 78,* 4679, 1973.

Inan, U. S., and D. L. Carpenter, Lightning-induced electron precipitation events observed at $L \sim 2.4$ as phase and amplitude perturbations on subionospheric VLF signals, *J. Geophys. Res., 92,* 3293, 1987.

Inan, U. S., D. L. Carpenter, R. A. Helliwell and J. P. Katsufrakis, Subionospheric VLF/LF phase perturbations produced by lightning-whistler induced particle precipitation, *J. Geophys. Res., 90,* 7457, 1985.

Inan, U. S., W. C. Burgess, T. G. Wolf, D. C. Shafer, and R. E. Orville, Lightning-associated precipitation of MeV electrons from the inner radiation belt, *Geophys. Res. Lett., 15,* 172, 1988.

Lohrey, B., and A. B. Kaiser, Whistler-induced anomalies in VLF propagation, *J. Geophys. Res., 84,* 5121, 1979.

Pappert, R. A., and F. P. Snyder, Some results of a mode-conversion program for VLF, *Radio Sci., 7,* 913, 1972.

Poulsen, W. L., T. F. Bell, and U. S. Inan, Three-dimensional modeling of subionospheric VLF propagation in the presence of localized *D* region perturbations associated with lightning, *J. Geophys. Res., 95,* 2355, 1990.

Tolstoy, A. The influence of localized precipitation-induced D-region ionization enhancements on subionospheric VLF propagation, Ph.D thesis, U. of Maryland, 1983.

Tolstoy, A., and T. J. Rosenberg, A quasi three-dimensional propagation model for subionospherically propagating VLF radio waves, *Radio Sci., 20,* 535, 1985.

Tolstoy, A., T. J. Rosenberg, and D. L. Carpenter, The influence of localized precipitation-induced D-region ionization enhancements on subionsopheric VLF propagation, *Geophys. Res. Lett., 9,* 563, 1982.

Tolstoy, A., T. J. Rosenberg, U. S. Inan, and D. L. Carpenter, Model predictions of subionsopheric VLF signal perturbations resulting from localized, electron-precipitation induced ionization enhancement regions, *J. Geophys. Res., 91* 13473, 1986.

Wait, J. R., On phase changes in very low frequency propagation induced by an ionospheric depression of finite extent, *J. Geophys. Res., 69,* 441, 1964.

Wolf, T. G., Remote sensing of ionospheric effects associated with lightning using very low frequency radio signals, Ph.D. dissertation, Stanford Univ., Stanford, Calif., 1990

AD-P006 321

VLF/LF REMOTE SENSING OF TRANSIENT AND LOCALIZED DISTURBANCES OF THE LOWER IONOSPHERE ASSOCIATED WITH LIGHTNING DISCHARGES

Umran S. Inan
Space, Telecommunications and Radioscience Laboratory
Stanford, California, 94305
USA

ABSTRACT

High resolution measurements of subionospheric VLF/LF signals are now established as a powerful remote sensing tool for investigation of transient and localized ionospheric disturbances associated with atmospheric lightning discharges. Events usually consist of sudden (<2 s) changes in amplitude and/or phase followed by relatively slow (10-100 s) recovery to background levels. In this paper, we discuss recent results from a new series of observations conducted at Arecibo, Puerto Rico and at Huntsville, Alabama. The subionospheric VLF paths from two different transmitters to these two sites are nearly collinear, allowing for identification of the location of the disturbed ionospheric regions. Events observed on some very short paths indicate that selective excitation of waveguide modes can be used to extract information on the energy spectra of precipitating particles. In one case, on 23 February 1990, an extremely high rate of event activity (> 60 events/hr) was observed at Huntsville on the 48.5 kHz signal from Silver Creek, Nebraska, with amplitude changes of as much as 15 dB in individual events.

INTRODUCTION

A substantial body of evidence has recently emerged indicating that the nighttime lower ionosphere is regularly perturbed by energetic (>40 keV) electron bursts that are precipitated out of their trapped radiation belt orbits in gyroresonant interactions with lightning-generated whistlers. While electron precipitation bursts have been observed on satellites [*Voss et al.*, 1984] and on rockets [*Rycroft*, 1973; *Goldberg et al.*, 1986], the overwhelming evidence for the occurrence characteristics and geographic distribution of lightning-induced electron precipitation (LEP) events, and the causative association of the observed ionospheric signatures with whistlers and lightning, have been obtained by means of ground-based VLF remote sensing of the nighttime D-region [*Carpenter and LaBelle*, 1982; *Inan et al.*, 1985; *Inan and Carpenter*, 1986; 1987; *Carpenter and Inan*, 1987; *Inan et al.*,1988a,b,c; *Dowden and Adams*, 1989;1990]. Examples of ground-based signatures of LEP events are shown in Figure 1, together with the associated whistler data that was available for the case shown. The characteristic VLF signatures of these events are commonly referred to as *Trimpi* events and involve a rapid (< 2 s) onset followed by a relatively longer (10-100 s) recovery, representing the recombination and/or attachment of the enhanced secondary ionization produced by the energetic electrons. The typically < 1 s delay between the causative lightning (or radio atmospheric) represents the kinematics of the waves and particles to and from the high altitude interaction region, consistent with theoretical predictions of whistler-particle interaction theory [*Chang and Inan*, 1985]. The duration of the onset (< 2 s) is believed to represent the duration of the LEP burst, again consistent with theory [*Chang and Inan*, 1985]. The typical density profiles of ionospheric disturbances in events observed at mid latitudes ($2 < L < 3$) are given in another paper in these proceedings [*Poulsen et al.*, 1990].

In the past few years, Stanford University has developed a network of VLF/LF remote sensing observations in the northern hemisphere distributed across the continental United States and at Palmer Station, Antarctica, as shown in Figure 2. Observations at 5 locations in the north of signals from 6 different VLF transmitters (see Table 1) is used to monitor ionospheric disturbances occurring on or near a criss-crossing set of subionospheric paths, with the potential for 'imaging' the location and structure of the disturbed regions [*Inan et al.*,1990]. In this paper, we present some new observations from Huntsville (HU), Alabama and Arecibo (AR), Puerto Rico. As illustrated in Figure 2, the signal paths from the NLK (24.8 kHz) and 48.5 kHz transmitters to these sites are nearly collinear, allowing for identification of the location of the disturbed ionospheric regions. In addition, the HU site allows for observation of the 48.5 kHz signal on a relatively short path located near a very active thunderstorm region, leading to unusually high event rates (> 60 events/hr), unusually large (up to 15 dB) amplitude changes, and selective excitation of high order waveguide modes that may allow extraction of information on the altitude profile of the ionospheric disturbances.

91-09681

EXAMPLES OF EVENTS OBSERVED AT HUNTSVILLE AND ARECIBO

Figure 3 shows a 25-minute sequence of events observed on the 48.5-HU, NLK-HU, 48.5-AR, and NLK-AR signals. The NAU-HU signal is also shown in the middle panel, with no detectable event signatures. The characteristic perturbations on the NLK-HU and 48.5-HU signals are remarkably well correlated, which indicates that they are due to the same ionospheric disturbance. This fact alone indicates that the disturbed region is most likely located between the 48.5 kHz transmitter and HU. The lack of events on the NAU-HU path is consistent with this hypothesis. The events observed on the 48.5-AR signal are very well correlated with those on 48.5-HU, indicating that the former are due to perturbation of the dominant waveguide modes at the same disturbed region (located between 48.5 and HU). The relative strength of the perturbed modes have apparently not changed in the course of the signal propagation from HU to AR, although it is interesting to note that some of the smaller perturbations are better defined at AR than at HU. We note that the size of the signal perturbation observed at a given site is a function of receiver location with respect to signal maxima and minima along the great-circle propagation path [*Poulsen et al.*, 1990]. In a later example (Figure 6), we show how the responses as seen at HU and AR can be very different, due to the selective excitation of higher order waveguide modes that are observable at HU but that are attenuated between HU and AR. We note that the NLK-AR signal basically exhibits the same events, although this signal at AR is relatively weak and identification of events is difficult.

The association of observed events with ducted whistlers from the magnetosphere is shown in Figure 4. Here, the first two of the events from the uppermost panel of Figure 3 are shown in the middle panel at higher resolution together with the associated ducted whistlers (top panel) observed at Palmer Station, Antarctica, and the intensity in the 28.5±0.15 kHz channel (NAU-HU) which shows the causative radio atmospherics from lightning. For the case shown, the whistler data from Palmer was only available during synoptic periods of 1-minute out of every five. The two events appear to be associated with enhanced radio atmospherics (lower panel) as well as the whistlers, although there are other atmospherics which do not seem to have produced whistlers or VLF amplitude changes (on the 48.5-HU signal). These additional radio atmospherics could be due to distant lightning, relatively further away from conjugate point of Palmer Station as well as the 48.5-HU path.

Another example of high rate of event occurrence is shown in Figure 5. Here, up to 25 events (variety of sizes, with the smaller events being recognized as characteristic rapid decreases in intensity or interruption of a recovery from a previous event) can be counted in the 13 minute period shown. The correlation between the 48.5-HU, NLK-HU, 48.5-AR and NLK-AR signals is again remarkable and by itself indicates that all of the perturbations are likely to be due to the same ionospheric disturbance. The disturbed region is again likely to be located in the close vicinity of the 48.5-HU path, since the NAU-HU path (not shown) did not register any perturbations. The NSS-HU path (see Figure 2) registered the strongest two of the events in this period, probably due to scattering from a large disturbance even though it was not along the NSS-HU great-circle path [*Poulsen et al.*, 1990].

The events shown in Figure 5 are rather unusual in terms of the magnitude of the perturbations. In the case of the largest events on the 48.5-AR signal, the amplitude changes are $\sim$ 15 dB. Such large changes in signal intensity in individual events, as well as the average reduction in intensity during periods of high event rate as shown, may be important impediments to practical subionospheric VLF/LF communication and navigation channels. We also note that although phase data from HU is not yet available, such large amplitude changes are also likely to be accompanied by equally substantial changes in signal phase[*Wolf*, 1990; *Wolf and Inan*, 1990].

The potential use of subionospheric VLF data to extract information concerning the shape (altitude profile) of the secondary ionization in the disturbed region is shown in Figure 6. The top panel shows a series of events on the 48.5-HU signal consisting of negative amplitude changes, except for one event which exhibits a positive change in amplitude as well as an unusual recovery signature. The second and third panels show this event at two different resolutions, better illustrating the overall recovery as well as the sharp initial recovery portion that takes place within the first second after the amplitude change. Such exceptionally rapid recoveries have in the past been interpreted as being due to MeV electrons that penetrate to and create ionization at altitudes as low as 50-70 km [*Inan et al.*, 1988a]. The bottom two panels show the 48.5 kHz signal observed at Arecibo (AR), with striking differences in signature. The amplitude change at AR is negative, followed by a typical exponential recovery. The differences in signatures between AR and HU are interpreted to be due to the high energy (deep penetrating) component of the precipitation burst, which produces ionization at low altitudes and selectively perturbs the higher order waveguide modes. These modes add up to produce a positive amplitude change at HU. By the time the signal reaches AR, the higher order modes are attenuated and the signal consists of lower order modes which are not as sensitive to ionization at the lower altitudes. It appears that the diversity of waveguide modes available on these relatively short VLF paths may reveal such information concerning the energy spectrum of the precipitation, especially when used in connection with quantitative models of VLF propagation in the earth-ionosphere waveguide *Poulsen et al.*,1990].

SUMMARY

New VLF/LF data from Huntsville and Arecibo further illustrates that the nighttime lower ionosphere is regularly disturbed by lightning-induced electron precipitation bursts from the magnetosphere. Observations at multiple points along collinear signal paths allow the identification of the location of the disturbed regions and may also reveal new information concerning the altitude profile of the disturbances. The event rates on paths in the vicinity of active thunderstorm centers can be > 60 per hour, with amplitude changes of up to 15 dB in individual events. The potential effects on VLF/LF subionospheric communications are yet to be evaluated, but are likely to be important in view of the size of the amplitude changes and the reduction in the average signal level in cases of high event rate.

TABLE 1. VLF/LF Transmitters.

Call Sign	Transmitter	Frequency	Latitude	Longitude
NSS	USN Maryland	21.4 kHz	39°N	76°W
NPM	USN Hawaii	23.4 kHz	21°N	158°W
NAU	USN Puerto Rico	28.5 kHz	18°N	67°W
NAA	USN Maine	24.0 kHz	45°N	67°W
NLK	USN Washington	24.8 kHz	48°N	122°W
	USAF Nebraska	48.5 kHz	42°N	98°W

REFERENCES

1. Carpenter, D.L., and J. W. LaBelle, A study of whistlers correlated with bursts of electron precipitation near $L = 2$, *J. Geophys. Res., 84,* 4427, 1982.

2. Carpenter, D.L., and U.S. Inan, Seasonal, Latitudinal and Diurnal distribuiton of whistler-induced precipitation events at low latitudes, *J. Geophys. Res., 92,* 3429, 1987.

3. Chang, H.C., and U.S. Inan, Lightning-induced electron precipitation from the magnetosphere, *J. Geophys. Res., 90,* 1531, 1985.

4. Dowden, R., and C.D.D. Adams, Modal effects on amplitude perturbations on subionospheric signals (Trimpis) deduced from two-frequency measurements, *J. Geophys. Res., 94,* 1515, 1989.

5. Dowden, R.L., and C.D.D. Adams, Location of lightning-induced electron precipitation from measurement of VLF phase and amplitude perturbations on spaced antennas and on two frequencies, *J. Geophys. Res., 95,* 4135, 1990.

6. Goldberg, R.A., J.R. Barcus, L.C. Hale, and S.A. Curtis, Direct observation of magnetospheric electron precipitation stimulated by lightning, *J. Atmos. terr. Phys., 48* 293, 1986.

7. Inan, U.S., H.C. Chang, R.A. Helliwell, J.P. Katsufrakis, W.L. Imhof, J.B. Reagan, M. Walt, D.w. Datlowe, and J. Mobilia, Wave-induced precipitation as a loss process for radition belt particles, *Adv. Space Rev., 5,* 243, 1985.

8. Inan, U.S., and D.L. Carpenter, On the correlation of whistlers and associated subionospheric VLF/LF perturbations, *J. Geophys. Res., 91,* 3106, 1986.

9. Inan, U.S., and D.L. Carpenter, Lightning-induced electron precipitation events observed at $L \simeq 2.4$ as phase and amplitude perturbations on subionospheric VLF signals, *J. Geophys. Res., 92,* 3293, 1987.

10. Inan, U.S., W.C. Burgess, T.G. Wolf, D.C., Shafer, and R.E. Orville, Lightning-associated precipitation of MeV electrons from the inner radiation belt, *Geophys Res Lett. 15,* 172, 1988a.

11. Inan, U.S., T.G. Wolf, and D.L. Carpenter, Geographic distribution of lightning induced electron precipitation observed as VLF/LF perturbation events, *J. Geophys. Res., 93,* 9841, 1988b.

12. Inan, U.S., D.C. Shafer, W.Y. Yip, and R.E. Orville, Subionospheric VLF signatures of nighttime D-region perturbations in the vicinity of lightning discharges, *J. Geophys. Res., 93,* 11455, 1988c.

13. Inan, U.S., F.A. Knifsend, and J.Oh, Subionospheric VLF 'Imaging' of lightning-induced electron precipitation from the magnetosphere, *J. Geophys. Res.,* (in press) 1990.

14. Poulsen, W. L., T. F. Bell, and U. S. Inan, Three-dimensional modeling of subionospheric VLF propagation in the presence of localized *D* region perturbations associated with lightning, *J. Geophys. Res.,95,* 2355, 1990.

15. Rycroft, M.J., Enhanced energetic electron intensities at 100 km altitude and a whistler propagating through the plasmasphere, *Planet. space Sci., 21,* 239, 1973.

16. Voss, H.D., W.L. Imhof, J. Mobilia, E.E. Gaines, M. Walt, U.S. Inan, R.A. Helliwell, D.L. Carpenter, J.P. Katsufrakis, H.C. Chang, Lightning induced electron precipitation, *Nature, 312,* 740, 1984.

17. Wolf, T. G., Remote sensing of ionospheric effects associated with lightning using very low frequency radio signals, Ph.D. dissertation, Stanford Univ., Stanford, Calif., 1990

18. Wolf, T.G., and U.S. Inan, Path dependent properties of subionospheric VLF amplitude and phase perturbations associated with lightning, *J. Geophys. Res.,* (in review) 1990.

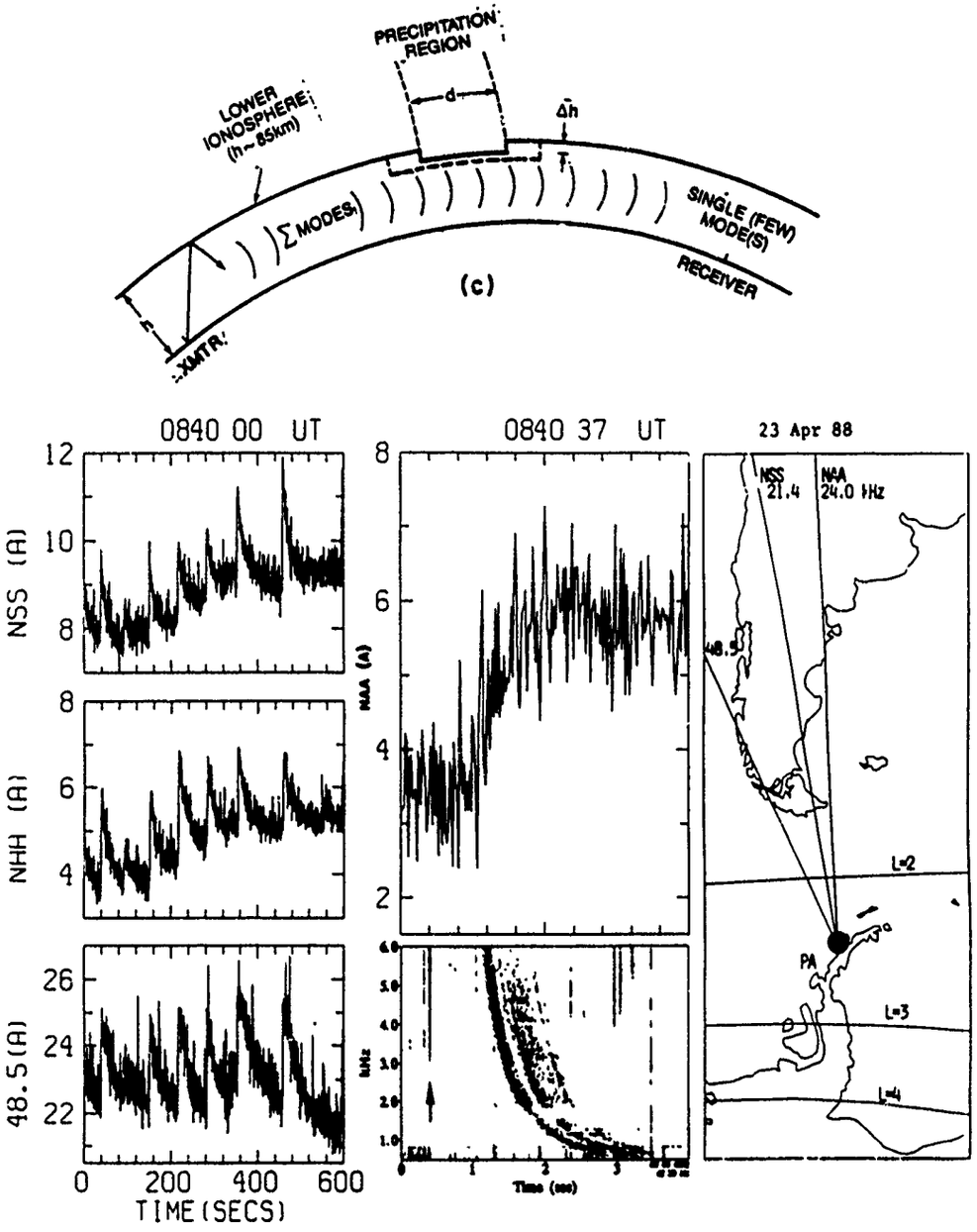


Figure 1. Illustration of the subionospheric VLF remote sensing technique as utilized at Palmer Station, Antarctica. the top panel shows a schematic description of a localized perturbation in the lower ionosphere which affects the phase and amplitude of VLF signals propagating in the earth-ionosphere waveguide. The lower panel on the right shows the observation geometry at Palmer Station with three signal paths arriving from distant transmitters in the northern hemisphere are shown. The left hand panels show the measured amplitude of these three signals (displayed on a linear amplitude (A) scale with $A = 0$ indicatring the absence of signal). The vertical scale shows signal amplitude A on a linear scale with arbitrary units, with $A = 0$ representing absence of signal. The time $T = 0$ corresponds to the UT time given in the upper right corner. The characteristic perturbations associated with electron precipitation bursts are recognized readily on the basis of their rapid (< 1 s) onset followed by relatively slow (10-100 s) recovery to unperturbed signal levels. In the case shown, a sepuence of at least 8 events are observed in succession on all three signals. The middle panel shows a high resolution record of one of the events, together with the causative magnetospheric whistler also observed at Palmer. The arrow indicates the arrival time of the causative atmospheric generated by the lightning discharge that launched the whistler.

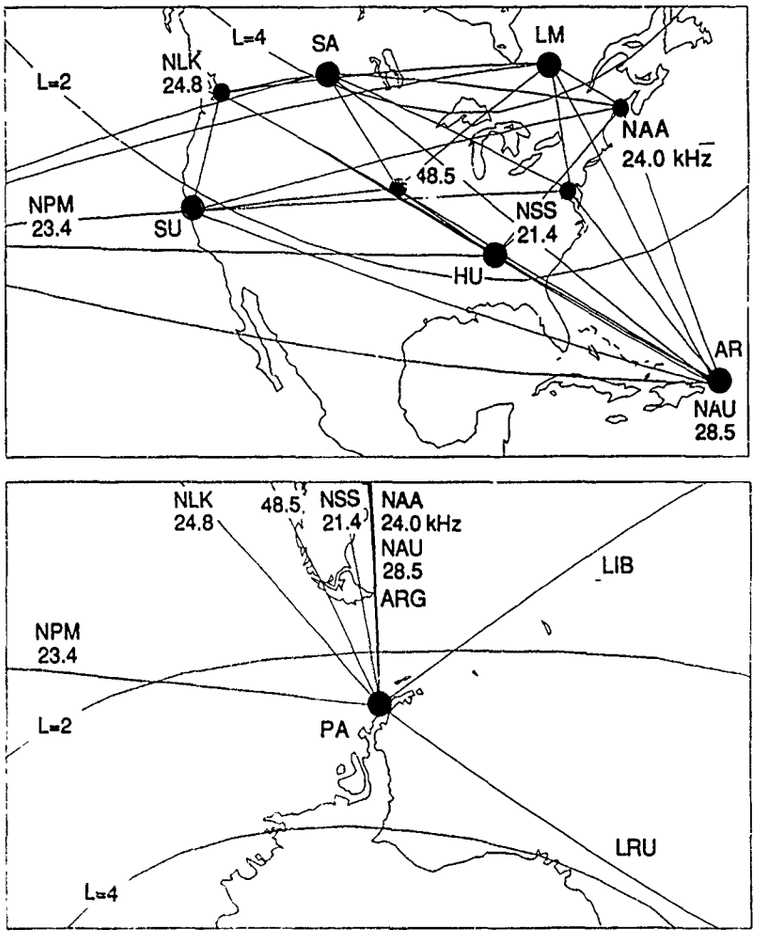


Figure 2. Stanford University VLF network observation sites and signal paths. The sites are Saskatoon (SA), Lake Mistissini (LM), Quebec, Arecibo (AR), Puerto Rico, Stanford (SU), California, and Huntsville (HU), Alabama. At each site signals from multiple transmitters are monitored and amplitude and phase information is stored with 50/20 Hz resolution. Recordings are generally made for the 0000-1200 UT period where most of the monitored paths are propagating under a nighttime ionosphere. Summary plot data is produced at the end of each night.

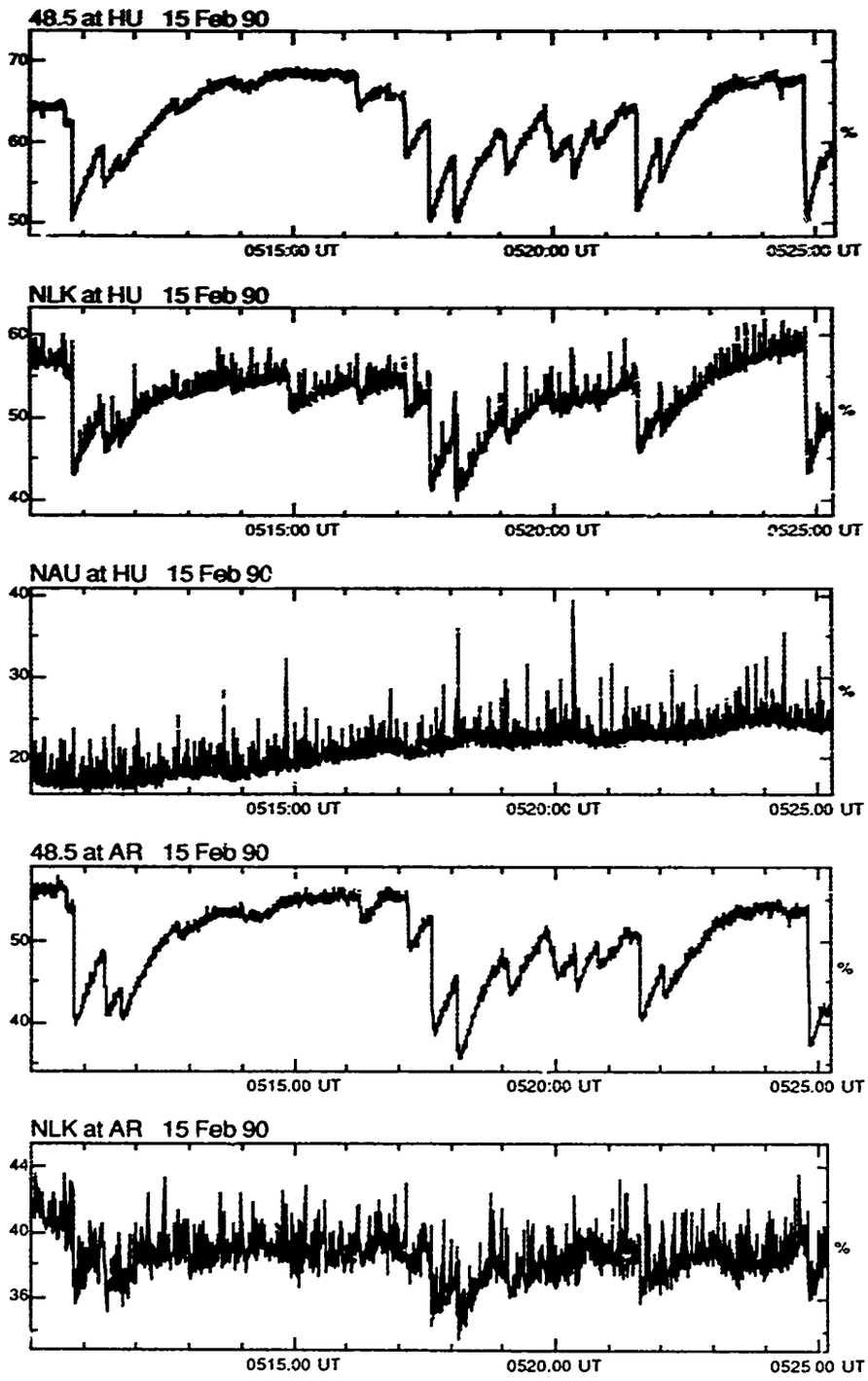


Figure 3. Simultaneous observations of lightning associated perturbation events at Huntsville (HU) and Arecibo (AR) on the 48.5 kHz and NLK (24.8 kHz) signals (see Figure 2). The disturbed ionospheric region is interpreted to be located somewhere near the 48.5-HU path, since the other paths such as NAU-HU (shown in the middle panel) did not register these events. The vertical scales shown linear amplitude (A) in linear arbitrary units, with A = 0 representing the absence of signal.

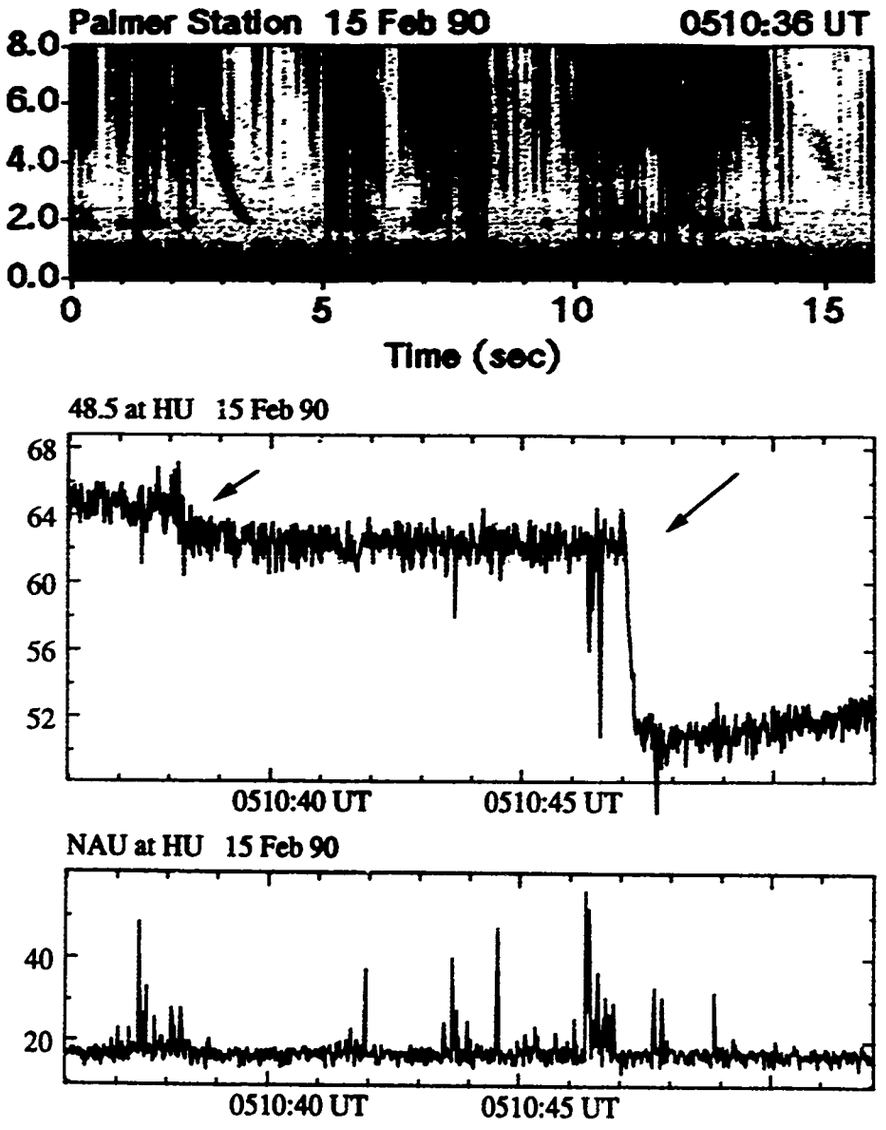


Figure 4. Whistlers observed at Palmer correlated with VLF perturbations on the 48.5 kHz signal observed at Huntsville (HU). The middle panel shows the first two events from the uppermost panel of Figure 3 on an expanded scale. The top panel shows the broadband (0-8 kHz) spectra observed at Palmer showing two whistlers in succession, time correlated with two negative amplitude changes on the 48.5-HU signal. The broadband data was only available on a synoptic basis (1 minute of every five). The lower panel shows the signal intensity in the 28.5±0.15 kHz channel (NAU-HU channel) as a measure of the impulsive radio atmospherics from lightning.

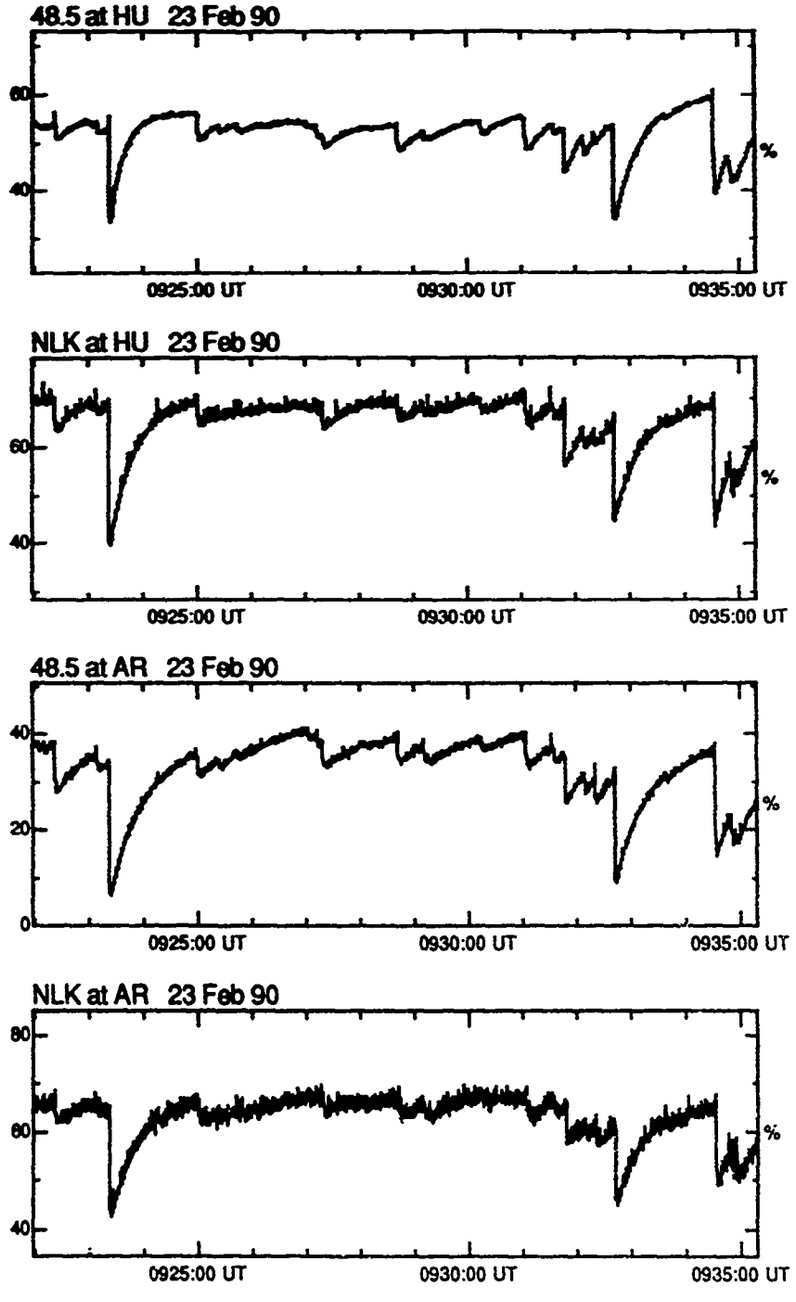


Figure 5. Simultaneous observations of lightning associated perturbation events at Huntsville (HU) and Alabama (AR) on the 48.5 kHz and NLK (24.8 kHz) signals (see in Figure 2). The disturbed ionospheric region is interpreted to be located somewhere near the 48.5-HU path, since the other paths such as NAU-HU did not register these events. The format of the amplitude records is similar to those of Figure 3. In terms of the effects of lightning-induced disturbances on subionospheric VLF/LF communication, it is important to note that the events shown represent very large (up to 15 dB in the case of the 48.5-AR signal change near 0923:20 UT) changes in signal amplitude.

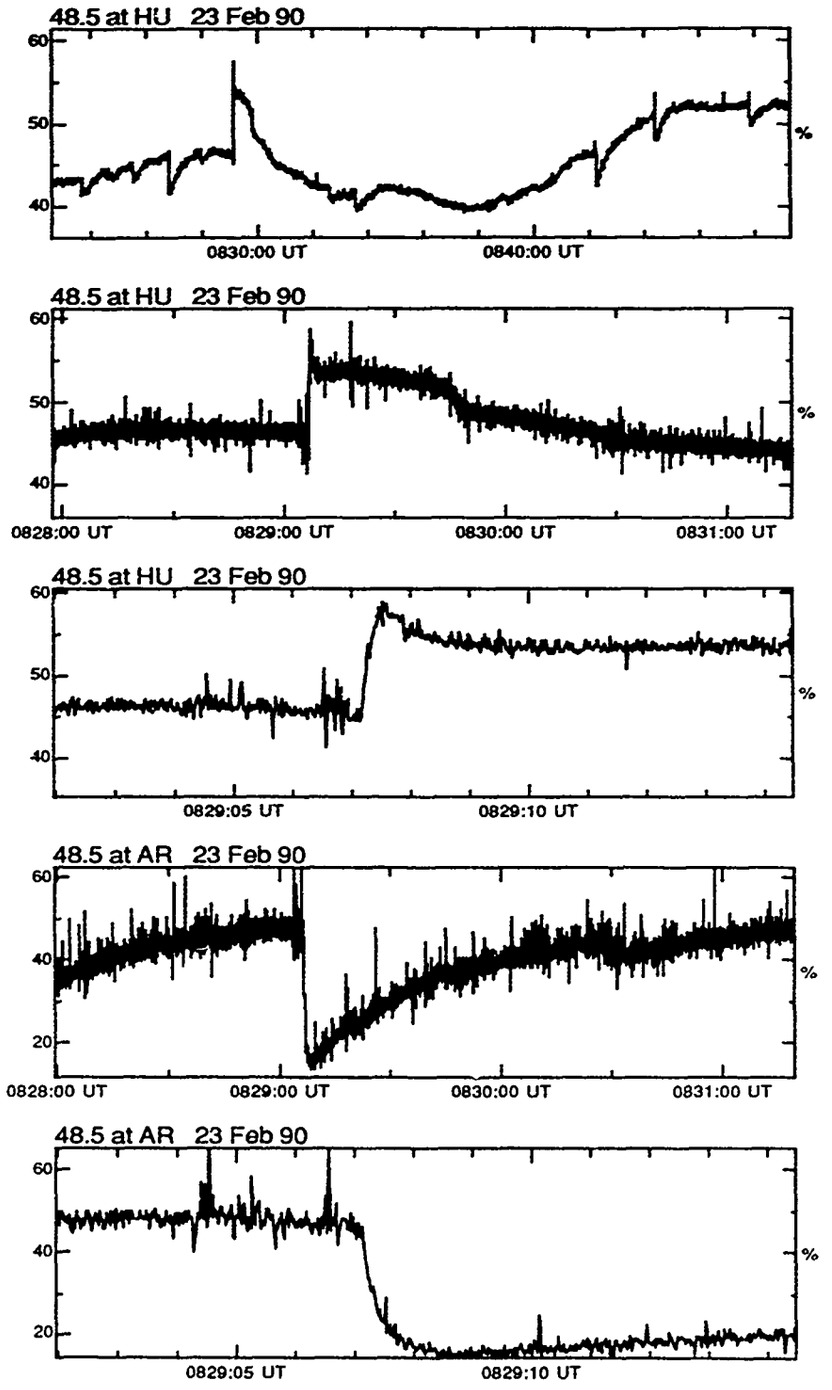


Figure 6. Possible evidence of selective excitation of different earth-ionosphere waveguide modes by an unusually large ionospheric disturbance. The top three panels show the 48.5 signal at Huntsville (HU) illustrating an unusual event at different resolutions. The format of the amplitude records is similar to that for Figure 3. The illustrated event is a positive amplitude change occurring in the midst of smaller amplitude decreases and also exhibits a distinctly different recovery signature. The 48.5 kHz signal observed at Arecibo (AR) does not display any of the unusual features.

AD-P006 322

OBSERVATIONS OF THE CHARACTERISTICS OF PROPAGATION OF VLF SIGNALS DURING MERIDIAN TRANSITS BY SCO X-1

J.H. BEALL*
K.S. WOOD
F.J. KELLY

E.O. Hulburt Center for Space Research
Naval Research Laboratory
Wasington, D.C. 20375

* also at St. John's College, Annapolis, Md. 21404 and
SFA, Landover, MD. 20785

ABSTRACT

We have conducted a series of studies of the VLF propagation characteristics during coordinated radio, optical, ultraviolet, and X-ray observations of the X-ray star Sco X-1. The observations were taken from 9-12 March 1989, a period when Sco X-1 was particularly active, and during a time of intense day-side solar activity. VLF sites at the Naval Research Laboratory (NRL) and the Chesapeake Bay Division (CBD) of NRL were used to obtain phase and amplitude data from the VLF transmitter at Annapolis (21.4 kHz) and from more remote transmitters. In this paper we give preliminary results which compare the time-varying X-ray flux with the VLF phase data taken during the observation.

INTRODUCTION

The nature of the ionosphere's response to celestial sources of radiation is of interest on scientific and practical levels ranging from communication to the basic physics of the response of a highly non-linear system to perturbations. The response of the ionosphere to solar radiation is well documented. The effect of celestial X-ray and ultraviolet fluxes is less intuitively obvious. However, a perusal of the technical literature uncovers a number of results which lend credence to the supposition that such an effect exists.

Several groups have claimed to detect the effect of X-ray fluxes from celestial (non-solar) sources upon ionization rates in the ionosphere and its consequent effects on radio propagation.

Anantakrishnan and Ramanathan (1969) have reported data which show a decrease in the field strength of 164 kHz radio waves which could be correlated with meridian transits of Sco X-1. However, Poppoff and Whitten (1969) have called this result into question by means of a theoretical analysis which shows that the effect of the ionizing radiation from Sco X-1 would be nominal compared to the ionizing rate from local sources. Svennesson, Reder, and Crouchley (1972) revisited this issue with an analysis of the effects of three sources (Sco X-1, Cen X-2, and Cen X-4) on VLF (Very Low Frequency) signals. They concluded that measurable effects do exist, based on a monthly comparison of the average diurnal phase shifts of the VLF signals. Phase shifts of the order of a few degrees were reported.

Direct experimental evidence of ion density variations from Sco X-1 at elevations of 60 to 80 km have been reported by Goldberg, Barcus, and Mitchell (1985) via rocket flights. The ion density and conductivity measurements were interspersed with measurements of the flux from Sco X-1. In addition, Fishman and Inan (1988) have reported a significant depression in the VLF amplitude signal from a receiver station in

91 9 4 122

91-09680

Antarctica coincident with a γ-ray burst on 1 August 1983 at 22:14:18 UT.

The results cited above suggest the need for further study of the degree to which celestial X-ray sources produce night-side ionization at a level detectable by VLF reflection measurements. It is precisely this issue that has motivated the current study.

In this paper we report the first coordinated observations of VLF phase and amplitude data with satellite observations of a celestial X-ray source during meridian transits by the source of the receiver sites. The X-ray observations where conducted by the Ginga satellite as part of a world-wide coordinated campaign to monitor the source at frequencies from radio to X-ray. The ground based VLF data were taken at the Naval Research Laboratory at both the main facility in Washington, D.C. (hereafter referred to as NRL) and the Chesapeake Bay Division (CBD). The data in this study can detect variations in amplitude for frequencies from 2.5 to 50 mHz, and thus extend the analysis of the ionospheric effects of X-ray stars to substantially different time scales than previous studies.

The analysis we present is a preliminary one and establishes the sensitivity of the technique described herein to measurements of the effect on VLF propagation at 21.4 KHz to celestial X-ray sources such as Sco X-1.

OBSERVATIONS

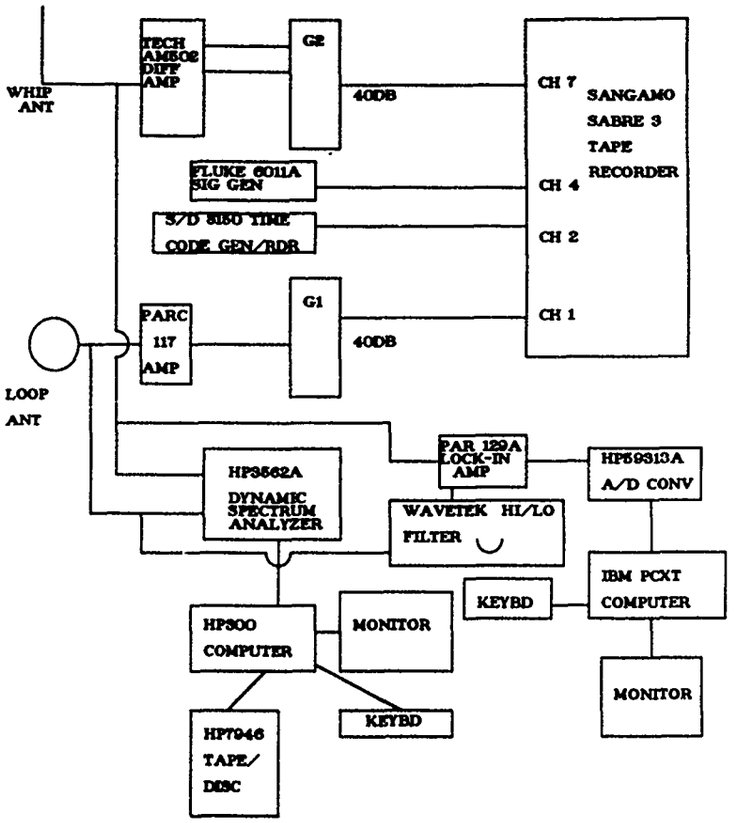
The VLF (20.4 kHz) phase and amplitude data from NRL and CBD were taken to coincide with a multiple-wavelength campaign to observe Sco X-1 on 10-12 March 1989.

The multi-wavelength campaign was conducted by an international collaboration which was able to obtain coverage in the X-ray (Ginga, Kvant/MIR), ultraviolet (IUE), optical (Wise, McDonald, ESO, Tokyo/Dodaira), and radio (VLA, Westerbork, Australia) bands. The multi-national campaign included observers and collaborators from 14 institutions. The list of collaborators includes T. Augusteijn (ESO), J.H. Beall, P. Hertz, F.J. Kelly, J.P. Norris, and K.S. Wood (NRL), N. Brosch and E. Leibowitz (Wise Obs.), T. Dotani and K. Mitsuda (ISAS), G. Hammerschlag-Hennsberge, W. Penninx, and J. van Paradijs (Amsterdam), G. Hasinger (MPE), R. Hjellming (NRAO), K. Horne (STScI), S. Kikuchi (Tokyo Obs.), W.H.G. Lewin (MIT), P.F. Michelson and B.A. Vaughan (Stanford), J.C. Raymond and S.D. Vrtilek (CfA), R. Stewart (Australian NRAO), and R. Strom (NFRA).

The spectrum of Sco X-1 has its greatest energy in the X-ray band. For this reason, and because that energy band produces ionization at the height of the D-layer (Popoff and Whitten 1969), we have chosen to concentrate our analysis on these data.

During the observations, the Ginga Large Area Counter (LAC) was configured in several different modes. The two data formats used were MPC mode, which gives spectra in 12 spectral channels with various time resolutions and gain settings, and the PC mode, which gives high time resolution in two spectral channels (1 ms resolution for low energy and 2 ms resolution for high energy). The data from the low gain setting have energy channels between 1-6 keV and 6-16 keV, while for photometric studies, the data have been binned into four energy channels. We have used the two highest "photometric" channels for the VLF data comparisons, so that the data we present here represent X-ray energies from 5.8 to 18.4 keV.

The VLF data have been used in numerous studies as a measure of the reflection coefficient of the lower ionosphere (see, e.g., Watt 1967, Best, Ratcliffe, and Wilkes, 1936, Bracewell et al 1951, Kelly and Rao, 1981, and Jones, Davies, and Weider, 1972). The data were taken with the experimental apparatus shown in Figure 1. The loop antenna is oriented so that the ground wave magnetic field from a local VLF transmitter is nulled out. The residual wave



VLF EQUIPMENT AT NRL CHESAPEAKE BAY DIVISION

MARCH 10,1989 X-RAY STAR EFFECTS MEASUREMENT

Figure 1. Configuration of VLF receiver site for Sco X-1 coordinated observations.

sensed by the loop antenna arises from a polarization conversion at the anisotropic reflection surface of the D-region ionosphere at a point midway between the transmitter and receiver sites. The intensity of this wave is proportional to the cross polarized reflection coefficient of the ionosphere, and the phase of the wave is dependent on the effective height of the reflecting "surface." By using a second (whip) antenna to pick up the dominant ground wave field from the transmitter, it is possible to detect any change in the effective height and reflection coefficient.of

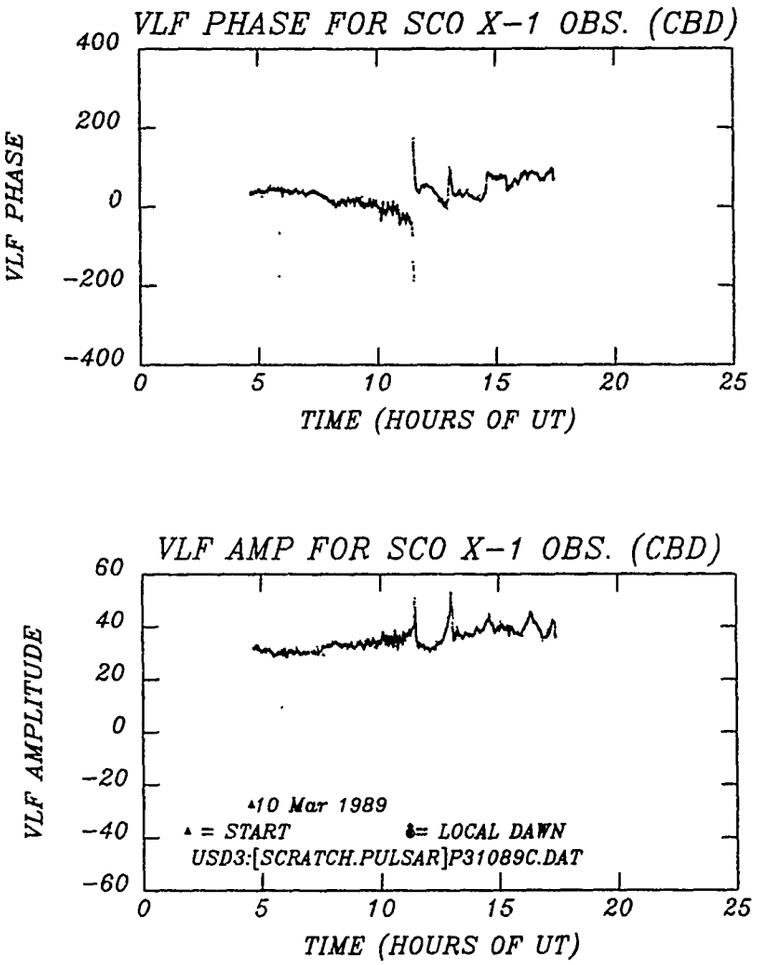


Figure 2. VLF phase and amplitude data taken on 10 March 1989 from the CBD site. Sco X-1 rose at 05:05 UT and transited the meridian at 10:17 UT.

the ionosphere. In the present experiment the signals from the whip antenna and the loop were sent into dual-channel HP3562A dynamic signal analyzers. The relative signal amplitude ratios and phase differences between the two channels were then digitally recorded. As an additional precaution against equipment failure at the remote site, the signals were recorded on an analog Sangamo tape recorder.

Since the VLF signal propagation paths are essentially local (from a transmitter in Annapolis, MD to two near by receivers at NRL and CBD), the times of interest for comparison with the X-ray data are times from local "Sco-rise" until the time of local dawn. At NRL and CBD (77°4.0' W. Longitude and 38°55.3' N. Latitude) Sco X-1 rose on 10 March 1989 at approximately 05:05 UT at an azimuthal angle of 109.65° and transited the meridian at 10:17 UT, reaching an elevation above the horizon of 35.9°. Local dawn occurs at approximately 13:00 UT, but the effect of the solar flux striking the ionosphere can be clearly seen in both the phase and amplitude from the CBD data, beginning at 11:00 UT (see Figure 2).

In Figure 3 we show the X-ray light curve for Sco X-1 during the Ginga observations for the time interval from 05:00 through 12:00 UT on 10 March 89, and the VLF phase data from both CBD and NRL. Gaps in the data are due to earth occultations and incomplete telemetry coverage. These data are shown on the same time scale as the 10 March 89 VLF data taken from CBD.

Sco X-1 is highly variable during the 10 March 89 observations on time scales from 0.1 to 10^3 seconds. Because the response of the lower ionosphere to perturbations from X-ray fluxes is likely to be on the order of a second or longer (Fishman and Inan 1988), and because the sampling interval of the VLF data is every 10 s, we have binned the X-ray data on that time scale.

The data from Sco X-1 are aperiodic. It is therefore not possible to "fold" the VLF data at a known period of the X-ray star in order to detect phase shifts on that time scale by integrating long stretches of data.

We have therefore used the flares in the Sco X-1 data to determine the times when we begin to bin the VLF data, allowing the binning to continue until another flare triggers a "reset" of the interval. The data are divided into phase bins which are normalized to a constant "exposure" - the amount of time spent by the data in the bin window. Thus, if each X-ray flare temporarily reduces the amplitude of the VLF signal (as did the Fishman and Inan γ-ray burst), that depression will be seen as a decrease in the appropriate phase bin. Failure to find such a signature can be converted into an upper limit for the response.

For the purposes of our analysis, the X-ray data are assumed to consist of a series of pulses of the same (possibly complex) shape but of varying amplitudes which occur at unknown intervals. This pulse shape is extracted from the data by an iterative method (Scargle 1981), and the residue of that deconvolution determines the

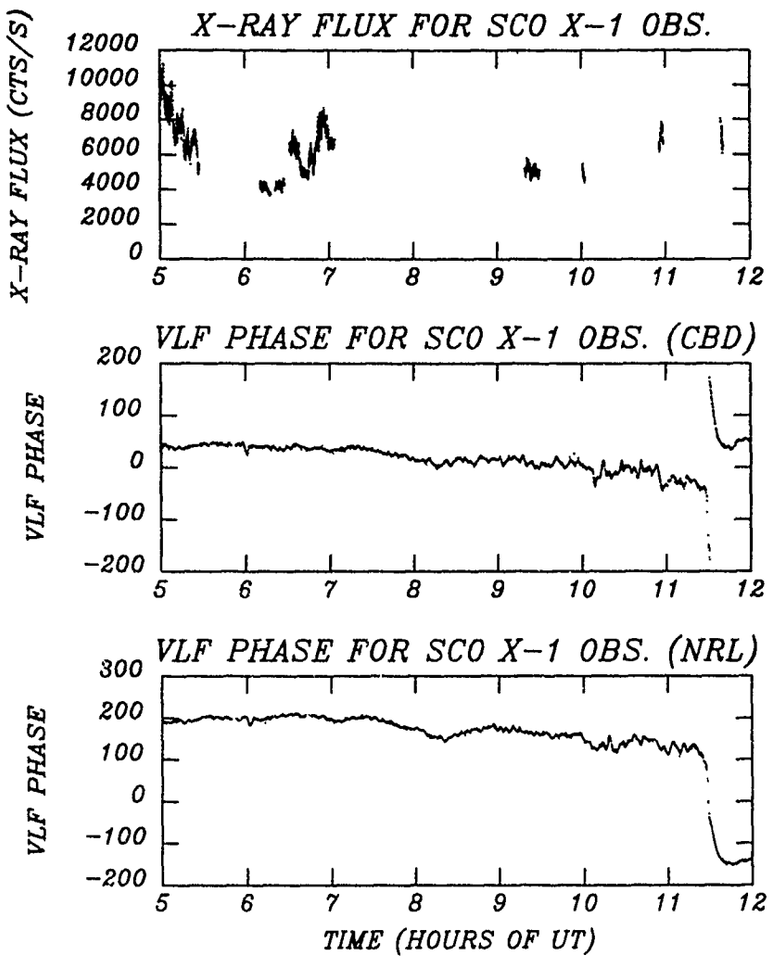


Figure 3. X-ray flux from Sco X-1 during the 10 March 89 observations and comparable trace from the VLF data taken from CBD and NRL.

amplitudes of the pulses and the times the pulses occurred. The pulse shape extracted by the algorithm from the data taken between 05:00 to 05:30 UT is shown in Figure 4, while the fit of the data by pulse shape

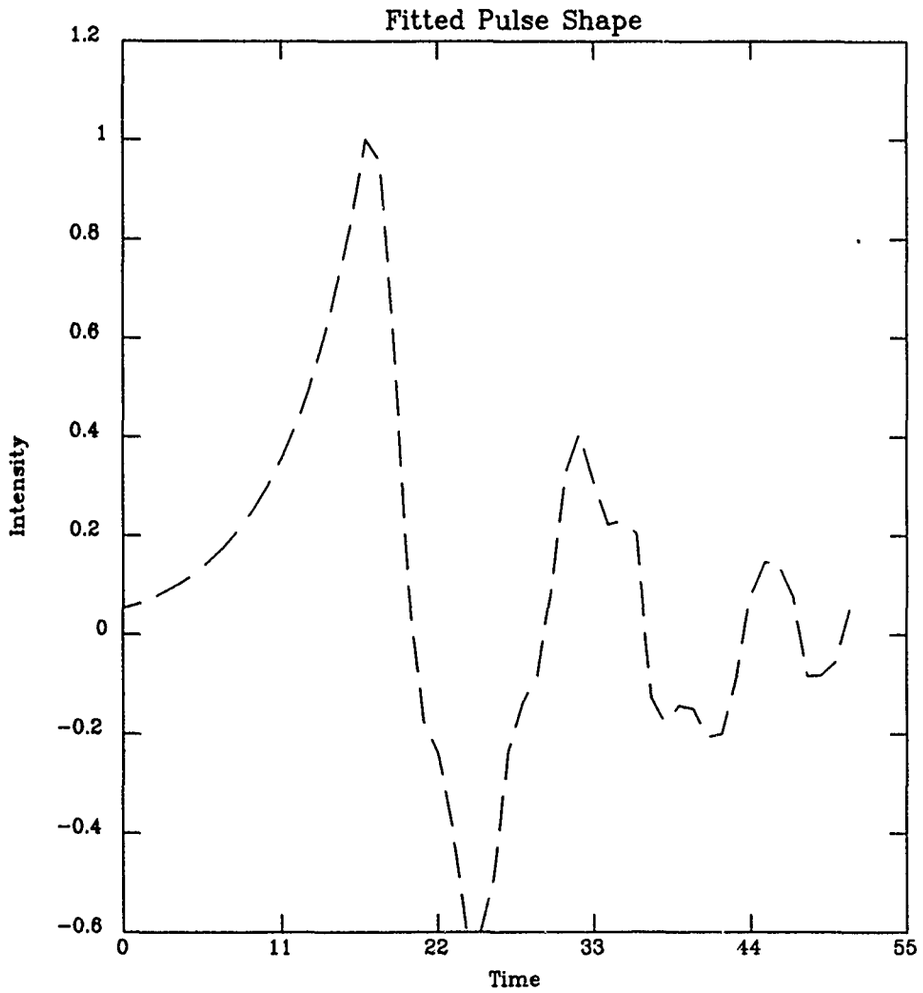


Figure 4. Deconvolved pulse shape for the data taken from 05:00 to 05:30 UT for the GINGA observations of Sco X-1.

convolved with the pulse times (that is, the reconstructed data - dashed lines) is shown in Figure 5 along with the actual data plotted as a solid line.

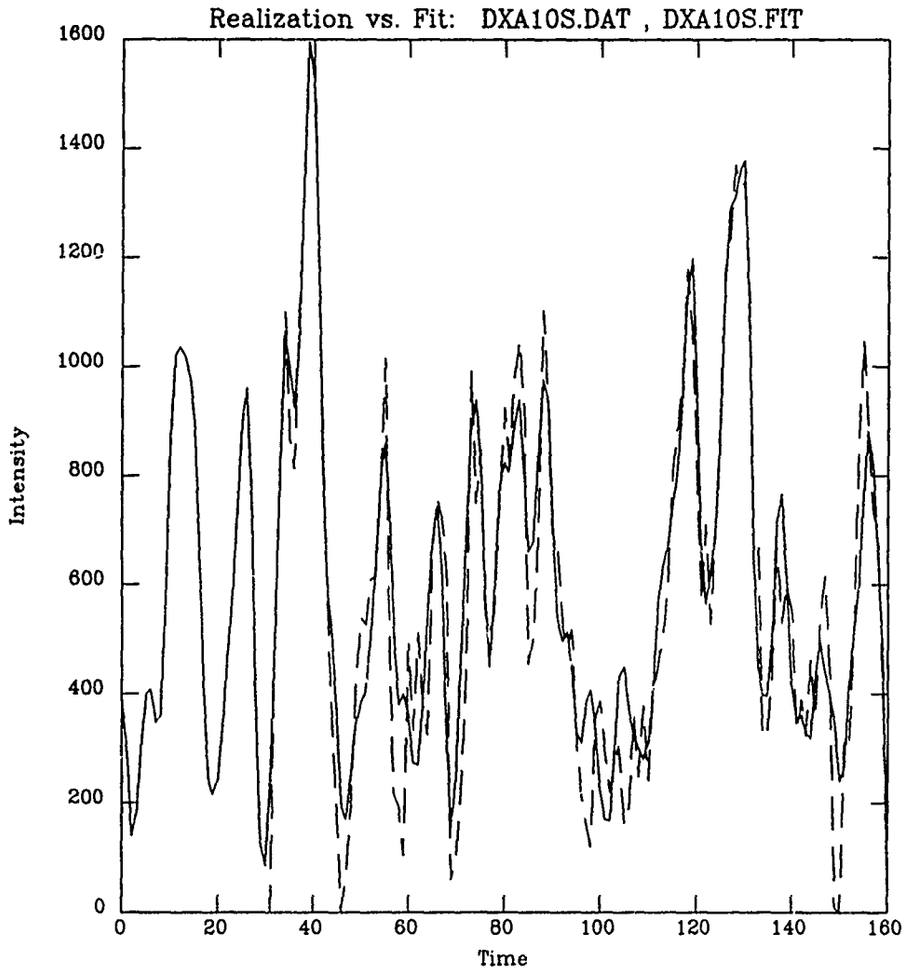


Figure 5. The original Sco X-1 data (solid line) with a slowly varying component removed), and the reconstruction of that data (dashed line) from a convolution of the pulse times and the pulse shape shown in Figure 4.

The X-ray flux and VLF phase data (both CBD and NRL) used in this segment of the analysis and shown in Figure 3 are from the time interval of 05:00 to 05:30 UT. These data are taken at a time when Sco X-1 is very near the horizon. Thus, it is unlikely that any effect of the X-ray star can be present in the data. The binned phase data plotted in 30 second phase bins show (as expected) no systematic variation in the VLF phase with respect to the X-ray flares, since the X-ray flux is at this time not directly illuminating the ionosphere.

During this interval of the data, the algorithm found 18 flares. The time between flares varied from roughly 30 to 60 seconds, with the most common interval being roughly 30 seconds, the minimum time scale for the binning of this particular run.. It is thus typically possible to search for ionospheric variations over roughly 100 to 200 seconds following flare onset, and the method is sensitive to mean phase angle variations as low as 0.7 degrees in 100 seconds, i.e., 0.007 degrees per second.

This analysis is preliminary, and does not represent a thorough investigation of possible systematic effects which may be present in the data. Work on this issue is proceeding apace with the investigation of the other segments of the data.

CONCLUSIONS

A multi-wavelength campaign to observe Sco X-1 during 10-11 March 1989 succeeded in obtaining simultaneous observations in radio through X-rays from 3 orbiting observatories and 7 ground based observatories. A portion of the X-ray data have been used to test the hypothesis that the X-ray star Sco X-1 can produce detectable fluctuations in the VLF phase and amplitude data gathered at the Naval Research Laboratory's main site (NRL) and at the Chesapeake Bay Division (CBD).

We have developed a method to detect the time of occurrence of aperiodic pulses in the X-ray light curve, and used these times as reference marks with which to fold the VLF data. The upper limits we obtain for a time interval when the X-ray star is effectively below the horizon (and therefore can have no effect on the measurement) show that the method can detect mean phase angle drifts as low as 0.7 degrees in 100 seconds (.007 degrees per second). It is therefore possible to use the method described herein to detect slight variations in the ionospheric phase over intervals of 100 to 200 seconds.

The method we use will be applied to more data from this observing sequence in subsequent publications

REFERENCES

Ananthakrishnan, S. and Ramanathan, K.R., 1969, Nature, 223, 488.

Ananthakrishnan, S., Chakravarty, S., and Ramanathan, K., 1970, Proc. Ind. Acad. Sci, 71A, 69.

A. D. Watt, VLF Radio Engineering, 1967(Pergamon Press, Oxford), p212ff.

Fishman, G.J. and Inan, U.S., 1988, Nature, 331, 418.

Goldberg, R.A., Barcus, J.R., and Mitchell, J.D., 1985, Journal of Atmospheric and Terrestrial Physics,

F. J. Kelly and Rao, D.N.M., Ionospheric Heating Effects on LF Propagation, 1981, in Effect of the Ionosphere on Radio Wave Systems based on the Ionospheric Effects, pp.100-104.

J. E. Best, Ratcliffe, J. A. , and Wilkes, M. V., Experimental Investigation of Very Long Waves Reflected from the Ionosphere, 1936,Proc. Roy. Soc. A, 156,614.

Poppoff, I.G. and Whitten, R.C., 1969 Nature, 224, 1188.

R. N. Bracewell, Budden, K. G., Ratcliffe, J. A., Straker, T. W. and Weekes, K., The Ionospheric Propagation of Low and Very-Low Frequency Waves over Distances less than 1000km, 1951, Proc. IEE, 98, part III, 53.

Svennesson, J., Reder, F., and Crouchley, J., 1972, Journal of Atmospheric and Terrestrial Physics 34, 49.

T. B. Jones and Davies,K., 1972, Nature,238.

Acknowledgements: We thank our many NRL coworkers who assisted us with this experiment and the reduction of the data. Specifically we express our appreciation to J. Ahearn, J. Barb, M. Daehler, L. DeBlasio, J. Eskinses, M. Lord, R. Nabor, Dr. Jay Norris, L. Quinn, T. Priddy, and M. Daehler.

AD-P006 323

HF PROPAGATION THROUGH ACTIVELY MODIFIED IONOSPHERES

Paul E Argo, T. Joseph Fitzgerald
John H. Wolcott, and David J. Simons
Los Alamos National Laboratory
Los Alamos, New Mexico 87545

Steve Warshaw and Ralph Carlson
Lawrence Livermore National Laboratory
Livermore, California

ABSTRACT

We have developed a computer modelling capability to predict the effect of localized electron density perturbations created by chemical releases or high-power radio frequency heating upon oblique, one-hop hf propagation paths. We have included 3-d deterministic descriptions of the depleted or enhanced ionization, including formation, evolution, and drift. We have developed a homing ray trace code to calculate the path of energy propagation through the modified ionosphere in order to predict multipath effects. We also consider the effect of random index of refraction variations using a formalism to calculate the mutual coherence functions for spatial and frequency separations based upon a path integral solution of the parabolic wave equation for a single refracted path through an ionosphere which contains random electron density fluctuations.

A three-dimensional Hamiltonian Raytracing program for Ionospheric Radio Propagation

This computer program tracks (hence is named TRACKER) the three-dimensional paths of radio waves through model ionospheres by numerically integrating Hamilton's equations, which are a differential expression of Fermat's principle. The Hamiltonian method, by using continuous models, avoids false caustics and discontinuous raypath properties often encountered in conventional raytracing methods. In addition to computing the ray path, TRACKER also calculates the group path (or pulse travel time), the phase path, geometrical (or "real") path length, and Doppler shift (if the time variation of the ionosphere is explicitly included). Computational speed can be traded for accuracy by specifying the maximum allowable integration error per step in the integration.

This program is an extension of a three-dimensional Hamiltonian integration code developed in the late '1960s' by R. Michael Jones and J. M. Stevenson, commonly called the "Jones code." We have substituted a modern linear differential equation solving routine (Hindmarsh, 1980) for the Runge-Kutta solver in the Jones code. This not only allows the program to integrate accurately with larger step sizes, but also enables the program to integrate through spatially steeper electron density gradients with smaller integration errors.

This program, TRACKER, has followed the "Jones' code" concept of modularity. The program can be extended to integrate any geophysical form of raytracing by simply replacing the subroutine defining the Hamiltonian. The models of the background index of refraction, and perturbations thereon, are also in a few replaceable subroutines.

TRACKER operates implicitly in several coordinate systems. The interface to the user is done in the usual geographic coordinates, but internally the actual raypath is calculated in the more physical geomagnetic coordinates (note here that we use a tilted dipole, rather than the more complex and accurate "corrected" geomagnetic coordinates). Because we are specifically interested in the effects of perturbations on the raypaths we also include coordinate systems relevant to each perturbation, making those calculations more tractible. TRACKER calculates gradients in the index of refraction explicitly, thus allowing nonanalytic forms of ionospheres and perturbations.

Raytracing codes cannot in general compute the raypath that connects a specified source and receiver. This "homing" capability is usually included by launching a fan of rays at small increments of azimuth and elevation, and linearly interpolating to find the ray that reaches the receiver location. In many cases raypaths must be found that approach to within fractions of a wavelength of the receiver. TRACKER includes a formalism to treat this "homing" problem as a differential equation of its own. In this case the one dimensional zero crossings of the rayfans are calculated (azimuth and elevation are treated independently). These are solved sequentially and iteratively until a "homed" ray of sufficient accuracy is found.

91-09679

Unlike the original "Jones' code," this program is not batch oriented, but rather is interactive. When the program is started, the user specifies what form of graphics terminal is being used. This allows to program to most efficiently display the outputs. The user then has approximately seventy five (75) commands with which to tailor his operating session. These commands cover an extensive range including controlling plot outputs (from none to directing it to special metafiles), modifying the physical raytrace parameters (such as frequency, receiver or transmitter locations, ray type, etc.), the numerical integration parameters (step size, error tolerances, etc.), and complex operations such as homing, finding multipath situations, etc.

Even in situations where ray theory does not apply, a picture of the radio raypaths often provide a "quick look" at the radio energy goes in the ionosphere. Such a picture may give the operator or scientist some insight into the specific problem, and may suggest a way to solve the individual problems encountered.

What is Raytracing?

Raytracing is a method of simulating the propagation of waves through a continuously varying refractive index structure. It uses the WKB, or high-frequency, approximation to find solutions to elliptic or hyperbolic equations that describe the flow of energy in the medium. It is closely related to the "method of characteristics," because the calculated raypaths are the characteristics of the source differential equation. Without higher order corrections, ray theory accounts only for refraction by large-scale gradients in the background medium. In these cases raytracing will give accurate information about the geometrical raypath followed by the radio wave, about shadow boundaries (skip zones) and reflections from surfaces, phase path lengths and pulse travel times and focusing effects. In cases where multiple rays reach the receiver, additional theory will be required to combine the field information from the multiple rays. In the worst cases, where the actual applicability of the ray theory is in question, the ray path picture can help determine which regions must be treated by more complete means.

The Hamiltonian raytracing technique involves numerically integrating Hamilton's equations, given varying sets of initial conditions. In general, the Hamiltonian equations are constructed from an expression of Fermat's principle of least action. In this case, the variation principle is expressed as an integral of a Lagrangian along some as yet unspecified path, in terms of a set of generalized coordinates, q_i. As in the generalized Hamiltonian formulation, the generalized moments p_i are also defined; in this case they are directional components of the wave number vector k. Then a Hamiltonian $H(q_i, p_i)$ is constructed from the Lagrangian. Integrating the equation $H(q,p) = 0$ then gives a path that satisfied Fermat's principle.

In Cartesian coordinates, the raytracing form of Hamilton's equations are

$$\frac{dx_i}{d\tau} = \frac{\partial H}{d.}; \quad \frac{dk_i}{d\tau} = -\frac{\partial H}{\partial x_i}, \quad i = 1 \text{ to } 3$$

where τ is a parameter usually proportional to time, k_i are the wave number components, and x_i are the raypath physical coordinates. In a full spherical coordinate frame, the equations become significantly more complicated. This program uses the set of equations due to Haselgrove (1954) and Jones (1982).

Most Hamiltonian raytracing programs use group path $P' = ct$ as the independent variable because partial derivatives of the Hamiltonian with respect to P' are independent of the form of the Hamiltonian. This choice not only allows the program to change Hamiltonians in the middle of a raypath, it also automatically forces the integration to take smaller "real" path length steps when the ray is near the reflection region.

In order to solve these equations for the raypath, initial values for the six quantities x_i and k_i are selected (i.e. start location and initial direction of propagation), and a numerical integration of the system of six differential equations is performed.

In this form of raytracing, the user can trade accuracy for speed by changing the maximum allowable integration error per step. The numerical integrators automatically adjust the integration step length along the raypath to keep the integration error below the specified bounds. In regions where the index of refraction is rapidly varying, small steps are necessary, while in slowly varying regions the step sized may be quite large.

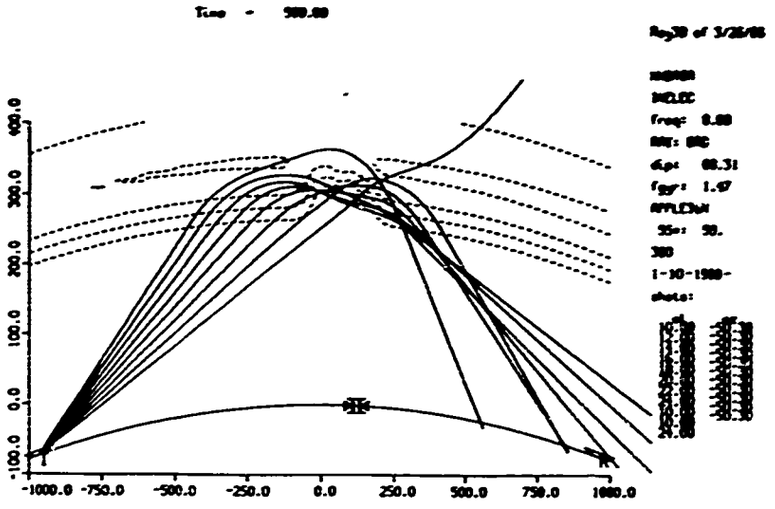


Figure 1. Rayfan through a pair of chemical releases, seen in the plasma density contours.

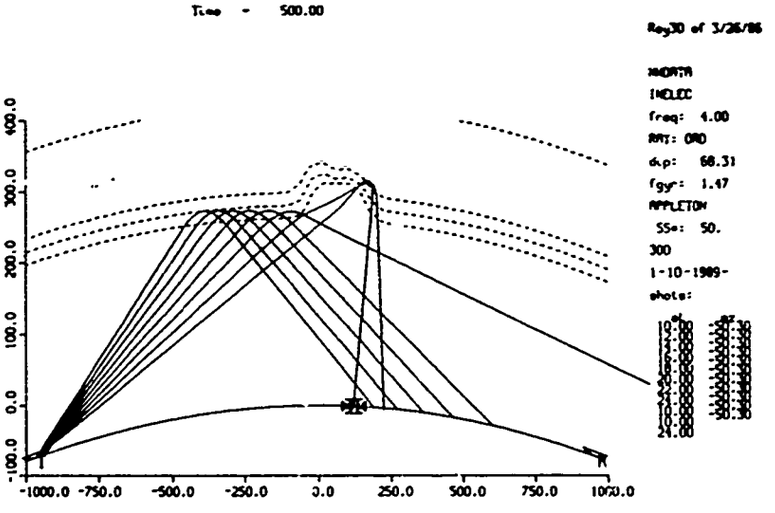


Figure 2. Similar to Figure 1, but at a higher frequency (8 MHz). Note that the depletions appear to act as a "focusing" lens.

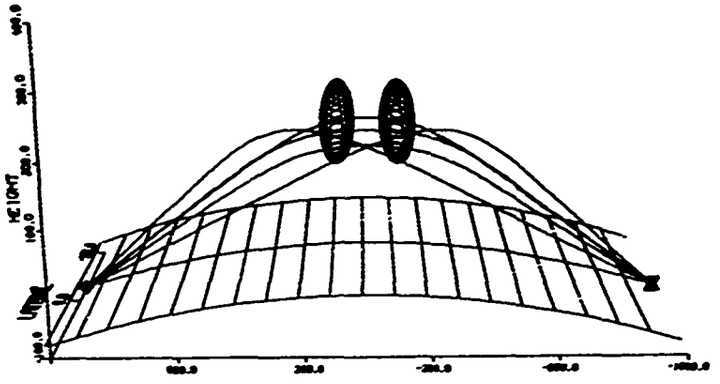


Figure 3. Chemical depletions can cause the formation of new propagation paths, above and beyond the "high" and "low" rays.

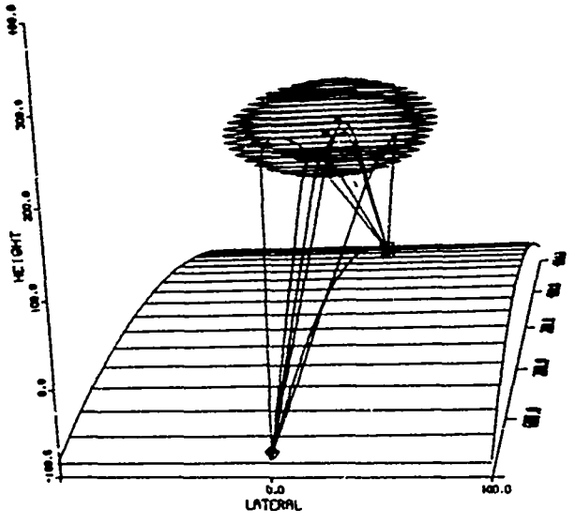


Figure 4. These new propagation paths are usually non-great circle, and are due to the focussing lens effects of the ionospheric depletions.

Examples of raypaths through chemical depletions

The outputs of the TRACKER raytracing program are varied, and tailored to studying propagation through various ionospheric perturbations. Figure 1 shows a rayfan, "shot" into a pair of chemical releases. The ionospheric plasma density contours are shown, and the depletions are evident. The effects on the rayfan are marked. Figure 2 shows the same ionosphere, but models propagation at a higher frequency (8 MHz). Note that there are still marked effects on the rayfan. These effects can in general be described as those of a focusing lens.

Figures 3 and 4 show the multiplicity of paths that can be formed by a pair of chemical releases. Two of the five paths are disturbed "low" and "high" rays, the other three are new modes generated by the focusing effects of the depletions. The time delays and signal strengths from each mode can be added to provide an estimate the multipath fading, and the stochastic model described in the next sections can be applied to each of the paths.

Effect of Random Electron Density Fluctuations

A complete prediction of hf propagation with TRACKER will ultimately be limited by the fact that the ionosphere is subject to random distortions over a wide range of scale sizes and time regimes. Such distortions arise from electron density fluctuations caused by a number of physical processes depending on location and time; the net result is that the hf communication path deteriorates due to fading, multipath, pulse spreading, etc. To predict the statistics of hf paths which reflect from an ionosphere which contains random macroscopic fluctuations in electron density we employ a method developed to describe acoustic propagation in the oceanic sound channel [*Flatté*, 1983]. In this method the scalar wave equation is approximated for long distance propagation with only small deflections by the parabolic wave equation which can be cast in the form of Schrödinger's equation and thus has a path integral solution [*Flatté*, 1986]. The statistics of the received amplitude may be calculated from the formal path integral solution without calculating a realization of the path integral solution.

This formalism allows us to calculate the statistics of single-mode hf transmissions over oblique one-hop paths. The background electron density may have an arbitrary spatial dependence although our assumption of small angular deflection precludes steep gradients in the refractive index. The random spatial structure of the electron density is assumed to be characterized statistically by a power law spectrum with outer and inner scales [Shkarofsky, 1968]. The spectral index, outer scale, inner scale, and the total variance may be a function of spatial location and altitude. The time variation of the random electron density is assumed to arise from the transport of fluctuations at a constant velocity and thus is controlled by the spatial statistics and the transport velocity. The lowest order statistics that are of interest are the second moments of the wave amplitude for spatial, time, and frequency separations. These are called the mutual coherence functions and characterize the angular spread of the received transmissions (spatial separation), Doppler spread (time separation) and coherent bandwidth (frequency separation). Our results predict the ensemble average behavior of the mutual coherence functions [Dashen, Flatté, and Reynolds, 1985].

Spatial and Frequency Coherence: Sample Results

As an example of the effects predicted by the formalism we will present numerical results for a 1950 km path in the polar region. The ionosphere is modelled as a single nighttime F layer with a critical frequency of 5.1 MHz at a height of 310 km. The electron density profile is assumed to be a Chapman layer with a scale height of 37 km. Figure 5 shows the ray path of an 11.1 MHz transmission in the rectangular coordinate system used for the calculation; the curved arcs indicate the surface of the earth and the altitude one scale height below the layer maximum. The low-ray, which we consider here, has an elevation angle with respect to the earth's surface of 12.55°. The ray path is oriented magnetic east-west; we assume a dip angle of 80° for the field.

For this example we assume that the variance of the relative electron density fluctuations is independent of altitude and equal to 0.01. The density fluctuations follow the Shkarofsky distribution with an outer scale of 10 km, a spectral index of 1.65, and an inner scale of 0.01 km. For these parameters, the variance of the phase fluctuations, Φ^2, is 4160 rad^2. The spatial correlation lengths of the fluctuations are 10.2 km in the vertical direction and 2.1 km in the horizontal direction. The relatively large correlation length in the vertical direction is a consequence of the elongation of the structure along the steeply sloping magnetic field.

The calculated mutual coherence for receiver separations in the transverse and vertical directions are shown in Figure 6. In the transverse direction, the coherence length is approximately 80 m while in the vertical direction it is

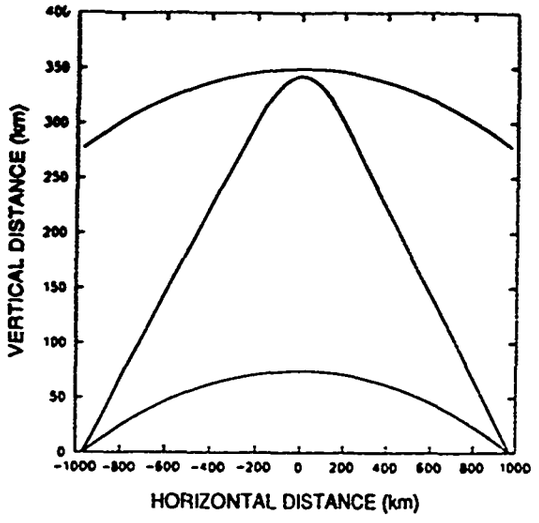


Figure 5. **Raypath in rectangular coordinates used for numerical simulation. We assume a single layer, Chapman-profile, electron density distribution with a maximum at 312 km altitude. The lower curve represents the surface of the earth; upper curve is at an altitude of one scale height below that of the maximum**

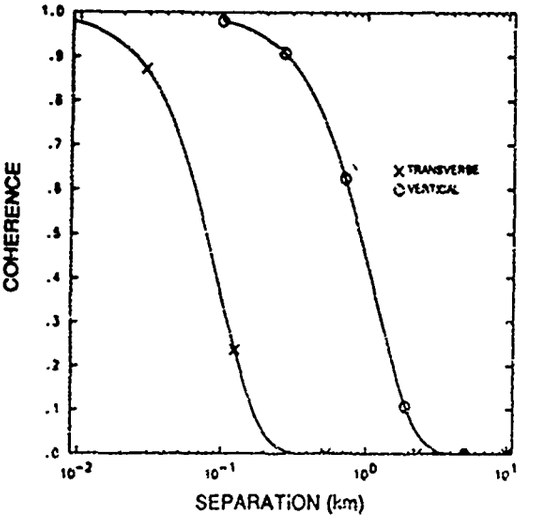


Figure 6. Caculated mutual coherence for receiver separations in the transverse and vertical directions for the ray shown in Figure 5.

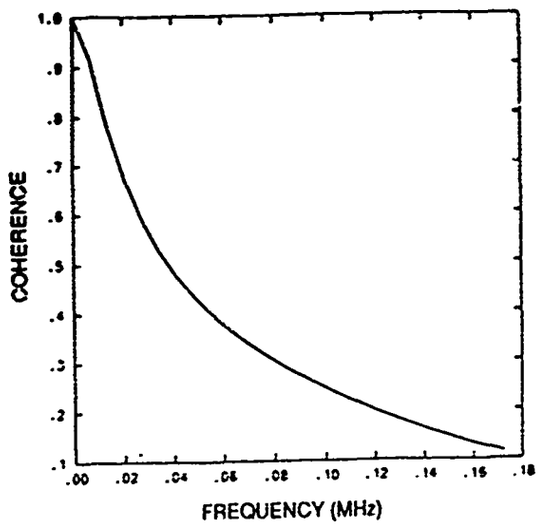


Figure 7 Calculated frequency coherence for the ray shown in Figure 5; the estimated bandwidth is 40 kHz.

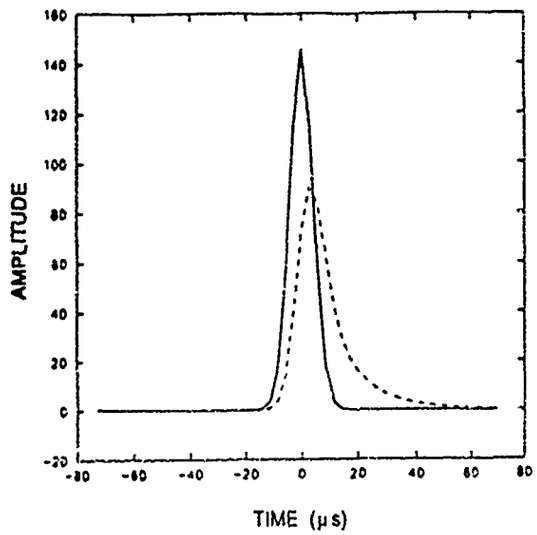


Figure 8 Ensemble average pulse response for ray shown in Figure 5. The solid curve represents the amplitude of the transmitted pulse, a Gaussian with a full width at half-maximum of 10 μs. The dashed curve represents the ensemble average of the received pulse which shows a delay and spread relative to the transmitted pusle

approximately 900 m. The difference in coherence lengths for the two directions reflects the fact that the structures are elongated in the vertical direction. The angular spread of the received signal in the transverse direction, given the calculated coherence length, corresponds to a half angle of approximately 10°.

The results of the simulation for the frequency coherence are shown in Figure 7 where we estimate a bandwidth of 40 kHz. The ensemble average pulse response is illustrated in Figure 8 where we assume a transmitted pulse of Gaussian profile with a full width at half maximum of 10μs. The average delay, τ_1, is 5.3 μs while the average spread, τ_o, is 6.8 μs. These parameters depend greatly on the electron density profile; for example, the Chapman layer has a relatively steep bottom edge so that paths reflecting below the nominal ray path do not contribute as much as these reflecting above the nominal path. Thus the ensemble average pulse tends to be delayed compared to the unperturbed pulse. A less steep profile, for example, a parabolic profile, would tend to enhance lower paths and lead to a spread towards early arrivals.

Acknowledgment: This work was supported by the U.S. Department of Energy.

REFERENCES

Dashen, R., S. M. Flatté, and S. R. Reynolds, Path-integral treatment of acoustic mutual coherence functions for rays in a sound channel, *J. Acoust. Soc. Am.*, *77*, 1716–1722, 1985.

Flatté, S. M., Wave propagation through random media: contributions from ocean acoustics, *Proc. IEEE*, *71*, 1267–1294, 1983.

Flatté, S. M., The Shrödinger equation in classical physics, *Am. J. Phys.*, *54*, 1088–1092, 1986.

Hindmarsh, A. C., LSODE and LSODI, two new initial value ordinary differential equation solvers, ACM-SIGNUM Newsletter, Vol. 15, No. 4, 1980, pp. 10–11.

Shkarofsky, I. P., Generalized turbulence space-correlation and wave-number spectrum-function pairs, *Can. J. Physics*, *46*, 2133–2153, 1968.

AD-P006 324

IONOSPHERIC MODIFICATION EXPERIMENTS WITH SATELLITE-BORNE AND GROUND-BASED TRANSMITTERS

M. C. Lee
Plasma Fusion Center
Massachusetts Institute of Technology
Cambridge, Massachusetts 02139

ABSTRACT

A research program is outlined to investigate the nonlinear VLF propagation and interaction with ionospheric plasmas with the satellite-borne and ground-based transmitters. Coordinated space and ground-based ionospheric modification experiments have been planned to study nonlinear propagation and mode conversion of VLF waves into lower hybrid waves in the presence of HF heater created ionospheric irregularities. Space experiments with the mother (ACTIVE)-daughter (CZAK) satellites are also discussed to examine the excitation of parametric instability by powerful VLF waves. Proposed ground-based radar, optical, and VLF measurements can provide effective diagnoses of the induced ionospheric effects. It is expected that electron acceleration caused by the lower hybrid waves will produce intense airglow emissions. In addition, a broad height distribution of enhanced plasma lines is also expected for radars to detect.

1. Introduction

Active VLF wave injection experiments have been conducted with powerful ground-based transmitters for over two decades and have centered on the study of VLF wave propagation and interaction with energetic charged particles in the magnetosphere [see, e.g., *Helliwell*, 1983 and references therein]. By contrast, the possible interaction between VLF waves and the ionospheric plasmas received little attention. However, it is shown theoretically that intense lower hybrid waves and plasma density striations can be favorably excited by either injected or lightning-induced VLF waves [*Lee and Kuo*, 1984]. The wavelengths of the excited lower hybrid waves (or field-aligned density striations) are found to be a few meters in the lower ionosphere, a few tens of meters in the upper ionosphere, and a few hundreds of meters in the magnetosphere at L=4 [*Lee and Kuo*, 1984]. In fact, observations of VLF wave-excited lower hybrid waves in the ionosphere were indicated in the experiments of the Franco-Soviet ARCAD 3 satellite [*Berthelier et al.*, 1982; *L. R. O. Storey and F. Lefeuvre*, personal communication, 1984] and others [e.g., *Chmyrev et al.*, 1976, 1986].

Spectral broadening of injected monochromatic VLF waves [*Bell et al.*, 1983; *Inan and Bell*, 1985] is a recently discovered phenomenon, arising from the nonlinear VLF wave propagation in the structured ionospheric plasmas. Our recent theoretical analysis of this phenomenon has identified two potential source mechanisms [*Groves et al.*, 1988]. One is the nonlinear scattering of VLF waves by ionospheric

91

91-09678

density fluctuations which renders the nonlinear mode conversion of VLF waves into lower hybrid waves. This mechanism generates single-peaked spectra centered at the carrier frequency. In the absence of ionospheric irregularities, if the injected VLF waves are intense enough, a parametric instability can be excited and generate the lower hybrid waves. This process tends to produce a spectrally broadened transmitted pulse with peaks at a discrete set of frequencies on both sides of the nominal carrier frequency. Both types of spectra were observed [see, *Bell et al.*, 1983]. Another interesting phenomenon to be mentioned is the so-called explosive spread *F* [*Woodman and La Hoz*, 1976; *Woodman and Kudeki*, 1984] which, according to our theoretical investigation of it [*Liao et al.*, 1989], is primarily caused by the lightning-induced VLF (whistler) waves. Some puzzling and unexplained features of the phenomenon can be reasonably understood with the characteristics of VLF (whistler) wave propagation and nonlinear interaction with the ionospheric plasmas.

Experiments for ionospheric plasma modification by the combined operation of powerful HF and VLF transmitters were discussed by M. C. Lee et al. in the First Suzdal Symposium on "Ionospheric Modifications by Powerful Radio Waves," which was held in Suzdal/Moscow, U.S.S.R. during 8–12 September, 1986. What was proposed were stimulated ionospheric plasma processes under the action of HF and VLF waves injected from ground-based transmitters. In brief, the concerned physical processes include the nonlinear VLF (whistler) scattering by the HF wave-induced ionospheric density striations and the nonlinear VLF wave-plasma interaction which further enhances the ionospheric plasma modifications caused by the HF heater waves.

Several ionospheric heating facilities are currently operated at the frequencies of a few MHz in the HF or MF band in the U.S.A., Europe, and the U.S.S.R., while VLF transmitters have been used at different locations for the study of wave-particle interaction in the magnetosphere. We thought that it was impossible to use the existing HF (or MF) and VLF transmitters for the proposed experiments. Therefore, we suggested that the experiments be carried out with available ionospheric heating facilities during lightning storms [*Lee et al.*, 1988]. It is not necessary to have lightning storms occurring nearby. If lightning occurs at the conjugate location in the opposite hemisphere, the lightning-induced VLF waves can propagate through the HF modified ionospheric region in ducted whistler modes and bounce back and forth for several cycles before they die out.

Although the proposed experiments, in principle, can be carried out, they are by no means easy technically. Nevertheless, these conceptualized experiments become possible and controllable after the launch of the ACTIVE satellite, which carries a powerful VLF transmitter. As illustrated in Figure 1, the satellite-borne transmitter can launch the VLF waves propagating through the HF heater-modified ionospheric region, when the orbiting satellite is in appropriate positions. Outlined in Section 2 are experiments we propose to conduct. The theoretical basis for proposed experiments is elucidated and the diagnosis of predicted ionospheric effects is described. A further description of diagnostic instruments and planning for the experiments is presented in Section 3.

2. Proposed Experiments

(A) Coordinated Space and Ground-based Ionospheric Heating Experiments

Ionospheric heating experiments have been planned at Arecibo, Puerto Rico with the combined operation of a ground-based HF transmitter and a satellite-borne VLF transmitter. The purpose of these coordinated space and ground-based heating experiments is to investigate the nonlinear VLF wave

propagation through the ionospheric region, which is highly disturbed by a powerful HF wave. The proposed experimental scheme is delineated in Figure 1 with the following scenario.

The ionosphere is first illuminated by a vertically transmitted HF heater wave and, then, a VLF wave from the ACTIVE satellite. Anomalous absorption of the HF wave occurs near its reflection height in the ionospheric F region. It has been known both theoretically and experimentally that short-scale (*viz.*, less than a few meters) ionospheric irregularities can be generated within seconds by HF heater waves, while it takes tens of seconds or longer for relatively large-scale (typically, hundreds of meters up to a few kilometers) irregularities to develop [see, e.g., *Gurevich*, 1978; *Fejer*, 1979; *Lee and Kuo*, 1983a; *Frey et al.*, 1984]. Distinct VLF propagation effects in the HF heater modified ionosphere are expected, depending on the scale lengths of the HF-induced ionospheric irregularities. As elaborated below, nonlinear scattering of VLF waves is most effective by short-scale (meter-scale) irregularities [*Groves et al.*, 1988]. It is also our interest to study VLF wave propagation in the presence of large-scale irregularities whose scale lengths are comparable to the VLF wavelength ($\sim$ a few hundreds of meters in the F region).

For simplicity, we assume that the VLF wave transmitted from the ACTIVE satellite propagates along the Earth's magnetic field in a ducted whistler mode. The whistler wave field is represented by

$$\vec{E}_o = (\hat{x} + i\hat{y})E_o \exp[i(k_o z - \omega_o t)] \tag{1}$$

where ω_o and k_o are the wave frequency and wave number, respectively; the z-axis of the coordinate system coincides with the geomagnetic field. Further, the HF heater-induced field-aligned ionospheric irregularities are denoted by

$$\delta n = \delta\hat{n} \exp(iky) \tag{2}$$

where the scale length of the irregularities is given by $2\pi/k$. We consider the VLF (whistler) propagation in the presence of ionospheric irregularities.

Solving the wave equation, we find that ionospheric irregularities can scatter nonlinearly the circularly polarized (whistler) wave into an elliptically polarized wave. A linearly polarized component of the scattered VLF (whistler) wave is introduced by the HF heater-induced ionospheric irregularities. The scattered wave field is found to be

$$\vec{E}_s = \left\{(\hat{x} + i\hat{y})E_{cp} + \hat{y}E_{lp}\right\} \exp[i(k_o z - \omega_o t)] \tag{3}$$

where E_{cp} and E_{lp} designate, respectively, the field intensity of the circularly polarized component and that of the linearly polarized component; they are related by

$$E_{cp} = 2\left(\frac{\lambda}{\lambda_o}\right)^2 E_{lp} = 2E_o\left(\frac{\delta n}{N_o}\right)\left(\frac{\lambda}{\lambda_o}\right)^2 \tag{4}$$

where $\lambda = 2\pi/k$ and $\lambda_o = 2\pi/k_o$ are the scale length of ionospheric irregularities and the wavelength of the whistler wave, respectively; N_o represents the unperturbed plasma density. It is clear from (4) that E_{lp} dominates over E_{cp} when $\lambda \ll \lambda_o$. The required ionospheric plasma conditions can be determined as follows. If a VLF wave at the frequency of 10 kHz is injected, and the ionospheric conditions are assumed to be f_c (electron gyrofrequency) = 1.3 MHz and N_o (background plasma density) = 4.5×10^{11} m^{-3}, the whistler wavelength in the ionosphere is found to be about 500 meters. Therefore, the HF heater-induced

meter-scale irregularities can effectively render the nonlinear mode conversion of a VLF (whistler) wave into an electrostatic (lower hybrid) wave.

To focus on the the study of nonlinear VLF scattering, we will operate the HF heater for at least a few seconds which assures the full development of meter-scale irregularities. Prominent ionospheric effects are associated with the appearance of lower hybrid waves. Since lower hybrid waves can accelerate electrons effectively along the Earth's magnetic field , the impact excitation of neutral particles and Langmuir waves may lead to intense airglow and a broad height distribution of plasma line enhancement as schematically illustrated in Figure 2. The monitoring of these expected ionospheric effects by radar and an all sky imager will be discussed in the next section.

We should point out that lower hybrid waves produced by whistler (VLF) waves via the nonlinear scattering process are significantly weaker than those excited by parametric instabilities, which will be examined in other proposed experiments. The field intensity of the lower hybrid waves, as shown in (4), is given by $E_o(\delta n/N_o)$. Recall that E_o and $(\delta n/N_o)$ represent the incident field intensity of the whistler wave and the fractional density fluctuations excited by the HF heater. Hence, the optimum conditions for these experiments require that (1) the position of the ACTIVE satellite be as close as possible to the HF heater-modified region (to achieve large E_o) and (2) the full development of the HF-induced meter-scale irregularities (to assure large $\delta n/N_o$).

As mentioned earlier, we also plan to investigate the VLF propagation in the disturbed ionosphere having large-scale irregularities. To distinguish it from the afore-said nonlinear scattering process, the HF heater should be operated in CW mode long enough to excite the large-scale irregularities. Moreover, the position of the ACTIVE satellite can be distant from the HF modified ionospheric region. However, we should make sure that the VLF signal, after its transionospheric propagation to reach the ground, is detectable above background VLF noise. Several VLF receivers will be deployed on the ground at the estimated footprints of the transmitted VLF (whistler) signals. Signals are going to be recorded before and after the HF heater is turned on.

(B) Space Experiments with the Mother (ACTIVE)-Daughter (CZAK) Satellites

Space experiments with the mother (ACTIVE)-daughter (CZAK) satellites can also be conducted at Arecibo (Puerto Rico). These experiments, as illustrated in Figure 3, are dedicated to the study of nonlinear VLF wave interactions with the ionospheric plasmas. In other words, stimulated ionospheric plasma processes triggered by powerful VLF waves will be investigated in the experiments.

Specifically, we plan to study the parametric excitation of lower hybrid waves by VLF (whistler) waves discussed in *Lee and Kuo* [1984]. This process may contribute to the better understanding of several outstanding phenomena, such as the spectral broadening of VLF waves, the puzzling explosive spread *F*, lightning-induced lower hybrid waves [*Liao et al.*, 1989; M. C. Kelley, personal communications, 1989], and the possible particle precipitation from the ionosphere. Before presenting the experimental plans, we highlight the concerned physical processes as follows.

For simplicity, consider a whistler wave propagating along the Earth's magnetic field, which is taken to be the z-axis of the rectangular coordinate system. The parametric excitation of lower hybrid waves by whistler (VLF) waves involves a four-wave interaction process which can be described by the

following wave-frequency (ω) and wave-vector ($\vec{k}$) matching relations:

$$\omega_+ - \omega_s = \omega_o = \omega_- + \omega_s^* \tag{5a}$$
$$\vec{k}_+ - \hat{x}k_s = \hat{z}k_o = \vec{k}_- + \hat{x}k_s \tag{5b}$$

where the subscripts o, $+/-$ and s represent the whistler wave, the Stokes/anti-Stokes components of the lower hybrid sidebands and the concomitantly excited field-aligned zero-frequency mode, respectively. The parametric instability is characterized by a two dimensional process and the excited low-frequency mode is a purely growing mode. Hence, both the Stokes and anti-Stokes components of the lower hybrid waves have to be taken into account in the analysis of the instability. Further, whistler propagation along the geomagnetic field provides the most favorable symmetric configuration for the four-wave interaction process to occur.

Two frequency domains of the whistler (VLF) wave have been identified, wherein one nonlinear effect dominates over the other in driving the concerned parametric instability:

(a) Frequency domain 1:

$$\omega_{lh}\left(1 + \frac{M}{m}\frac{v_{te}^2}{c^2}\frac{\omega_{pe}^2}{\Omega_e^2}\right)^{1/2} < \omega_o < \omega_{pi} \tag{6}$$

with the non-oscillatory beating current to be the dominant nonlinear effect. Here $M\,(m)$, v_{te}, c, $\omega_{pe}\,(\omega_{pi})$, Ω_e, ω_{lh} are the ion (electron) mass, electron thermal velocity, speed of light, electron (ion) plasma frequency, electron cyclotron frequency and lower hybrid resonance frequency, respectively. The threshold electric field for exciting the parametric instability in this frequency domain is found to be

$$E_{th} = 1.2(m/e)\nu_e v_{te}|\eta|^{1/2} \tag{7}$$

where η is defined by $[1 + (M/m)(k_o/k_s)^2]/[1 - (M/m)(k_o/k_s)^2(\omega_e/\omega_{pe})^2]$.

(b) Frequency domain 2:

$$\omega_{pi} < \omega_o$$

with the thermal pressure force to be the dominant nonlinear effect. The corresponding threshold field has the following form:

$$E_{th} = 0.86(k_s^2 v_{te}^3/\Omega_e)(m/e)|1 + (1 + 4\Omega_e^2\nu_e^2\eta_e^2/k_s^4 v_{te}^4)^{1/2}|^{1/2}/|\eta|^{1/2} \tag{8}$$

The growth rates of the instability are given by

$$\gamma \sim 0.5(\nu_e k_s^2 v_{te}^2/\Omega_e^2)[E_R^2 - 1] \quad \text{for } E_R^2 \ll 10 \tag{9a}$$
$$\gamma \sim 1.4(\nu_e k_s^2 v_{te}^2/\Omega_e^2)E_R \quad \text{for } E_R \gg 10 \tag{9b}$$

where E_R is the incident field intensity of the whistler wave normalized by the threshold field of the instability. Based on these characteristics of the parametric instability, we can make the following predictions for experimental corroboration. Both in-situ and ground-based diagnosis of the predicted ionospheric effects are proposed.

First of all, in the spectral broadening phenomenon, the nearly monochromatic VLF was seen to have a bandwidth expansion as high as 100 Hz [*Bell et al.* , 1983; *Inan and Bell*, 1985]. We believe that

the off-carrier components, having electrostatic nature, are induced lower hybrid waves. If the aforementioned parametric instability is excited, we expect that the VLF receivers carried by the daughter (CZAK) satellite will detect a spectrally broadened VLF signal from the mother (ACTIVE) satellite with peaks at a discrete set of frequencies on both sides of the nominal carrier frequency. Secondly, from (7): $E_{th} \propto [1 + (M/m)(k_o/k_s)^2]/[1 - (M/m)(k_o/k_s)^2(\Omega_e/\omega_{pe})^2]$ and (9a & 9b): $\gamma \propto k_s^2$, it is clear that the excitation of plasma modes with larger wave number, k_s (i.e., small scale lengths, λ_s) requires lower thresholds; consequently, the short-scale modes can be quickly excited. In the upper ionosphere, the threshold fields are typically less than 1 mV/m and the growth rates are about 1 Hz for $E_R \sim 30$.

Since both the Stokes and anti-Stokes components of lower hybrid waves are excited simultaneously by the parametric instability, a standing wave pattern is formed across the Earth's magnetic field. Therefore, the excited lower hybrid waves have a zero perpendicular phase velocity. However, the excited lower hybrid waves can propagate along the Earth's magnetic field. From the dispersion relations of lower hybrid waves and whistlers, we find the group velocity of the lower hybrid waves is related to that of whistlers by

$$\left(\frac{dw_o}{dk_o}\right)_{lh} = \frac{1}{2}\left[\frac{(M/m)(k_o/k_s)^2}{1 + (M/m)(k_o/k_s)^2}\right]\left(\frac{dw_o}{dk_o}\right)_{whistlers} \tag{10}$$

Since $(M/m)(k_o/k_s)^2 \sim O(10^{-2}) \ll 1$, the whistler waves have a group velocity much greater than those of the excited lower hybrid waves. If the in-situ measurements of VLF waves are made by the CZAK subsatellite, it is expected that a significant delay exists in detection of the primary wave (i.e., the whistlers) and the decay waves (i.e., the excited lower hybrid waves).

In addition to the suggested in-situ measurements for the CZAK subsatellite to carry out, we also offer ground-based diagnoses with radars, all-sky imager, and VLF receivers. It is expected that the field intensities of lower hybrid waves excited by the parametric instability are much greater than those produced by the nonlinear (mode conversion) scattering process. Electron acceleration by lower hybrid waves in this case may yield intense airglow [*Chernyev et al.*, 1986]. Consequently, a very broad height distribution of enhanced plasma lines can be detected by radars [*Carlson et al.*, 1982]. It requires that the ACTIVE–CZAK satellites reach their low orbital positions for the best diagnosis of the proposed experiments. The VLF transmitter is operated in pulsed mode; the duration of the pulses needs to be long enough ($\sim$ a few seconds) for the excitation of the parametric instability. The growth time of the instability can be determined by varying the pulse duration and monitoring the intensities of VLF waves, airglow, and plasma lines recorded on the ground.

3. Diagnostic Instruments and Planning for Experiments

Ground-based optical imaging measurements are proposed as part of the coordinated diagnostics during the RF transmissions from the ACTIVE satellite. st electrons accelerated during these transmissions can cause impact excitation of various spect species. Optical emissions (artificial auroras) from these species can then be detected by sensitive imaging systems to yield information on the location, intensity and energy of the electron fluxes. For instance, the excitations of the O (^1D and ^1S) states of atomic oxygen are produced by fluxes of electrons with energies exceeding 1.96 and 4.17 eV, respectively, and lead to optical emissions at 6300$\AA$ and 5577$\AA$. Other possible processes include excitation of the O ($^3P\ ^5P$) (threshold energy of $\sim$ 9.6 eV and radiation of 7774$\AA$) and N_2^+ (ING) (threshold $\sim$ 18.7 eV and radiation at 4278$\AA$). Thus, optical measurements at multiple wavelengths can

be used to detect characteristics of the spectrum of accelerated electrons, at least in the sense that several threshold energies have been exceeded.

Optical imaging measurements have been successfully employed during ground-based HF heating experiments [*Bernhardt et al.*, 1988]. The AFGL All Sky Imaging Photometer (ASIP II), which will be used for the proposed experiments, has also successfully detected emissions from recent heating experiments at Arecibo, Puerto Rico [*Weber et al.*, 1989]. ASIP employs a variable field of view with interchangeable lenses covering 180°, 90°, 40°, 20°, and 12° full field. A filter wheel rotates up to four interference filters into the optical axis of the images by successive exposures. An image intensifier is used to increase the gain followed by a 512 × 512 CCD (charge-coupled device) for signal detection.

Some unique aspects of detection of optical emission features from a moving satellite must be considered, as compared to detection of these same features generated by a fixed ground-based transmitter. The main difference is related to exposure time. Typical exposure times are $\sim$ 10 sec to detect weak emission features. Successive images can be added separately during heater-on and heater-off periods to increase signal to noise; the heater-on and heater-off images are then subtracted to separate the heater induced component from the sky background. This technique requires that the heater-on duration is greater than the photometer exposure time which is easily achieved for a ground based heater. The life time for build up of emission from forbidden excited states ($\sim$ 100 sec for $O(^1D)$) must also be considered.

For a moving satellite the factor limiting the exposure time is the motion of the satellite (the motion of the beam of accelerated electrons) across the ASIP II detector. For a satellite at 500 km perigee, with a nominal velocity of 7 km/sec, the satellite will change zenith angle at a rate of .15°/sec near the horizon and up to $\sim$ .80°/sec near the zenith. Assuming a 40° field of view lens and that 500 CCD pixels are available for the image, maximum exposure times are only .01 sec/pixel near the zenith and .53 sec/pixel near the horizon. For 180° field of view lens, these times increase to .45 sec/pixel and 2.4 sec/pixel, respectively. Thus, spatial resolution is sacrificed for longer exposure times with the 180° lens.

These numbers depend on the exact orbit achieved by the ACTIVE satellite after launch. In addition, the actual photo emission altitude may be lower than 500 km since the accelerated electrons must penetrate to low enough altitudes to collide with neutral atomic oxygen. These simple calculations serve to illustrate, however, the limiting effect of the satellite motion on the effective exposure time of each image. Any active control of the location of interaction region, such as maintaining the antenna orientation at a fixed position with respect to the ground station, would increase the probability of detecting optical emissions.

The task of the ground-based VLF receivers is to monitor the signal transmitted from the orbiting ACTIVE satellite. Doppler shifting of the signal frequency due to the orbital velocity (of the order of 7600 m/s) and inclination (83 °) will be small (of the order of $\pm$0.25 Hz) and will depend on the direction of the observer relative to the satellite motion. Thus, the receiver should have a bandwidth of at least 400 Hz in order to record the transmitted signals. Ideally, a receiving system might consist of two orthogonal loop antennas, each feeding a low-noise, 400 Hz bandwidth amplifier/receiver whose output voltage is downconverted to baseband; the two output voltages are sampled simultaneously (both in-phase and quadrature phase) and digitally stored in computer memory at a rate sufficient to insure

no aliasing occurs (> 800 Hz). The crossed-loops allow the determination of direction of arrival of the signals (but have ±180° ambiguity). In two seconds, 1600 samples will have been recorded, so an 800 point discrete Fourier transform (DFT) can be calculated for the 400 Hz band, each point separated by 0.5 Hz. In effect, the two second integration has narrowed the bandwidth from 400 Hz to 0.5 Hz and improved the SNR by 29 dB. Should a greater SNR be required, longer integration times could be used, but the time resolution of the satellite signals would be correspondingly coarser.

A portable FFT computing spectrum analyzer (Nicolet Model 446), available to AFGL, can provide digital sampling of the VLF signals. This spectrum analyzer has selectable bands of 10, 50, 100, 500, 1K, 2K, 10K, or 20K Hz wide. It calculates a 400 point spectrum from 1024 time samples. The nominal bandwidth is 1/400 of the frequency range; for example, if the 500 Hz band is chosen, each spectral point is separated by 500/400 Hz = 1.25 Hz. Other instrument features include several averaging modes, such as selection of number of spectra averaged, and various digital outputs.

Radar diagnoses of proposed experiments are available to us at Arecibo (Puerto Rico). Although radars are primarily used to detect enhanced plasma lines in our experiments, incoherent scatter observations will be made to map electric field, conductivities, and ambient ionospheric plasma parameters before the commencement of RF transmissions. This research program, devoted to the investigation of nonlinear VLF propagation and interaction with ionospheric plasmas, have been planned by a team which consists of M. C. Lee (serving as the Principal Investigator, MIT), W. J. Burke (AFGL), H. C. Carlson, Jr. (AFGL), J. L. Heckscher (AFGL), P. A. Kossey (AFGL), and E. J. Weber (AFGL). Three graduate students, K. M. Groves, K.L. Koh, and C. Yoo (all from MIT) will participate in the project as Research Assistants for part of their Ph.D. thesis research. The dedicated research areas by the team include: Radar diagnosis (Carlson, Lee), Airglow measurements (Weber), VLF detection (Heckscher, Kossey), and Theory (Burke, Lee).

Acknowledgments. This work was supported by the NASA grants NAG 5-1336 and NAG 5-1055, and the AFOSR grant AFOSR-88-0217.

References

Bell, T. F., H. G. James, U. S. Inan, and J. P. Katsufrakis, The apparent spectral broadening of VLF transmitter signals during transionospheric propagation, *J. Geophys. Res.*, *88*, 4813, 1983.

Bernhardt, P. A., L. M. Duncan, and C. A. Tepley, Artificial airglow excited by high-power radio waves, *Science*, *242*, 1022, 1988.

Berthelier, J. J. et al., Measurements of the VLF electric and magnetic components of waves and dc electric field on board the AUREOL-3 spacecraft: The TBF-ONCH experiment, *Ann. Geophys.*, *38*, 643, 1982.

Carlson, H. C., V. B. Vickwar, and G. P. Mantas, Observation of fluxes of suprathermal electrons accelerated by HF excited instabilities, *J. Atmos. Terr. Phys.*, *44*, 1089, 1982.

Chmyrev, V.M. et al., Artificial injection of very low frequency waves into the ionosphere and the magnetosphere of the Earth, *JETP Lett.*, *23*, 409, 1976.

Chmyrev, V. M. et al., Particle precipitation, optical emissions, and VLF noises in the mid-latitude ionosphere stimulated by ground-based VLF transmitters, *Proceedings of the International Symposium on "Modification of the ionosphere by powerful radio waves"*, Suzdal/Moscow, U.S.S.R., 8–12 September, 1986.

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

Frey, A., P. Stubbe, and H. Kopka, First experimental evidence of HF produced electron density irregularities in the polar ionosphere; diagnosed by UHF radio star scintillation, *Geophys. Res. Lett., 11*, 523, 1984.

Groves, K. M., M. C. Lee, and S. P. Kuo, Spectral broadening of VLF radio signals traversing the ionosphere, *J. Geophys. Res., 93*, 14683, 1988.

Gurevich, A. V., Nonlinear phenomena in the ionosphere, (Physics and Chemistry in Space 10), 370 pp., Springer-Verlag, 1978.

Helliwell, R. A., Controlled stimulation of VLF emissions from Siple Station, Antarctica, *Radio Sci., 18*, 801, 1983.

Inan, U. S., and T. F. Bell, Spectral broadening of VLF transmitter signals observed on DE 1: A quasi-electrostatic phenomena?, *J. Geophys. Res., 90*, 1771, 1985.

Kossey, P. A., Some aspects of long wave propagation in ionospheric radio wave propagation, *AFGL In-House Technical Report, May 1984*; also published in *AFGL Handbook of Geophysics and Space Environments*, 1986.

Leclert, G. P., C. F. Karney, A. Bers, and D. J. Kaup, Two dimensional self-modulation of lower hybrid waves in inhomogeneous plasmas, *Phys. Fluids, 22*, 1545, 1979.

Lee, M. C. and S. P. Kuo, Ionospheric irregularities and geomagnetic field fluctuations due to ionospheric heating, *Proceedings of International Symposium on "Active Experiments in Space," Spec. Publ. 195*, pp81–89, Eur. Space Agency, Neuilly, France, 1983a.

Lee, M. C. and S. P. Kuo, Excitation of upper hybrid waves by a thermal parametric instability, *J. Plasma Phys., 30*, 463, 1983b.

Lee, M. C. and S. P. Kuo, Production of lower hybrid waves and field-aligned plasma density striations by whistlers, *J. Geophys. Res., 89*, 10873, 1984.

Lee, M. C., J. A. Kong, and S. P. Kuo, Enhanced ionospheric modifications by the combined operation of HF and VLF heaters, *Proceedings of the International Symposium on "Modification of the ionosphere by powerful radio waves"*, Suzdal/Moscow, U.S.S.R., 8–12 September, 1986.

Lee, M. C., K. M. Groves, C. P. Liao, D. R. Rivas, and S. P. Kuo, Combined operation of two ground transmitters for enhanced ionospheric heating, *J. Geomag. Geoelectr., 40*, 1141, 1988.

Liao, C. P., J. P. Friedberg, and M. C. Lee, Explosive spread F caused by the lightning induced electromagnetic effects, Accepted for publication in the J. Atmos. Terr. Phys., 1989.

Woodman, R. F. and C. La Hoz, Radar observations of F region equatorial irregularities, *J. Geophys. Res., 81*, 5447, 1976.

Woodman, R. F. and E. Kudeki, A causal relationship between lightning and explosive spread F, *Geophys. Res. Lett., 11*, 1165, 1984.

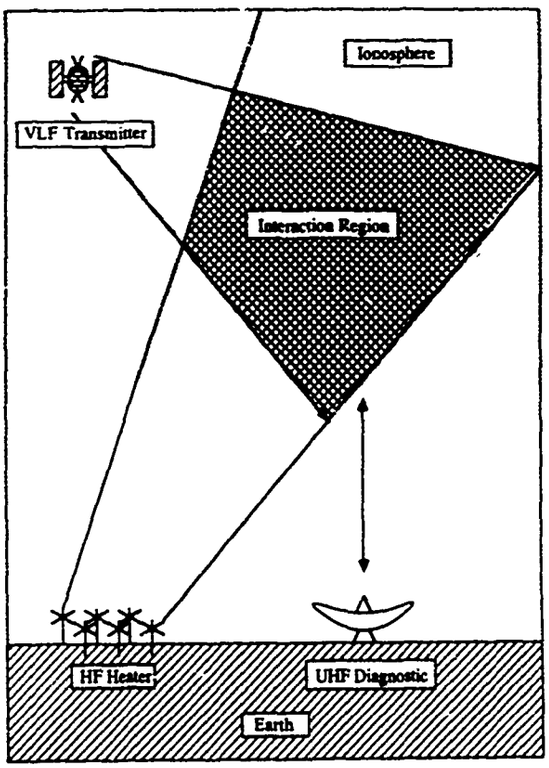


Figure 1. Coordinated space and ground-based ionospheric modification experiments.

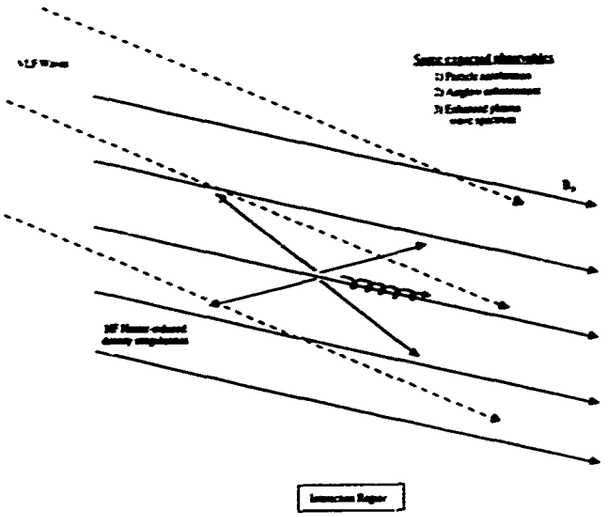


Figure 2. Expected ionospheric plasma effects including electron acceleration, airglow, and enhanced plasma lines.

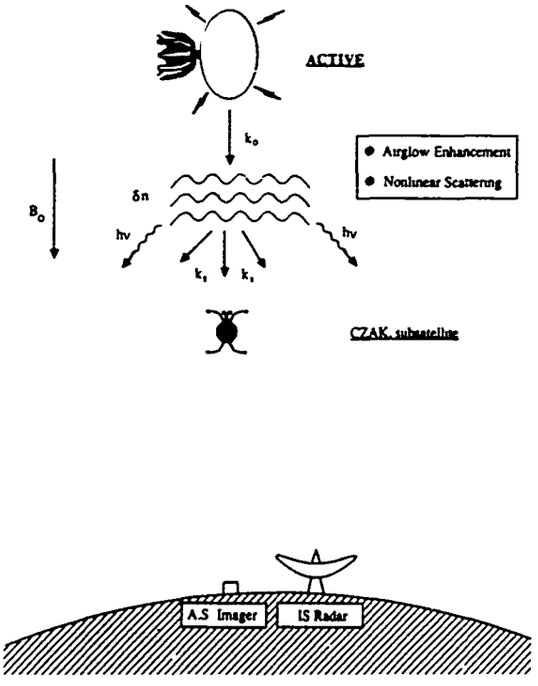


Figure 3. Space experiments with the mother(ACTIVE)-daughter(CZAK) satellite.

AD-P006 325

WIDEBAND NOISE OBSERVED AT GROUND-LEVEL IN THE AURORAL REGION

Robert F. Benson
Michael D. Desch

Laboratory for Extraterrestrial Physics
NASA - Goddard Space Flight Center
Greenbelt, MD 20771

ABSTRACT

A wideband noise event was detected at ground level from the Andøya Rocket Range in Norway in January, 1989. The signals were observed using 4 commercial communication receivers (tuned to 159, 515, 905 and 1200 kHz), an ionosonde (200 kHz to 3.5 MHz interference-free observations) and a riometer (32.5 MHz). The event, which occurred during a period of magnetic disturbance near magnetic midnight, was the only one observed during nearly 3 weeks of operations. This low frequency-of-occurrence is attributed partly to high local noise levels. The ease with which this event was identified on the ionograms produced by the local ionosonde suggests that routine ionosonde recordings should be inspected in search for such events. Such an effort would enhance existing research directed toward developing techniques for identifying quiet communication channels and help to identify the origin and frequency-of-occurrence of high latitude wideband noise events.

INTRODUCTION

Reber and Ellis [1956] apparently were the first to report the ground-level detection of high latitude terrestrial signals with frequencies in the range of a few hundred kilohertz. They observed signals down to 520 kHz in their low-noise cosmic noise experiment. Dowden [1959] observed signals at 180 kHz which he attributed to the aurora. Jorgenson [1968], Morgan [1977], and Sato and Hayashi [1985] also reported the ground-level detection of high-latitude terrestrial signals with $f \lesssim 100$ kHz. The reviews of Ellyett [1969] and LaBelle [1989] discuss these observations and many others (including those at much higher frequencies).

Most scientific routine ground-level monitoring of the high-latitude radio environment has either been in the frequency range $f \lesssim 10$ kHz (in search of ducted whistler (W) mode waves) or $f \gtrsim 10$ MHz (as a by-product of riometer cosmic radio wave absorption experiments). Interest in the frequency range between these extremes was stimulated by the theoretical work of Wu et al. [1983] indicating that W mode waves propagating downward along the direction of the earth's magnetic field B can be generated in the low-altitude auroral region by the cyclotron maser instability. (By contrast, the Cerenkov process, considered to be the generating mechanism for W mode auroral hiss observed by spacecraft in the topside ionosphere [Swift and Kan, 1975; Maggs, 1976], generates W mode waves propagating oblique to B which have difficulty reaching the ground [Swift and Kan, 1975].) Waves propagating downward, along B, are only restricted by the constraint $f < f_H$, where f_H is the electron gyrofrequency (see Figure 13.1 of Booker [1984]), and the parallel propagation restriction can be relaxed when ionospheric irregularities and electron collisions are considered (see p. 27 of Helliwell [1965]). Benson et al. [1988] attempted to detect such waves and to determine the maximum emission frequency in an experiment performed in Alaska. (This frequency is an important parameter in the Wu et al. [1983] mechanism yielding information on the minimum altitude of wave generation because it is related to f_H). During the spring of 1986 they detected signals attributed to ionospheric origin on 7 out of 15 nights of operation. These signals were observed on one or more of the following frequencies: 150, 291, 500 or 700 kHz. A one-to-one correlation with geophysical phenomena was not always evident but the strongest emissions occurred near midnight magnetic local time (MLT). The results of an attempt to repeat this experiment in Norway will be presented in the next section.

91 9 4 161

91-09719

OBSERVATIONS

Four Realistic (Radio Shack) phase lock loop synthesized communication receivers with 10 m horizontal ground-level monopoles were used in Andenes, Norway to search for aurorally generated emissions with f ≥ 150 kHz (the low frequency limit of the receivers). The experiment was conducted in January, 1989 in conjunction with the flight of the Oedipus A rocket from the Andøya Rocket Range. This rocket consisted of a dual payload connected by a conducting tether. One of the payloads included a radio receiver covering the frequency range from 25 kHz to 5 MHz. This receiver was derived from the WISP (Waves in Space Plasmas) Space Shuttle program [James et al., 1987]. The goal of the ground-based radio observations was to relate topside and bottomside observations of auroral radio emissions with f ≥ 100 kHz. The equipment used was the same as in the Alaskan experiment of Benson et al. [1988] except that a 16 bit analog to digital converter and a personal computer replaced the multichannel chart recorder. Data were recorded at the rate of 20 samples/s; one of the 4 receivers was sampled in turn every 50 ms. Thus the time resolution on any single receiver was 200 ms.

A preliminary investigation of ambient noise levels was conducted on the island of Andøya in September, 1988. Three sites operated by the Andøya Rocket Range were tested and the science building at the rocket range was found to be the quietest in the evening hours due to the shielding of signals from European radio stations provided by the adjacent mountains. During the January, 1989 experimental campaign, however, interference from the Norwegian telemetry transmitter caused a problem. The ground-based receivers were operated during this period for approximately 93 hours when this transmitter was off the air. A scan of these data indicate that auroral radio noise of the type observed near Fairbanks during the spring of 1986 was less common at Andenes, Norway in January, 1989. The data were scanned for simultaneous enhancements detected on all 4 receivers (each tuned to a different frequency between 150 and 1500 kHz). A visual scan was made of plots based on long term averages of data, extending over time intervals ranging from 5 minutes to several hours, and a computer scan was performed on the individual data samples separated by 50 ms. In the later scan, the approach was to identify a number of consecutive occurrences of 4 consecutive enhancements (one for each receiver) above specified background levels. The background levels were determined by visual inspections of the unaveraged data at the beginning and end of the interval to be scanned. Confidence in the technique, and system operation in general, was obtained by the detection of signals radiated from the Andøya Rocket Range ionosonde for a few seconds at 29 and 59 minutes past each hour.

The Andøya ionosonde also provided an extra receiving system for searching for noise of ionospheric origin. Just as the ISIS topside sounder receivers have proven to be valuable in detecting natural signals such as auroral kilometric radiation [Benson and Calvert, 1979] and type III solar radio storms [James et al., 1990] in the topside ionosphere, conventional ground-based ionosondes can be used to detect noise of ionospheric origin reaching the ground. In fact, the only event atttibuted to ionospheric noise detected during January, 1989 on the island of Andøya was first identified from an inspection of the Andøya ionograms. Three consecutive ionograms from the time period of interest are presented in Figure 1. The ionogram recorded at 2229 UT in Figure 1 reveals a wideband speckled noise pattern covering the entire virtual height scale over the frequency range from 200 kHz to about 3.5 MHz, whereas the adjacent ionograms (which are tyical of those preceding and following 2229 UT) show this type of noise band only in the frequency range beyond the point of strong E and/or F layer echoes and with the greatest intensity up to about 10 or 12 MHz.

Signals were also detected during this time from 159 to 1200 kHz using the 4 ground-based receivers as shown in Figures 2, 3 and 4. The individual points in Figures 2 and 3 correspond to averages of 24 s and 2 s of raw data from each receiver, respectively. The raw data with 0.2 s resolution presented in Figure 4 indicate that the noise event persisted for about 3 min centered near 2227:30 UT.

The Andøya 32.5 MHz riometer also recorded a large signal enhancement that started just before 2227 UT and persisted for about 1 minute (see Figures 5 and 6).

The above event occurred during a time period of magnetic disturbance near midnight MLT (see Figure 7). The most intense events observed during the Alaskan experiment were also observed near midnight MLT. These results are consistent with satellite observations of Laaspere et al. [1971] of the diurnal variation of the invariant latitude of the maximum intensity of 200 kHz auroral hiss (see Figure 8).

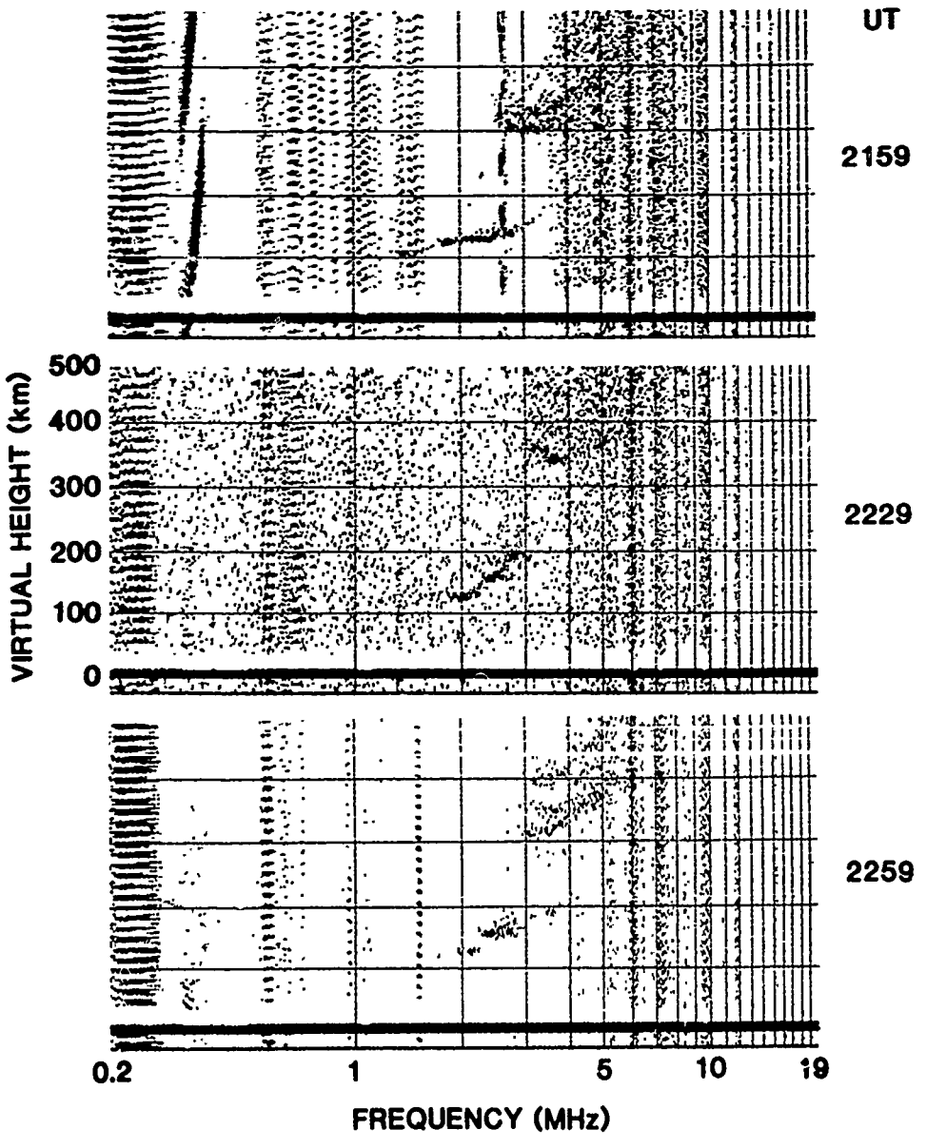


Figure 1. Three consecutive ionograms from the Andøya Rocket Range, Andenes, Norway recorded on 19 January 1989.

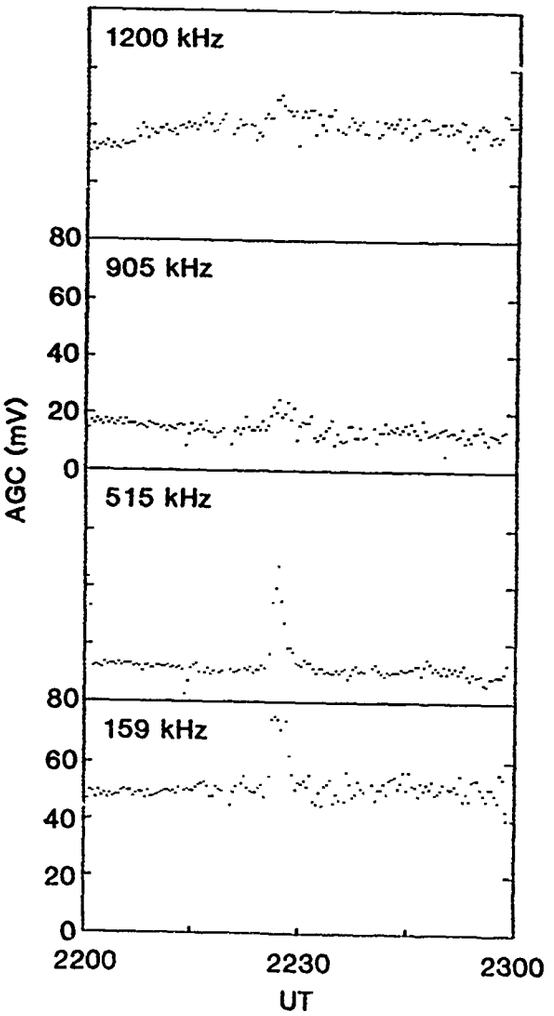


Figure 2. Automatic gain control voltage (AGC) from the ground-based receivers on 19 January 1989 at Andenes, Norway. Each point represents a 24 s average.

DISCUSSION

Benson et al. [1988] pointed out that the lowest W mode emission frequency from the cyclotron maser instability using the parameters of Wu et al. [1983] would be about 350 kHz. Thus this model will not explain the 159 kHz emissions observed in Norway. The work of Wu et al. [1983] was extended by Wu et al. [1989] to include the effect of the trapped auroral electrons in their wave generation mechanism. The altitude range over which emission (near the local f_H) can occur in this later work extends from approximately 0.3 to 2 R_E. Even this model, however, will not explain the emissions observed above the maximum ionospheric value for f_H, i.e., it cannot explain the ionosonde observations that extend at least to 3.5 MHz or the riometer observations at 32.5 MHz. Additional theoretical work is called for to explain emissions over such a broad frequency range. Additional observations are also required on a long term nearly continuous basis in order to better determine the geophysical conditions required for the emissions to be detected on the ground and to determine the exact frequency spectrum. Measurements of this type are underway in Alaska and are planned for Antarctica [J. LaBelle, personal communication, 1990].

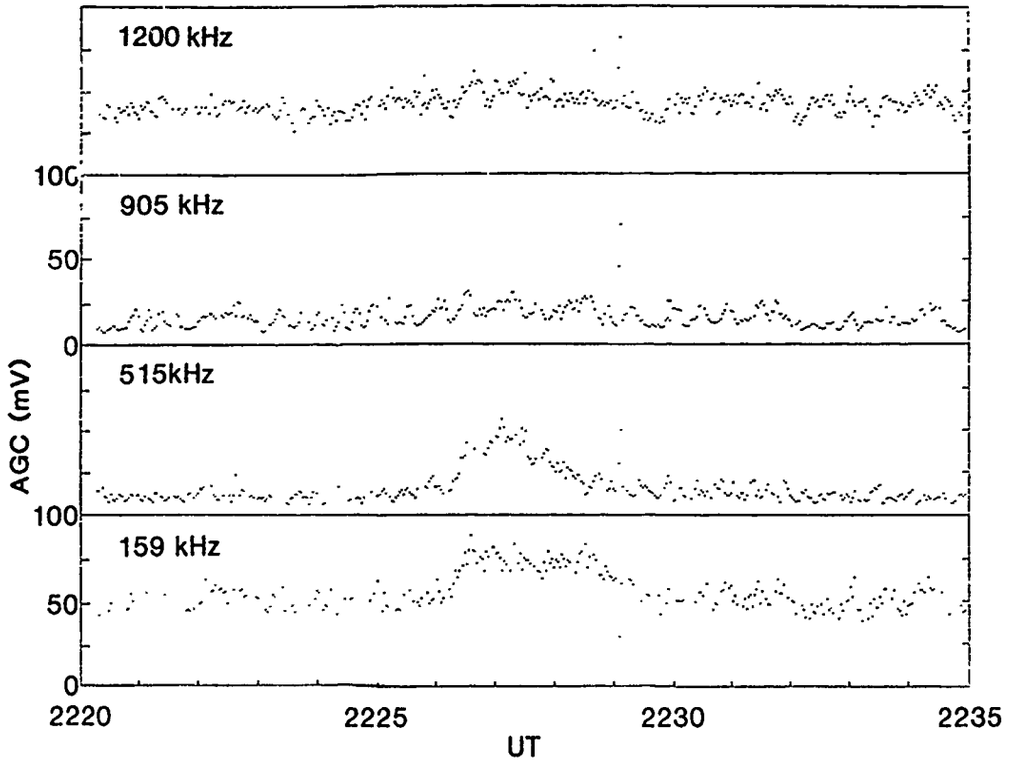


Figure 3. Same as Figure 2 except each point represents a 2 s average.

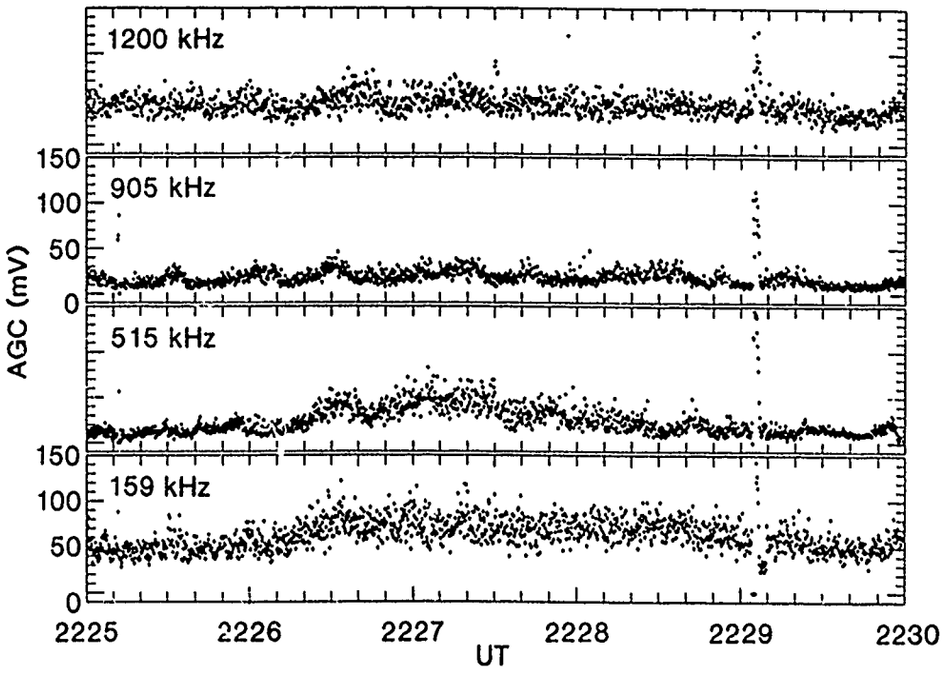


Figure 4. Same as Figure 2 except each point represents a raw data sample taken every 0.2 s on any one receiver.

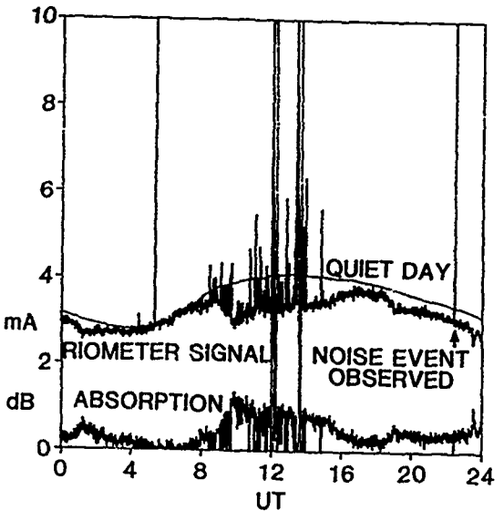


Figure 5. Andenes 32.5 MHz riometer record for 19 January 1989.

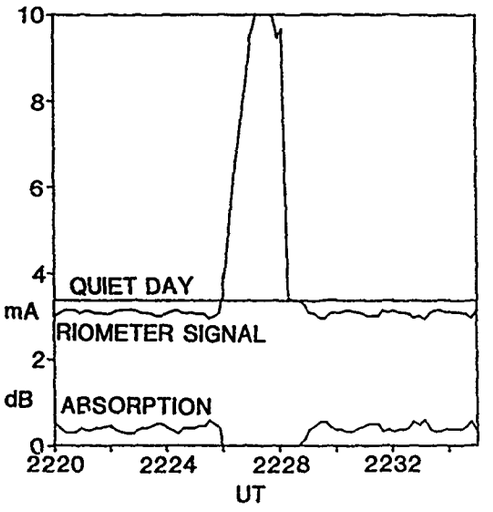


Figure 6. An expanded portion of Figure 5 centered on the noise event observed on 19 January 1989.

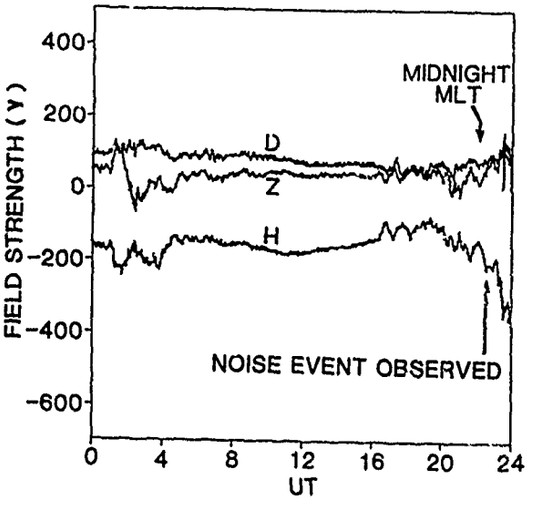


Figure 7. Andenes magnetometer record for 19 January 1989.

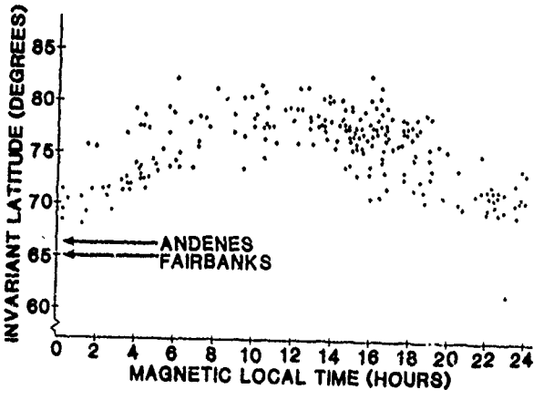


Figure 8. The location of 200 kHz auroral hiss maximum intensity as observed by OGO 6 between June 1969 and May 1970 when Kp ≤ 2 [after Laaspere et al., 1971] with Andenes and Fairbanks locations indicated.

Measurements of the noise environment have also been made during experiments directed toward developing techniques for identifying quiet communication channels [Borgmann and Lindner, 1989]. These techniques usually consist of sophisticated adaptive frequency change procedures to rapidly search the radio spectrum for interference-free channels. Their success depends on the character of the interference. A technique that can successfully find a quiet channel in the midst of many narrow-band interfering signals may not be able to function during an extended interval where the interference is due to a wideband source such as the high latitude noise reported in this paper. The experimental work of Borgmann and Lindner [1989] supports earlier observations that the noise levels are much higher at night than during the day. The unanswered question, however, is what percentage of this noise is due to naturally occurring wideband signals as opposed to man-made narrowband signals? There may be data available in the communication-oriented experiments that would help to answer the scientific question concerning the origin of the natural emissions; an answer to this scientific question would most certainly benefit communication efforts designed to elude interference.

ACKNOWLEDGMENTS

We are grateful to C. Moyer for engineering assistance and to J. Deering for engineering assistance, software development and field operations support.

REFERENCES

Benson, R. F., and W. Calvert, ISIS 1 observations at the source of auroral kilometric radiation, Geophys. Res. Lett., 6, 479-482, 1979.

Benson, R. F., M. D. Desch, R. D. Hunsucker, and G. J. Romick, Ground-level detection of low- and medium-frequency auroral radio emissions, J. Geophys. Res., 93, 277-283;2767, 1988.

Booker, H. G., Cold plasma waves, 349 pp., Martinus Nijhoff Publishers, Boston, Mass., 1984.

Borgmann, D., and J. Lindner, The characteristics of the HF radio channel and its influence on modern HF data communication system design, AGARD Conf. Proc., AGARD-CP-453, 52-1 - 52-12, 1989.

Dowden, R. L., Low-frequency (100 kc/s.) radio noise from the Aurora, Nature, suppl. 184, 803, 1959.

Ellyett, C. D., Radio noise of auroral origin, J. Atm. and Terr. Phys., 31, 671-682, 1969.

Helliwell, R. A., Whistlers and Related Ionospheric Phenomena, 349 pp., Stanford University Press, Stanford, Calif., 1965.

James, H. G., R. F. Benson, J. Fainburg, and R. G. Stone, The ionosphere as a focusing lens - a case study involving simultaneous type-III solar radio storm measurements from the ISIS-1 and -2 and ISEE-3 satellites, Radio Sci., in press, 1990.

James, H. G., T. R. Darlington, C. H. Hersom, R. S. Gruno, and J. V. Gore, Exploring space plasmas - The WISP/HF experiment, Proc. IEEE, 75, 218-235, 1987.

Jorgensen, T. S., Interpretation of auroral hiss measured on OGO 2 and at Byrd Station in terms of incoherent Cerenkov radiation, J. Geophys. Res., 73, 1055-1069, 1968.

Laaspere, T., W. C. Johnson, and L. C. Sempreton, Observations of auroral hiss, LHR noise, and other phenomena in the frequency range 20 Hz - 540 kHz on Ogo 6, J. Geophys. Res., 76, 4477-4493, 1971.

LaBelle, J., Radio noise of auroral origin: 1968-1988, J. Atm. and Terr. Phys., 51, 197-211, 1989.

Maggs, J. E., Coherent generation of VLF hiss, J. Geophys. Res., 81, 1707-1724, 1976.

Morgan, M. G., Wide-band observations of LF hiss at Frobisher Bay (L = 14.6), J. Geophys. Res., 82, 2377-2386, 1977.

Reber, G., and G. R. Ellis, Cosmic radio-frequency radiation near one megacycle, J. Geophys. Res., 61, 1-10, 1956.

Sato, N., and K. Hayashi, Band-limited ELF emission burst (auroral roar), J. Geophys. Res., 90, 3531-3535, 1985.

Swift, D. W., and J. R. Kan, A theory of auroral hiss and implications on the origin of auroral electrons, J. Geophys. Res., 80, 985-992, 1975.

Wu, C. S., D. Dillenburg, L. F. Ziebell, and H. P. Freund, Excitation of whistler waves by reflected auroral electrons, Planet. Space Sci., 31, 499-507, 1983.

Wu, C. S., P. H. Yoon, and H. P. Freund, A theory of electron cyclotron waves generated along auroral field lines observed by ground facilities, Geophys. Res. Lett., 16, 1461-1464, 1989.

AUTHOR INDEX

AARONS, J.,39
ANDERSON, D.N.,351,360,366,436
ARGO, P.E.,542
AUSTEN, J.R.,12
BAILEY, A.D.,329
BAIRD, F.G.,209
BAKER, K.B.,100
BASU, B.,69
BASU, S.,92,124,470
BASU, Su.,92,124
BEALL, J.H.,533
BELL, T.F.,515
BENCH, P.M.,329
BENSON, R.F.,561
BERNHARDT, P.A.,24
BISHOP, D.F.,273
BISHOP, G.J.,470
BROWN, L.D.,343
BUONSANTO, M.,460
CARLSON, R.,542
CLYNCH, J.R.,470
COCO, D.S.,470
COHEN, D.J.,322
COKER, C.,470
CONKRIGHT, R.,1
CORIATY, A.J.,329
COSTA, E.,92
COSTER, A.J.,460
DAEHLER, M.,418
DANA, R.A.,142
DANIELL, R.E., Jr.,343,351
DATLOWE, D.W.,82
DAVIES, K.,1,378
DECKER, D.T.,351,436
DESCH, M.D.,561
DETRICK, D.L.,76
DOHERTY, P.H.,343,360,366
DUNCAN, L.M.,24
FEDDER, J.A.,108
FERGUSON, J.A.,505
FITZGERALD, T.J.,542
FOSTER, J.C.,266,436
FOX, M.W.,343,388
FRANKE, S.J.,12
GAINES, E.E.,82
GAPOSCHKIN, E.M.,460
GERSON, N.,491
GOODING, D.J.,181
GOODMAN, J.M.,281
GORMAN, F.J.,401
GREENWALD, R.A.,100
GROVES, K.M.,266
HAGN, G.H.,189
HARGREAVES, J.K.,76
HOLMES, J.D.,481
HOUPIS, H.L.F.,237,247
HUBA, J.D.,108
HUNSUCKER, R.D.,46
IMHOF, W.L.,82
INAN, U.S.,515,523
JASPERSE, J.R.,69,351,436
JEFFREY, Z.R.,300
KATAN, J.R.,329
KELLER, C.M.,171
KELLY, F.J.,533
KERSLEY, L.,118
KESKINEN, M.J.,24,108,133
KLOBUCHAR, J.,12,343,360,366,470
KNEPP, D.L.,237,257
KOSSEY, P.A.,329
KULLSTAM, P.A.,152
LEE, M.C.,266,396,550
LI, S.W.,329
LIN, K.H.,427
LIU, C.H.,12
LIU, X.M.,378
LIVINGSTON, R.C.,124
McNAMARA, L.F.,300
MENDILLO, M.,388
MEYER, R.L.,481
MILLER, K.L.,442
MITCHELL, H.G.,108
MOBILIA, J.,82
MOODY, T.N.,257
MORRISON, J.M.,209
NICHOLS, H.E.,181
NICKISCH, L.J.,219,247
OSTERGAARD, J.C.,329
PAKULA, W.A.,360
PAUL, A.K.,410
POULSEN, W.L.,515
PRYSE, S.E.,118
RAND, T.W.,201
RASMUSSEN, J.E.,329
RAYMUND, T.D.,12
REILLY, M.H.,281,290
REINKING, J.T.,257
RETTERER, J.M.,436
RHOADS, F.J.,281
RICH, F.J.,436
RICHARDS, P.G.,442
ROBINSON, R.M.,82,452
ROESLER, D.P.,313
ROSENBERG, T.J.,76
ROWLAND, H.L.,24
RUOHONIEMI, J.M.,100
RUSSELL, C.D.,118
SAILORS, D.B.,160
SAMSON, J.C.,100
SATYANARAYANA, P.,108

AUTHOR INDEX

SCALES, W.A.,24
SCHLEGL, W.S.,427
SCHLEHER, J.S.,229
SCHUNK, R.W.,351
SIMONS, D.J.,542
SINGH, M.,281
SOICHER, H.,401
SOJKA, J.J.,351
SPALLA, P.,388
STALKER, J.,12
STEHLE, R.H.,189
TETENBAUM, D.,460
THORNTON, L.E.,460
TORR, D.G.,442
VONDRAK, R.R.,82
VOSS, H.D.,82
WARSHAW, S.,542
WEITZEN, J.A.,329
WHALEN, J.A.,60
WITTWER, L.A.,142
WOLCOTT, J.H.,542
WOOD, K.S.,533
YALOWITZ, J.S.,209
YANG, K.S.,427
YEH, K.C.,201,427
YOO, C.,396
ZALESAK, S.T.,108

☆ U. S. GOVERNMENT PRINTING OFFICE : 1990 O - 278-020